Psychological Bulletin
2009, Vol. 135, No. 3, 347–368

© 2009 American Psychological Association
0033-2909/09/$12.00 DOI: 10.1037/a0014211

Implicit Measures: A Normative Analysis and Review

Jan De Houwer
Ghent University

Sarah Teige-Mocigemba
University of Freiburg

Adriaan Spruyt and Agnes Moors
Ghent University

Implicit measures can be defined as outcomes of measurement procedures that are caused in an automatic manner by psychological attributes. To establish that a measurement outcome is an implicit measure, one should examine (a) whether the outcome is causally produced by the psychological attribute it was designed to measure, (b) the nature of the processes by which the attribute causes the outcome, and (c) whether these processes operate automatically. This normative analysis provides a heuristic framework for organizing past and future research on implicit measures. The authors illustrate the heuristic function of their framework by using it to review past research on the 2 implicit measures that are currently most popular: effects in implicit association tests and affective priming tasks.

Keywords: implicit measures, automaticity, IAT, affective priming

Most psychologists would argue that a full understanding of the behavior of an individual requires knowledge not only of the external situation in which the individual is present but also of the internal psychological attributes of the individual. Throughout the history of psychology, researchers have therefore attempted to measure interindividual differences in the psychological attributes of people (e.g., Anastasi, 1958; Eysenck & Eysenck, 1985; Mischel & Shoda, 1995). During the past decade, a major development in this research has been the introduction of so-called implicit measures. These measures were originally put forward mainly within the social psychology literature (e.g., Fazio, Jackson, Dunton, & Williams, 1995; Greenwald, McGhee, & Schwartz, 1998) but have since then spread to various other subdisciplines of psychology, including differential psychology (e.g., Asendorpf, Banse, & Mücke, 2002), clinical psychology (e.g., Gemar, Segal, Sagrati, & Kennedy, 2001), consumer psychology (e.g., Maison, Greenwald, & Bruin, 2004), and health psychology (e.g., Wiers, van Woerden, Smulders, & de Jong, 2002).

Despite the widespread use of implicit measures, the actual meaning of the term *implicit measure* is rarely defined. On the basis of the work of Borsboom (Borsboom, 2006; Borsboom, Mellenbergh, & van Heerden, 2004) and De Houwer (De Houwer, 2006; De Houwer & Moors, 2007; Moors & De Houwer, 2006), we first provide a normative analysis of the concept "implicit measure." The analysis is normative in the sense that it stipulates the properties that an ideal implicit measure should have. As such, the analysis provides a heuristic framework for reviewing and evaluating existing research on implicit measures. By examining the extent to which a particular implicit measure exhibits these normative properties, one can clarify the way in which the measure is an implicit measure and highlight those issues on which further research is required. In the second part of this article, we perform this exercise with regard to the two types of implicit measures that are currently most popular: effects in implicit association tests (IATs; Greenwald et al., 1998) and affective priming tasks (Fazio et al., 1995). Before we present and apply our normative analysis, we provide a brief description of these two measures.

During a typical IAT, participants see stimuli that belong to one of four categories and are asked to categorize each stimulus by pressing one of two keys. Two of the four categories are assigned to the first key, and the two other categories are assigned to the second key. The core idea underlying the IAT is that categorization performance should be a function of the degree to which categories that are assigned to the same key are associated in memory. Hence, by examining which combinations of categories result in the best categorization performance, one should be able to infer which categories are more closely associated in memory. Take the example of a racial IAT designed to measure attitudes toward Black and White individuals (e.g., Mitchell, Nosek, & Banaji, 2003; Monteith, Voils, & Ashburn-Nardo, 2001). There are four categories of stimuli: stimuli related to Black individuals (e.g., the face of a Black person), stimuli related to White individuals (e.g., the face of a White person), positive words (e.g., *summer*), and negative words (e.g., *cancer*). In the Black-positive task, participants press the first key whenever a Black face or a positive word appears and the second key whenever a White face or a negative word is presented. In the White-positive task, the first key is assigned to White faces and positive words and the second key is assigned to

Jan De Houwer, Adriaan Spruyt, and Agnes Moors, Department of Psychology, Ghent University, Ghent, Belgium; Sarah Teige-Mocigemba, Department of Psychology, University of Freiburg, Freiburg, Germany.

Portions of this article were presented by Jan De Houwer at the EAPP Summer School on Implicit Measures of Personality, 29 July 2007, Bertinoro, Italy, and at the GRK 1253 Summer School on Methods of Affective Neuroscience, 6 October 2007, Bronnbach, Germany. Preparation of this article was supported by Ghent University Grant BOF/GOA2006/001.

Correspondence concerning this article should be addressed to Jan De Houwer, Department of Psychology, Ghent University, Henri Dunantlaan 2, B-9000 Ghent, Belgium. E-mail: Jan.DeHouwer@UGent.be

Black faces and negative words. Participants who are faster on the Black-positive than on the White-positive task are assumed to have stronger associations in memory between the concepts "Black person" and "positive" than between the concepts "White person" and "positive" (or weaker associations between "Black person" and "negative" than between "White person" and "negative"). The reverse is assumed to be true for someone who is faster on the White-positive than on the Black-positive task. Given the additional assumption that racial attitudes are a function of the strength of the associations in memory between, on the one hand, the concepts "Black person" and "White person" and, on the other hand, the concepts "positive" and "negative," one can argue that the difference in performance on the Black-positive and White-positive tasks provides an index of the attitude toward Black persons relative to the attitude toward White persons.

In a typical affective priming task, participants categorize target stimuli as being positive or negative. Each target is preceded by a prime stimulus. The core idea underlying an affective priming measure is that one can estimate the attitude toward the prime stimulus by examining how the presence of the prime influences the affective categorization of the target stimuli. For instance, in order to measure attitudes toward Black and White persons, one can present on each trial the picture of a Black or a White person as a prime stimulus followed by a positive or a negative target word that participants categorize as being positive or negative in valence (e.g., Fazio et al., 1995). If Black faces facilitate responding to positive relative to negative target words, this effect would indicate a positive attitude toward Black persons. If Black faces facilitate the affective categorization of negative relative to positive words, this effect would suggest a negative attitude toward Black persons. The attitude toward White persons can be estimated in a similar manner by investigating whether White faces as primes facilitate responding to positive or negative targets.

Many other implicit measures have been proposed in recent years. As are IAT effects and affective priming effects, most of these implicit measures are based on performance in speeded reaction time tasks. That is, the psychological attributes of the individual are inferred from the speed or accuracy with which the individual responds to certain stimuli in certain tasks (e.g., De Houwer, 2003a; Nosek & Banaji, 2001; see De Houwer, 2003b, for a structural analysis and review). Other implicit measures, however, focus not on the speed of responding but on the content of responses (e.g., Payne, Cheng, Govorun, & Stewart 2005; Sekaquaptewa, Espinoza, Thompson, Vargas, & von Hippel, 2003; see De Houwer, 2008, for a discussion of dimensions on which implicit measures can differ). In the first part of this article, we try to determine what all these different measures have in common and what it means to say that something is an implicit measure.

Implicit Measures: A Normative Analysis

A normative analysis of the concept "implicit measure" implies the specification of a set of criteria that an ideal implicit measure should meet. Implicit measures are a subclass of all possible measures of psychological attributes. Hence, an ideal implicit measure should not only be an ideal measure but should have the additional characteristic of being implicit. We therefore first discuss the normative criteria that any perfect psychological measure should meet. Afterward, we specify the additional criterion that applies specifically to implicit measures. Because of their normative nature, the criteria that we specify in this section set very high standards for any type of measurement. It might well be that most psychological measures, implicit or otherwise, do not meet these standards. Even measures that are not perfect can, however, still be useful. The normative criteria should therefore not be interpreted as minimal conditions that must be met before a measurement outcome can be regarded as an implicit measure. Rather, they are ultimate goals. By specifying the characteristics of an ideal implicit measure, one can use the normative criteria to highlight the strengths and weaknesses of existing implicit measures and to provide direction for future research.

What Is a Measure?

Given that psychological measures are meant to reveal internal psychological attributes of individuals, an ideal psychological measure should provide an exact index of the extent to which an individual possesses the psychological attribute that the measure was designed to capture. For instance, an ideal measure of racial attitudes should reflect the extent to which a particular individual likes or dislikes particular racial groups. As pointed out by Borsboom et al. (2004, p. 1061), "a test is valid for measuring an attribute if and only if (a) the attribute exists and (b) variations in the attribute causally produce variations in the outcomes of the measurement procedure." Figure 1A provides a graphical representation of this statement. When a measurement procedure is applied to a certain person, a hypothetical attribute within the person causes an observable outcome (bottom arrow), which can then be used to make an inference about the attribute of the person (top arrow). This statement, though at first sight obvious, clarifies a number of important issues. We discuss these issues in detail and summarize the main conclusions at the end of this section.

First, a distinction can be made between a measure in the sense of a procedure and a measure in the sense of the outcome of a procedure (see also De Houwer, 2006). For instance, the racial IAT is a measure in the sense of a procedure. It is a set of objective guidelines about what someone should do in order to obtain an index of racial attitudes (e.g., what stimuli to present in what manner, what instructions to give, and how to register and analyze responses). The procedure generates an outcome, namely, a score

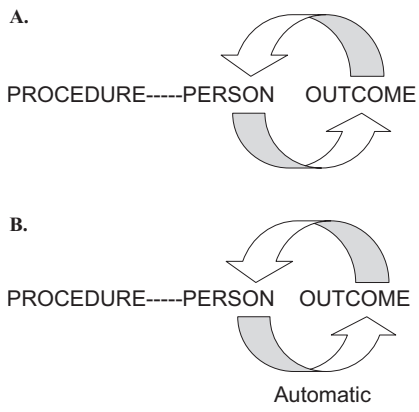


Figure 1. A schematic representation of the definition of (A) a measure and (B) an implicit measure.

reflecting the difference in performance on the two IAT tasks (e.g., the Black-positive task and the White-positive task). The outcome is a measure in the sense that it is meant to reflect racial attitudes. To avoid confusion, one should always clarify whether the term *measure* is used to refer to a procedure or to an outcome of a procedure. We use it to refer only to a measurement outcome and use the term *measurement procedure* to refer to a procedure used to generate a measurement outcome.

Second, the claim that a measurement outcome provides a valid measure of a psychological attribute implies the ontological claim that the attribute exists in some form and influences behavior. There has been a lot of debate in philosophy about whether it is possible to substantiate ontological claims. Borsboom et al. (2004) argued, however, that there simply is no alternative to making ontological claims when measuring. It does indeed seem illogical to argue that the statement "the outcome is a valid measure of attribute X" and the statement "attribute X does not exist" are both true. If the attribute does not exist or does not cause variation in the outcome, the outcome cannot be a measure of the attribute. If the outcome is a measure of the attribute, the attribute must exist and must causally influence the outcome (see Borsboom et al. for a more detailed critique of positivist and constructivist views on ontological claims regarding measurement).

Ontological assumptions might of course be incorrect or incomplete. To evaluate those assumptions, one can engage not only in conceptual analyses but also in empirical research using measures that are assumed to be somehow related to psychological attributes. As such, ontological assumptions can also depend on empirical measurement. Borsboom et al. (2004) nevertheless argued that ontological assumptions are primordial. There can be no measurement without ontological assumptions, but there can be ontological assumptions without measurement (e.g., those based solely on conceptual considerations). Moreover, as we discuss later, measurement allows for strong conclusions regarding psychological attributes only under very specific conditions.

Regardless of the ontological status that one assigns to assumptions about psychological attributes, it is important to realize that claims about the validity of a measure do imply assumptions about the nature of psychological attributes. In psychology, relatively little is known about the attributes that are assumed to underlie behavior. Psychological attributes, such as attitudes, stereotypes, and personality traits, cannot be observed directly. Instead, they are inferred indirectly from the observation that there are systematic differences in behavior that are not merely a function of differences in the external environment. Although most psychologists will agree that there are internal attributes that codetermine (human) behavior, there is less agreement about what these different attributes are and how they should be defined. For instance, a special issue of the journal *Social Cognition* (Gawronski, 2007) was recently devoted to the question "What is an attitude?" even though the concept "attitude" has been measured in numerous ways ever since Thurstone (1928). It is important to realize that the validity of a measure of psychological attributes can go only as far as the validity of the assumptions about the attributes it is assumed to measure. If these assumptions turn out to be incorrect or incomplete, the old interpretation of the measure is no longer valid and claims about the validity of the measure need to be abandoned or altered.

The statement of Borsboom et al. (2004) has a third important implication. It clarifies that validity implies causality. Variations in a measurement outcome can reveal something about variations in a psychological attribute only if the attribute somehow causes the outcome. To verify the validity of a measure, we thus need evidence that variations in the to-be-measured attribute indeed cause variations in the measurement outcome. The most suitable way to obtain such evidence is through experimental research (i.e., research in which the attribute is manipulated experimentally and the effects of the manipulation on the measurement outcome are examined). This research should reveal not only that variations in the attribute cause variations in the measurement outcome but also how they do so (see also Wentura & Rothermund, 2007). Knowledge about the causal mechanisms provides more certainty about the fact that an attribute causes an outcome and allows one to optimize the measure in the sense of maximizing the effects of the attribute on the measurement outcome.

Whereas Borsboom et al. (2004) promoted experimental studies as the primary way to study validity, until now, the correlational approach has dominated validation research. Borsboom et al. argued that the correlational approach as typically adopted in validation research is suboptimal for the study of the validity of measures because correlational evidence (a) does not allow for causal inferences and (b) is not directed at examining the relation between psychological attributes and measurement outcomes. With regard to the first point, there are many well-known reasons why correlations do not allow for causal conclusions. For instance, a correlation between two variables might be due not only to a causal relation between the two but to the presence of a third variable that causally influences both other variables. With regard to the second point, most correlational validation studies have been designed to examine how psychological constructs are related to each other rather than to measurement outcomes. Following the recommendations of Cronbach and Meehl (1955), researchers developed theories (so-called nomological networks) about whether a particular target attribute (e.g., intelligence) should or should not be related to other attributes (e.g., general knowledge). A measure of the target attribute was considered to be valid if it correlated in the expected way with measures of other attributes. It would lead us too far afield to discuss all the arguments that Borsboom et al. presented against this approach. For the present purpose, it is important to realize that correlational studies about the relations between (measures of) psychological constructs do have important limitations for the study of the validity of measures when validity is defined in terms of the causal impact of a psychological attribute on the measurement outcome.

This fact does not imply that correlational studies are worthless and that only experimental studies should be conducted from now on. First, correlational results can constrain hypotheses about the nature of the psychological attribute that causes variation in a measure. For instance, as the evidence increases that a measure correlates in the expected manner with measures of other attributes, it becomes less likely that these correlations are due to a hidden, third factor (see Nosek & Smyth, 2007). As pointed out by an anonymous reviewer, because correlational data are often much easier to obtain than experimental data, the correlational approach offers us an efficient way to learn more about the validity of a measure.

Second, experimental research also has limitations. Most prominently, experiments can provide conclusive information about the causal properties of psychological attributes only if the implemented manipulations (a) affect the to-be-measured attribute in the intended manner and (b) do not affect other attributes or processes that determine performance. When a manipulation does not impact on the attribute that is being measured, the absence of an effect of the manipulation on the measure says nothing about the validity of that measure. This fact implies that an experimental approach makes little sense for measures that capture attributes that are stable over time and impervious to situational factors. Also, when a manipulation influences psychological attributes and processes other than the intended ones, the presence of an effect on the measure provides little information about the validity of that measure. In such cases, it is not clear whether the observed effects are due to the fact that the measure captures the to-be-measured attribute or is determined by other attributes and processes. As is the case with correlational research, a third factor might thus be responsible for the observed relation between the independent and dependent variables.

The more we know about the determinants of psychological attributes and the processes by which attributes influence behavior, the more certain we can be that experimental manipulations will have only the intended effects. Hence, the merits of an experimental approach to examining the validity of measures depend on theoretical and conceptual knowledge. Despite the limitations of experimental research, Borsboom et al. (2004) convincingly argued that experimental research should be given a prominent place in validation research. Given that validation research has until now been dominated by correlational studies, this is an important conclusion regardless of whether one agrees with Borsboom et al.'s evaluation of correlational research.

In the previous paragraphs, we have argued that claims about the validity of a measure (a) refer to the properties of an outcome of a procedure rather than to the procedure itself, (b) imply ontological assumptions about the causal effect of psychological attributes and thus depend on the validity of these assumptions, and (c) can be examined not only with a correlational approach but with experimental studies. A final important point in the work of Borsboom et al. (2004) is that they distinguished between the validity of a measure and its overall quality. They argued that a valid measure is not necessarily reliable or predictive of criterion variables and could even measure different attributes in different groups of respondents (Borsboom et al., p. 1070). This is because a measure can be a valid index of a psychological attribute even if this attribute is not the only source of variation in the measure. Validity implies that the to-be-measured attribute causes variation in the measure but does not rule out the possibility that other attributes or situational factors are additional sources of variation. When variation in a measurement outcome has multiple sources, one can never be sure that a particular measurement outcome was caused by the to-be-measured attribute rather than by other sources of variance. Also, if the impact of the other sources of variation is time or context dependent, this will reduce reliability, predictive validity, and measurement invariance. In light of our aim to specify the normative criteria that an ideal (implicit) measure should meet, these criteria should take into account not only requirements of validity but the determinants of overall quality. In other words, we need to be sure not only that the to-be-measured psychological attribute causally influences the measurement outcome and how it does so; we should also know whether there are other sources of variation and how these sources impact on the measure.

To conclude, on the basis of the work of Borsboom and colleagues (Borsboom, 2006; Borsboom et al., 2004), we can now define a measure as a measurement outcome that is causally produced by the to-be-measured attribute. The overall quality of the measure also depends on whether there are other sources of variation. On the basis of these considerations, we argue that an ideal measure should conform to two normative criteria: It should be clear (a) which attributes causally produce the measurement outcome and (b) how these attributes causally produce the measurement outcome. We will refer to the first criterion as the "what criterion" and to the second criterion as the "how criterion" (also see De Houwer, Geldof, & De Bruycker, 2005). Both conceptual analyses and empirical research are required to verify whether a measure conforms to these criteria. There needs to be conceptual clarity about the attribute that is assumed to be measured. Not only correlational but also experimental studies can provide information about which attributes actually cause variation in the measurement outcome and how they do so (also see Wentura & Rothermund, 2007).

What Is an Implicit Measure?

The claim that a measurement outcome is an implicit measure implies not only that it is a measure (i.e., that it is causally produced by the to-be-measured attribute) but also that it is implicit in some sense. De Houwer (De Houwer, 2006; De Houwer & Moors, 2007) argued that the term *implicit* can best be understood as being synonymous with the term *automatic.* Both terms have been used to describe the features of psychological processes or, more precisely, the conditions under which psychological processes can be operative. For instance, a process can be called automatic in the sense that it can operate even when participants do not have particular goals, a substantial amount of cognitive resources, a substantial amount of time, or awareness (of the instigating stimulus, the process itself, or the outcome of the process; see Bargh, 1992; Moors & De Houwer, 2006). From this perspective, an implicit measure can be defined as a measurement outcome that is causally produced by the to-be-measured attribute in the absence of certain goals, awareness, substantial cognitive resources, or substantial time (De Houwer, 2006; De Houwer & Moors, 2007). This definition implies that the processes by which the attribute causes the measurement outcome are automatic in a certain sense of the word, an idea that is graphically represented in Figure 1B.

We have carefully avoided equating the concepts "implicit" and "automatic" with one particular feature or set of features. One reason for this is that the different features of automaticity do not always co-occur. For instance, evidence suggests that stereotype activation is automatic in that it does not depend on the conscious goal to activate the stereotype or on the presence of processing resources but is nonautomatic in that it depends on the presence of certain other goals (e.g., Bargh, 1992; Moskowitz, Salomon, & Taylor, 2000). Moreover, different processes can possess different features of automaticity and thus be automatic in different ways (Bargh, 1992). For these reasons, one cannot simply say that a

process is automatic. It is always necessary to specify in what sense a process can be considered automatic by specifying which automaticity features it possesses and which it does not possess. One could, of course, pick out one feature as being the defining feature. If this were done, it would mean that a process can be called automatic or implicit if it can be demonstrated that the process can operate under that specific condition. Although this approach is potentially useful, at present there is little agreement about what the central defining feature should be. Whereas some refer to a certain aspect of awareness (e.g., Greenwald & Banaji, 1995), others emphasize the lack of control (i.e., the lack of an impact of goals relating to the process; e.g., Fazio & Olson, 2003). In the absence of any convincing arguments to select one of the features as being central, the best solution seems to be to always specify the feature or features one is referring to when calling a process automatic or implicit.

It is important to point out that evidence regarding the implicitness of a measure does provide some information about the nature of the underlying psychological attribute (see also De Houwer, 2009). If the measurement outcome is causally influenced by a psychological attribute under a certain set of conditions (e.g., when proximal goals are absent), one can conclude that the psychological attribute can be activated and can influence behavior under those conditions. There is, however, not necessarily a one-to-one mapping between observed measurement effects and the properties of the underlying psychological attribute. When a measure is not causally influenced by the attribute under certain conditions, this could be because the attribute is not activated under those conditions or because processes by which the attribute influences the measure do not operate under those conditions (see also De Houwer, 2009). Likewise, if a variable influences the magnitude of the measurement outcome, this could be due to its effect on the to-be-measured attribute or to its effect on other processes that influence the magnitude of the outcome (see Gawronski, Deutsch, LeBel, & Peters, 2008, for a detailed discussion of this issue). Finally, the implicitness of a measure says little about how the underlying attribute is represented. For instance, it is difficult to determine whether different attributes underlie implicit and explicit measures or whether both measures reflect the same attributes under different conditions (see Nosek & Smyth, 2007; Payne, Burkley, & Stokes, 2008). Therefore, caution is needed when one draws conclusions about the properties of psychological attributes on the basis of empirical measurement results.

On the basis of these considerations, we can formulate a third normative criterion that should be met before a measure can be called an implicit measure: The to-be-measured attribute should cause the measurement outcome automatically. This implicitness criterion implies that (a) it is clearly specified which automaticity features one is referring to and (b) there is empirical evidence to back up the claim that the measure possesses those automaticity features (see De Houwer, 2006; De Houwer & Moors, 2007).

Implicit Measures: A Review

Now that we know the normative criteria to which an implicit measure should conform, we can examine whether each implicit measure that has been or will be proposed meets these criteria. Thus, our analysis provides a heuristic framework not only for past research but for future studies. In this section of the article, we review past research on implicit measures and highlight the questions that need to be addressed in future studies. Because most existing studies have focused on IAT and affective priming effects, we limit our review to these two measures. For each of these two measures, we evaluate whether and in what way they meet the three normative criteria of implicit measures: the what criterion, the how criterion, and the implicitness criterion. We do not aspire to describe or even refer to each individual IAT and affective priming study that has been conducted in the past (see Klauer & Musch, 2003; Lane, Banaji, Nosek, & Greenwald, 2007; and Nosek, Banaji, & Greenwald, 2007, for more extensive reviews). Instead, we summarize the main conclusions that can be reached on the basis of previous research and relate them to the normative criteria that we put forward in this article. For each conclusion, we refer to only a subset of the relevant evidence or, when available, to papers that provide a review of the literature relevant for that conclusion.

IAT Effects

The What Criterion: What Attributes Cause Variations in IAT Effects?

Associations in memory. In their seminal paper, Greenwald et al. (1998) argued that IAT effects (i.e., the difference in performance on the two tasks of an IAT procedure) reflect associations between concepts (hence the name Implicit Association Test). Although it is not entirely clear what Greenwald et al. meant by the term *association* (see the exchange between Greenwald, Nosek, Banaji, & Klauer, 2005, and Rothermund, Wentura, & De Houwer, 2005), they did suggest that psychological attributes, such as attitudes and stereotypes, are represented as associations in memory (e.g., Greenwald et al., 2002). Hence, the hypothesis that IAT effects reflect associations implies the hypothesis that variations in IAT effects can be caused by psychological attributes such as attitudes and stereotypes. Researchers have adopted three approaches in examining whether IAT effects do capture attitudes and stereotypes: experimental, semiexperimental, and correlational. In this section, we present a brief overview of these three lines of research.

The hypothesis that IAT effects are caused by the to-be-measured attributes can be examined by experimentally manipulating those attributes and examining whether the manipulations influence the IAT effects in the expected manner. Relatively few studies have adopted this approach. Perhaps the strongest evidence for the validity of IAT effects as a measure of attitudes comes from studies in which novel attitudes were created by pairing previously unknown stimuli with other, clearly positive or negative stimuli (e.g., Olson & Fazio, 2001). The results showed that IAT effects reflected these new attitudes, even when participants were unaware of the fact that the attitudes resulted from the stimulus pairings. These results are particularly convincing, because it is difficult to see how the observed effects could have been caused by attributes or processes other than the newly created attitudes.

Other experimental studies focused on the malleability of preexisting attitudes and stereotypes. The results of these studies showed that IAT effects are sensitive to manipulations of variables such as the experimental context and instructions (for a review, see Blair, 2002; Gawronski & Bodenhausen, 2006). For instance, IAT

measures of racial attitudes indicate that White participants are less prejudiced against Black persons (a) when interacting with a Black experimenter than when interacting with a White experimenter (e.g., Lowery, Hardin, & Sinclair, 2001); (b) after seeing movie clips of Black individuals in a positive compared with a negative situational context (e.g., Wittenbrink, Judd, & Park, 2001); or (c) after seeing pictures of admired Black individuals and disliked White individuals compared with pictures of disliked Blacks and admired Whites (e.g., Dasgupta & Greenwald, 2001). In studies on gender stereotypes, Blair, Ma, and Lenton (2001) showed that gender stereotypes as measured by IAT effects were less pronounced following counterstereotypic mental imagery but were stronger following stereotypic mental imagery. Note, however, that there is disagreement about whether these malleability effects provide evidence for the validity of IAT scores. In some cases, the effects on the IAT scores could have been due not to changes in the to-be-measured attribute but to changes in the extent to which the attribute caused variations in the IAT score (De Houwer, 2009; Gawronski et al., 2008). For instance, the presence of a Black experimenter could lead to an increase in the extent to which participants try to control the outcome of the IAT (see Blair, 2002). In future research, recently proposed componential accounts, such as diffusion model analysis (Klauer, Voss, Schmitz, & Teige-Mocigemba, 2007) or the quad model (Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005), might help to identify the processes that underlie malleability effects.

Most claims about the validity of IAT effects are based on semiexperimental and correlational data. The most straightforward semiexperimental way of examining the validity of (implicit) measures is by looking at whether variations in the type of stimuli influence the measures in a meaningful manner. For instance, on the basis of normative studies and a priori arguments, one can postulate that most people have more favorable attitudes toward flowers than toward insects. In line with the hypothesis that IAT effects can register attitudes, Greenwald et al. (1998) showed that an IAT designed to measure the attitudes toward flowers and insects indeed revealed more positive attitudes toward flowers than toward insects. Another popular semiexperimental approach for testing the validity of (implicit) measures is the so-called known-group approach (e.g., Banse, Seise, & Zerbes, 2001). It involves variations in the type of participants whose reactions are measured. For instance, one can argue on a priori grounds that White and Black individuals should differ in their racial attitudes.

In support of the validity of the racial IAT, studies have confirmed that White and Black individuals indeed show different racial IAT effects (Nosek, Banaji, & Greenwald, 2002). Although results such as these suggest that attributes such as attitudes can cause variations in IAT effects, their conclusiveness is limited by the semiexperimental design on which these studies are based. When one divides stimuli or participants into groups on the basis of one particular feature (e.g., valence or group membership), it is difficult to exclude the possibility that the groups differ also with regard to other, correlated features. Hence, one cannot be entirely confident that observed differences between the groups are due to the feature that the experimenter used to create the groups. Note, however, that the risk of unrecognized confounds also exists in fully experimental studies. Furthermore, the risk of confounds in semiexperimental studies can be reduced by carefully controlling for plausible correlated features.

A final set of studies used a correlational approach. We can divide correlational studies in two sets on the basis of the type of criterion variable that was used. The first set of studies focused on predictive validity, in that the magnitude of IAT effects was used as a predictor of a particular behavior thought to be indicative of the to-be-measured attribute. The second set of studies assessed convergent validity by examining the relation between IAT effects and other measures of the to-be-measured attribute. Recent meta-analyses (Greenwald, Poehlman, Uhlmann, & Banaji, in press; Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005; also see Lane et al., 2007, and Nosek et al., 2007, for reviews) showed that IAT effects do tend to correlate in a meaningful manner with such criterion variables. The results of numerous correlational studies are thus in line with the idea that IAT effects can capture attitudes and stereotypes.

Because of the evidence reviewed above, it is now generally accepted that variations in IAT effects are at least sometimes and at least to some extent caused by the attitudes or stereotypes that they were designed to measure. In those cases in which IAT effects can be interpreted as indices of attitudes and stereotypes, there are restrictions on the manner in which they should be interpreted. Most important, IAT effects at best allow for only relative conclusions, because they are determined by at least two attitudes or stereotypes. For instance, a racial Black–White IAT effect is determined not only by attitudes toward Black persons but also by attitudes toward White persons (Greenwald & Nosek, 2001). Blanton and Jaccard (2006) have argued that IAT scores are also relative in that it is impossible to interpret the absolute value and sign of an IAT score. An IAT effect does not reveal whether an individual has positive or negative attitudes as such (e.g., whether Person A likes White people more than Black people). This is so because it is not clear what psychological reality corresponds to a zero score on an IAT. For instance, even if racial attitudes are a direct cause of scores on a racial IAT, it is not certain that a zero score on the racial IAT means that the person likes Black and White individuals to the same extent. Because of this fact, the IAT effect shown by a particular individual can be interpreted only by comparing it with the IAT effects of other persons (e.g., Person A has more positive attitudes toward White persons or less positive attitudes toward Black persons than does Person B; but see Greenwald, Nosek, & Sriram, 2006, for a critique of Blanton & Jaccard, 2006).

There are at least two procedural factors that can bias IAT scores and thus complicate the interpretation of the absolute value and sign of an IAT score. First, there is evidence suggesting that IAT effects are determined not only by the attitudes and stereotypes concerning the categories (e.g., "Black person," "White person"; De Houwer, 2001) but also by the individual stimuli used to represent the categories (e.g., a particular Black or White face; Bluemke & Friese, 2006; Govan & Williams, 2004; Mitchell et al., 2003). Because it is often unclear which stimuli are most suitable for measuring a particular attitude or stereotype, it is difficult to correct for the impact of this factor. Second, IAT scores are known to depend on the order in which the two tasks of the IAT are presented. Although measures can be taken to reduce order effects (see Nosek, Greenwald, & Banaji, 2005), it is difficult to determine the magnitude of the order effect for a particular individual and thus to correct the IAT score for these order effects.

In addition, one should be aware that, as is possible with virtually all measures of psychological attributes, variations in IAT effects might be caused by attributes other than attitudes or stereotypes. As we discuss earlier, such variation could limit the overall quality of the IAT as a measure of attitudes and stereotypes. If IAT effects can be caused by different kinds of attributes or situational factors and if it is not clear which effects are caused by which factor, the meaning of the effects becomes ambiguous (see Fiedler, Messner, & Bluemke, 2006). In the following sections, we examine which other attributes might cause variations in IAT effects.

Extrapersonal knowledge. It has been argued that variations in IAT effects can be caused by extrapersonal knowledge (i.e., knowledge that the individual has but regards as irrelevant for his or her own responses to objects; see Gawronski, Peters, & LeBel, 2008, for a conceptual analysis). For instance, assume that the racial IAT effect of a particular individual suggests that this individual has a more negative attitude toward Black persons than toward White persons. Researchers such as Karpinski and Hilton (2001) and Olson and Fazio (2004) have argued that this effect does not necessarily reflect the personal attitudes of the individual but rather the fact that the individual possesses knowledge of societal views about Black and White persons. In Western societies, Black persons tend to be regarded in a less favorable manner than are White persons. Even individuals who claim that these societal views do not correspond with their personal views (e.g., Black individuals) might show racial IAT effects suggestive of pro-White attitudes simply because they possess knowledge of the pro-White societal views.

The hypothesis that IAT effects can be caused by extrapersonal, societal views has been supported by experiments in which the manipulation of extrapersonal views led to changes in IAT effects (e.g., Han, Olson, & Fazio, 2006; Karpinski & Hilton, 2001). Further support came from semiexperimental studies. When groups with diverging personal and societal views were tested, IAT effects at least sometimes seemed to be in line with societal views (see data on racial attitudes in Black persons and attitudes toward unhealthy but tasty foods such as candy; Olson & Fazio, 2004; Spruyt, Hermans, De Houwer, Vandekerckhove, & Eelen, 2007). Also, when the IAT procedure is changed in such a way that it should be less susceptible to the impact of societal views (i.e., by removing error feedback and using more personalized category labels), the evidence for the causal role of societal views becomes weaker (Olson & Fazio, 2004).

On the other hand, doubts have been raised about the theoretical significance and validity of the extrapersonal account of IAT effects. At the conceptual level, it has been argued that the distinction between personal and societal views actually makes little sense, especially when one considers the automatic effects of these views (Banaji, 2001; Gawronski & Bodenhausen, 2006; Gawronski et al., 2008; Nosek & Hansen, 2008a). At the empirical level, correlational studies provided little evidence for a link between IAT effects and measures of societal views (Nosek & Hansen, 2008a). Furthermore, Nosek and Hansen (2008b) recently obtained evidence that raises serious doubts about the correct interpretation of the experimental and semiexperimental data that were regarded as evidence for the extrapersonal account (Han et al., 2006; Olson & Fazio, 2004). Few arguments remain to support the claim that IAT effects are causally influenced by extrapersonal views.

Salience. A second alternative account was proposed by Rothermund and Wentura (2004), who argued that IAT effects are caused by salience asymmetries. The basic idea is that performance during an IAT will be fast when the categories assigned to the same key are similar with regard to their salience. Salience can be defined as the degree to which a stimulus pops out within a background of other stimuli. For instance, in a racial IAT, one can argue that Black faces and negative words are more salient for White participants than are White faces and positive words. Hence, White participants should be faster when Black faces and negative words are assigned to the same key (as is the case in the White-positive task) than when Black faces and positive words are assigned to the same key (as is the case in the Black-positive task). Rothermund and Wentura reported experimental data (i.e., manipulations of salience influence IAT effects), semiexperimental data (e.g., stimuli that differ only in salience lead to IAT effects), and correlational data (i.e., IAT effects are related to measures of salience) in support of their hypothesis.

As was acknowledged by Greenwald et al. (2005), it is thus beyond any doubt that at least some IAT effects are caused by salience asymmetries (see Kinoshita & Peek-O'Leary, 2006, and Houben & Wiers, 2006, for more recent evidence). There is still disagreement, however, about the pervasiveness of the impact of salience asymmetries (see Rothermund et al., 2005). Also, at the conceptual level, there is still uncertainty about how salience should be measured (e.g., Greenwald et al., 2005) and how it is related to other attributes, such as familiarity and polarity (e.g., Kinoshita & Peek-O'Leary, 2005, 2006; Proctor & Cho, 2006).

Similarity. On the basis of the findings of Lasaga and Garner (1983), De Houwer et al. (2005) put forward the hypothesis that IAT effects can be caused by similarity at the perceptual level. Even before the introduction of the IAT, Lasaga and Garner demonstrated that the task of categorizing four stimuli by pressing one of two keys is easier when perceptually similar stimuli are assigned to the same key than when perceptually dissimilar stimuli are assigned to the same key. Conceptually replicating this study, De Houwer et al. used four categories, each of which comprised several stimuli rather than four individual stimuli, and found the same effects (see also Mierke & Klauer, 2003, Experiment 1). Because there is no reason to assume that the stimuli used by Lasaga and Garner and by De Houwer et al. differed in any systematic manner other than with regard to their perceptual features, these studies do strongly suggest that similarity at the perceptual level can cause IAT effects.

In most IAT procedures other than those of Lasaga and Garner (1983) and De Houwer et al. (2005), stimuli are selected in such a way that the stimuli of the different categories do not differ perceptually. Hence, in all likelihood, perceptual similarity does not play a major role in most IAT studies. Nevertheless, the fact that perceptual similarity can cause IAT effects led De Houwer et al. to formulate the hypothesis that all IAT effects are caused by some type of similarity. Stimuli and categories can be similar or dissimilar not only with regard to their perceptual features but with regard to other features, such as their (affective) meaning and their salience. From this perspective, attitudes and stereotypes can cause IAT effects because they are a form of semantic similarity (regardless of whether they are personally endorsed). Likewise, salience asymmetries can drive IAT effects because they imply similarity with regard to salience. The similarity hypothesis put

forward by De Houwer et al. therefore encompasses all previously discussed hypotheses and is compatible with the fact that multiple attributes can cause IAT effects. Which attributes actually cause the IAT effect should depend on the types of similarity that are most salient in a given situation (see Medin, Goldstone, & Gentner, 1993). On the negative side, because similarity is unconstrained (everything is similar to everything else in some respect), the similarity account runs the risk of being unfalsifiable (but see De Houwer et al., 2005, for a response to this criticism).

Cognitive abilities. Finally, correlational studies suggest that IAT effects can be influenced by psychological attributes related to the general cognitive abilities of the individual. First, McFarland and Crouch (2002) observed a correlation between overall response speed and the magnitude of effects on a variety of IAT tasks. If it is assumed that overall response speed is determined to a large extent by the cognitive abilities of the participant, the results of McFarland and Crouch suggest that IAT effects are determined at least in part by the cognitive ability of the participant. Second, IAT effects tend to increase in magnitude when the age of the participants increases (e.g., Hummert, Garstka, O'Brien, Greenwald, & Mellott, 2002). Because cognitive abilities tend to decline with age, this finding also suggests that IAT effects are determined by general cognitive abilities (see Sherman et al., 2008, for evidence supporting this conclusion). Finally, several studies (Back, Schmukle, & Egloff, 2005; McFarland & Crouch, 2002; Mierke & Klauer, 2003) showed that effects in IAT tasks correlate even if the tasks are supposed to capture attributes that should not be correlated. This finding can be explained if it is assumed that a general factor, such as cognitive ability, influences all IAT effects, regardless of the attributes they were designed to measure (but see Klauer et al., 2007). It is important to note, though, that the impact of cognitive abilities depends on how the IAT effects are calculated. The correlations described above seem to be strongest when the difference in the mean reaction time on the two tasks of an IAT is taken as the IAT effect but almost disappear when this difference is standardized (i.e., when mean difference is divided by the standard deviation of all reaction times; see Back et al., 2005; Cai, Sriram, Greenwald, & McFarland, 2004; Mierke & Klauer, 2003, for evidence, and Greenwald, Nosek, & Banaji, 2003, for a recommended way of calculating IAT effects).

Summary. The available evidence supports the hypothesis that IAT effects at least sometimes and to some extent measure the attributes that they are supposed to measure (i.e., associations in memory such as those that underlie attitudes and stereotypes). There is also evidence that they can reflect other attributes, such as salience, perceptual similarity, and general cognitive skills. The fact that IAT effects can be caused by different types of attributes does complicate the interpretation of these effects (e.g., Fiedler et al., 2006) and thus the overall quality of the measure (Borsboom et al., 2004). In part, this problem can be solved by learning more about the conditions under which the various kinds of attributes influence IAT effects. An important challenge for future research is therefore to uncover the variables that determine when a particular kind of attribute causes variation in IAT effects. We already know that the way in which IAT effects are calculated determines the impact of general cognitive abilities (e.g., Mierke & Klauer, 2003). Future studies on this topic might find inspiration in the similarity account put forward by De Houwer et al. (2005). According to this account, IAT effects will reflect whatever types of similarity are most salient for an individual in a certain context. Now that we have discussed various attributes that could cause variations in IAT effects, we turn to the question of how attributes can cause variations in IAT effects.

The How Criterion: By Which Processes Do Attributes Cause Variations in IAT Effects?

Random walk model. Brendl, Markman, and Messner (2001) introduced an informal (i.e., not mathematically formalized) random walk model in which responding during the IAT is a function of (a) the rate at which evidence is accumulated for a particular response and (b) the response criterion or threshold that the accumulated evidence must reach before a response can be emitted. In this model, IAT effects can thus be due to factors that influence the rate of evidence accumulation or the setting of the response criterion. We explain and evaluate these options consecutively.

Let us return to the example of a racial IAT. For a participant who likes White persons, a White face not only belongs to the category White but also is positively valenced. In the White-positive task, both sources of evidence (i.e., the positive valence as well as membership in the category White) move the accumulation process toward the correct response (i.e., the common response for White faces and positive words). In contrast, in the Black-positive task, the two sources of evidence move the accumulation process in opposite directions, because White faces and positive words are now mapped on different responses. For this reason, the net evidence accumulation rate for White faces should be lower in the Black-positive task than in the White-positive task, and this fact should lead to slower responses in the former task than in the latter.

The empirical evidence regarding the role of evidence accumulation is mixed. On the one hand, diffusion model analyses performed by Klauer et al. (2007) provided support for the hypothesis that a process akin to evidence accumulation is one of the sources of IAT effects. On the other hand, it is unlikely that evidence accumulation operates according to the principles described by Brendl et al. (2001). In their random walk model, evidence accumulation cannot lead to IAT effects for stimuli that are related to only one of the categories, because such stimuli can influence evidence accumulation in only one way. For instance, in a racial IAT, positive words (e.g., *flower*) and negative words (e.g., *cancer*) are typically unrelated to racial groups. Hence, evidence accumulation for these stimuli should be determined only by their valence and should be the same in the White-positive task and the Black-positive task. Contrary to this prediction, many studies, including those reported by Brendl et al. (2001), have shown that IAT effects can be observed for stimuli that are related to only one category (see also Klauer et al., 2007).

In the random walk model of Brendl et al. (2001), IAT effects could also result from the fact that participants adhere to different response criteria in the different tasks of an IAT. For instance, if participants for some reason set the response criterion higher in the Black-positive task than in the White-positive task of the IAT, this would lead to slower responses in the former than in the latter task and thus to a racial IAT effect. One problem with this account is that it does not specify what determines the level of the response criterion that participants choose. Brendl et al. argued that participants would raise the response criterion if they believed or perceived that a task is difficult but did not specify how they would

arrive at this belief or perception. Is it because of associations in memory, extrapersonal knowledge, salience asymmetries, or (perceptual) similarity? At the empirical level, the available evidence suggests that shifts in the response criterion are at best only one source of IAT effects.

In an unpublished study, De Houwer (2000) added two categories (numbers and nonwords) to a standard flower–insect IAT (see Greenwald et al., 1998). If IAT effects are due mainly to a change in response criterion, the effect of the response assignments for flower and insect items (flowers assigned to same key as positive words or insects assigned to same key as positive words) should be as big for the items of the additional categories as for the flower and insect items themselves. Results showed, however, that the IAT effect for the additional categories was only marginally significant and was significantly smaller than that for the other items. Finally, the diffusion model analyses reported by Klauer et al. (2007) identified shifts in the response criterion as one of several processes underlying IAT effects. In the next paragraphs, we consider a number of additional processes.

Response activation account. De Houwer (2001, 2003b) pointed out that there are structural similarities between stimulus–response compatibility tasks, such as the well-known Stroop task (see MacLeod, 1991, for a review) and the IAT task. In both tasks, stimulus and response features are compatible on some trials and incompatible on other trials. Take the example of the racial IAT. If participants are asked to press a first key for positive words and a second key for negative words, the keys become associated with positive and negative valence, respectively. In other words, pressing the first key becomes a positive response (equivalent to saying "good") and pressing the second key becomes a negative response (equivalent to saying "bad"; see Eder & Rothermund, 2008, for evidence supporting this assumption). Hence, for participants who like White persons but dislike Black persons, stimuli and responses are compatible (in the sense of associated with the same valence) when they are asked to press the first (positive) key for White faces and to press the second (negative) key for Black faces (as is the case in the White-positive task). When the same participants are asked to press the first (positive) key for Black faces and the second (negative) key for White faces (as is the case in the Black-positive task), the stimuli and responses are always incompatible. From research on stimulus–response compatibility effects, we know that performance is better when stimuli and responses are compatible than when they are incompatible. There is strong evidence that such effects are due to processes at the response selection stage whereby elements of the stimulus activate the incorrect (in case of incompatible combinations) or correct (in case of compatible combinations) response alternative. Because stimulus–response compatibility varies between the different blocks of an IAT, De Houwer (2001, 2003b) put forward the hypothesis that IAT effects are due to the activation of responses by (relevant or irrelevant features of) the presented stimuli.

Unfortunately, there have been few if any direct tests of this hypothesis. De Houwer (2001) examined whether IAT effects reflect the properties of the individual stimuli or the categories to which those stimuli belong but did not test whether these effects were due to processes at the response selection stage. Nevertheless, the hypothesis could be tested by using strategies that have been applied to demonstrate the role of response selection processes in other stimulus–response compatibility tasks (e.g., negative priming; see Wentura, 1999). Also, componential approaches, such as the quad model (Conrey et al., 2005), could help isolate the impact of response activation.

As is the case with the random walk processes discussed in the previous section, several attributes could cause variations in IAT effects by means of the response activation mechanism. De Houwer et al. (2005) argued that the response activation account fits very well with the idea that IAT effects are driven by different kinds of similarity. In fact, the concept "compatibility" can be regarded as synonymous with the concept "similarity." Hence, it can be argued that stimuli activate responses to which they are similar in a certain respect. Therefore, all attributes that can be regarded as a particular type of similarity (e.g., with regard to meaning, salience, or perceptual form) can cause IAT effects as the result of the response activation mechanism. General cognitive abilities also could have an effect, because they determine how much impact the activated responses have on actual performance.

Differential task switching model. During an IAT, participants are instructed to pay attention to two stimulus dimensions in order to categorize the stimuli. For instance, in a racial IAT, they are asked to respond to faces on the basis of the racial group (Black or White) and to words on the basis of valence (positive or negative). Because faces and words are presented in alternating order, participants constantly need to switch between the tasks of responding to the racial features of faces and responding to the valence of words. Research on task switching has shown that performance deteriorates as the result of switching between tasks (e.g., Meiran, Chorev, & Sapir, 2000).

Mierke and Klauer (2001, 2003) pointed out that the need to switch between different tasks depends on which categories are assigned to the same response. Again, take the example of the racial IAT. When participants who like White persons and dislike Black persons are asked to press the first key for White faces and positive words and the second key for Black faces and negative words (White-positive task), they can capitalize on response synergy and simply respond to both faces and words on the basis of whether they like the presented face or word. Because there is a perfect confound between the valence of the faces and the racial category of the faces, responding to a face on the basis of its valence or on the basis of its racial group leads to the same response. In contrast, when the same individuals are to press the first key for Black faces and positive words and the second key for White faces and negative words (Black-positive task), they must pay attention to the racial feature of the faces, because responding on the basis of the faces' valence would lead to incorrect responses. In the Black-positive task, accurate responding therefore requires task switching. Because task switching leads to performance costs (e.g., Rogers & Monsell, 1995), performance will be less good in the Black-positive task than in the White-positive task.

Klauer and colleagues (Klauer & Mierke, 2005; Mierke & Klauer, 2001, 2003; see also Klauer et al., 2007) provided strong evidence in support of the task switching model of IAT effects. First, participants who are generally good in switching between tasks should generally be less affected by whether the response assignments force them to switch between tasks. Hence, regardless of the attribute that an IAT is supposed to measure, these participants should reveal smaller IAT effects than do participants who are poor in switching between tasks. In support of this idea, Mierke and Klauer (2001, 2003) found that effects on different IATs are correlated even when

those IATs were designed to measure different attributes that should not be correlated (e.g., political attitudes and attitudes toward flowers and insects; see Klauer et al., 2007).

More direct evidence comes from sequential analyses of performance during the IAT tasks. Because switching between tasks is associated with performance costs, reaction times on any given trial should be longer when another dimension was relevant on the previous trial (switch trials) than when the same dimension was relevant (repetition trials). These differences in reaction times are called task switching costs. In a racial IAT, for instance, responses on a trial with a face stimulus should be slower when it is preceded by a trial with a word than when it is preceded by a trial with another face. If the need for task switching depends on which categories are assigned to the same key, the task switching costs should be a function of the category–response assignments. This is exactly what Mierke and Klauer (2001, 2003) observed. To extrapolate their findings to a racial IAT that is completed by participants who like White persons and dislike Black persons, one would expect task switching costs to be smaller in the White-positive task than in the Black-positive task.

In another set of studies, Klauer and Mierke (2005) found aftereffects indicative of active task switching during the IAT. Let us again take the example of the racial IAT. When participants who like White persons but dislike Black persons complete the Black-positive task of the racial IAT, they should pay attention to the valence of the words but should avoid paying attention to the valence of the faces. This is so because categorizing stimuli according to valence leads to the correct response only for words and to the incorrect response for faces. On the basis of earlier findings, Klauer and Mierke predicted that the repeated act of avoiding the evaluation of stimuli should carry over to a subsequent task in which the same stimuli had to be evaluated as being good or bad. In line with this prediction, they found that participants evaluated stimuli (e.g., Black and White faces) more slowly following an IAT task in which valence of those stimuli had to be ignored (e.g., the Black-positive task for participants who like White persons more than Black persons) than after an IAT task in which the valence of those stimuli could be used to categorize stimuli fast and correctly (e.g., the White-positive task for those persons).

Although these data provide strong evidence for the hypothesis that IAT effects at least in part are due to differential task switching costs, it remains unclear to what extent the differential costs result from a conscious strategy or from automatic processes. In principle, participants may consciously decide to recode certain tasks in the IAT (see De Houwer, 2003a; Rothermund & Wentura, 2004; Wentura & Rothermund, 2007). In the case of the racial IAT, for example, such recoding would involve a conscious intention to categorize both faces and words on the basis of valence when one realizes that such a strategy results in correct (and fast) responses (e.g., in the White-positive task for people who like White persons and dislike Black persons). Such a strategic recoding would imply that IAT performance is driven to some extent by the consciously intentional evaluation of the stimuli. This implication would raise doubts about whether IAT effects actually provide an implicit measure of attitudes (see below).

Strategic recoding might be prevented in two ways. First, the IAT's block structure can be eliminated (as recently proposed by Rothermund, Teige-Mocigemba, Gast, & Wentura, in press;Teige-Mocigemba, Klauer, & Rothermund, 2008). In the so-called single block IAT, the assignment of the categories to the responses can change from trial to trial rather than remain fixed during an entire block of trials. Given that recoding processes rely on a consistent assignment of categories to response keys over trials (Strayer & Kramer, 1994), Teige-Mocigemba et al. (2008) hypothesized that such a change should impede any kind of strategic recoding and indeed found evidence for this assumption.

Another way of preventing strategic recoding is by avoiding a perfect confound between stimulus features. For instance, De Houwer (2001) presented names of British and non-British (foreign) persons to British participants. It is important that half of the British and half of the foreign persons were liked by the participants (e.g., Princess Diana, Mahatma Gandhi), whereas the other persons were disliked (Margaret Thatcher, Adolf Hitler). It was unlikely that participants would intentionally decide to respond on the basis of the valence of the names rather than their category (British or foreign), because in half of the cases this approach would have led to an incorrect response. In most IATs, however, there is a perfect confound between valence and category membership (e.g., for a particular person, in a racial IAT, all White faces will be more positive than all Black faces). Therefore, one should be aware that participants usually can strategically recode standard IAT tasks.

Regardless of the exact nature of the processes that underlie differential task switching costs in the IAT, these processes could be responsible for the impact of a variety of attributes on IAT performance. As pointed out by Mierke and Klauer (2001, 2003), participants (intentionally or unintentionally) exploit similarities between stimuli in an attempt to facilitate task switching in certain blocks of an IAT. These similarities could be related not only to attitudes or other associations in memory but to salience or perceptual features of the items. Because task switching depends heavily on the cognitive abilities of the participant, interindividual differences in these abilities also should have an important impact on IAT effects.

Summary. Several proposals have been put forward about the processes by which attributes can cause variations in IAT effects. Nevertheless, compared with the number of studies on the relation between IAT effects and criterion variables (see Greenwald et al., in press, and Hofmann et al., 2005, for reviews), relatively little research has directly examined the role of each of these processes. The available evidence provides the strongest support for the involvement of task switching processes, but the exact nature of these processes still needs to be determined. Moreover, task switching appears to be just one of the mechanisms that produce IAT effects (see Klauer et al., 2007). Hence, there is a clear need for more research on how IAT effects come about. This research also can help clarify which attributes influence IAT performance under which conditions. Note, however, that the what and how criteria do not overlap completely, because one attribute could exert an effect through various processes and one process could support the effect of various attributes (see De Houwer et al., 2005).

The Implicitness Criterion: In What Sense Do IAT Effects Provide an Implicit Measure of Attributes?

Above, we argue that the implicitness of a measure refers to the conditions under which a psychological attribute causes variations in the measure (and thus the conditions under which the measure reflects the psychological attribute). A measure can be called an implicit measure of a psychological attribute if it is caused by that

attribute even under conditions that are typically associated with automatic processes. In line with Moors and De Houwer (2006; see also De Houwer & Moors, 2007), we focus on conditions involving the presence of proximal and distal goals, awareness, processing resources, and time.

The presence of proximal goals. A proximal goal is a goal related to the process under study. Proximal goals include the goal to engage in, stop, alter, or avoid the operation of a process. Hence, a process can be automatic in that it operates independently of the proximal goal to engage in, stop, alter, or avoid the operation of that process. Processes that operate under those conditions can be called unintentional (in the case of the goal to engage in), uncontrolled (with regard to the goal to stop, alter, or avoid), or autonomous (when such processes are independent of all proximal goals; see De Houwer & Moors, 2007; Moors & De Houwer, 2006). In the case of implicit measures, the processes under study are those by which an attribute of the person causes variations in the measure. Hence, the question of whether IAT effects are implicit in the sense of unintentional, uncontrolled, or autonomous boils down to the question of whether the processes by which the to-be-measured psychological attribute causes IAT effects operate independently of the goal to engage in, stop, alter, or avoid these processes. In other words, does the attribute still cause IAT effects (i.e., is the measure still valid) even when the participants (a) do not have the goal to express the attribute in IAT effects, (b) have the goal to stop the expression of the attribute in IAT effects, (c) have the goal to alter the way in which the attribute is expressed in IAT effects, or (d) have the goal to avoid the expression of the attribute in IAT effects?

To the best of our knowledge, only the last two issues have been addressed in research. Whether IAT effects depend on the conscious goal to alter or avoid the expression of an attribute has been examined in studies on faking. The results of these studies have been mixed. Some showed that IAT effects were largely unaffected by instructions to fake a certain attitude (e.g., Asendorpf et al., 2002; Banse et al., 2001; Egloff & Schmukle, 2002; Kim, 2003), whereas others suggested that participants can intentionally influence IAT effects (e.g., De Houwer, Beckers, & Moors, 2007; Fiedler & Bluemke, 2005; Steffens, 2004). The extent to which IAT effects can be consciously controlled seems to depend on a variety of variables, such as how much experience the participants have with the IAT (e.g., Fiedler & Bluemke, 2005; Steffens, 2004; see Czellar, 2006; De Houwer et al., 2007; and Schnabel, Banse, & Asendorpf, 2006, for other moderating variables). Hence, the available evidence does not allow for the strong conclusion that IAT effects are implicit in the sense of being always independent of the goals to avoid or alter the expression of the to-be-measured attribute. Nevertheless, it does seem to be the case that IAT effects are more difficult to control than are most traditional (questionnaire) measures (e.g., Steffens, 2004). In this sense, IAT effects can be described as less controllable and thus more implicit than many other measures. Also, the fact that IAT effects can be controlled when participants are encouraged to do so does not imply that participants do try to control IAT effects when they do not receive instructions to do so.

The presence of distal goals. Distal goals are goals other than those related to the process under study. A process can be called goal independent when its operation does not depend on any (proximal or distal) goal. It should be clear that it is difficult if not impossible to demonstrate that a process is entirely goal independent. The best one can do is demonstrate that the process does not depend on particular (distal) goals and make those goals explicit when describing the process as goal independent. Possible distal goals that could be relevant for IAT effects are the goal to respond quickly to stimuli and the goal to make few errors. Apart from preliminary data of Popa-Roch (2008) showing that response-time-based IAT effects decrease in magnitude when the goal to avoid errors is removed, we do not know of any studies that examined whether IAT effects depend on the presence of distal goals.

The presence of awareness. Although the term *implicit* is often seen as being virtually synonymous with the term *unaware* (e.g., Greenwald & Banaji, 1995), it is rarely made explicit what it means to say that a measure is unaware. It is important to realize that describing a measure as unaware can mean several things (Bargh, 1992; De Houwer & Moors, 2007). It could point to the fact that the to-be-measured attribute causes the IAT effect even when participants are unaware of (a) the stimuli that activate the attribute (e.g., the attitude object that is presented during the task); (b) the origins of the attribute itself (e.g., the fact that participants possess a certain attitude or how they acquired the attitude); (c) the fact that the attribute influences performance (e.g., that the outcome reflects a certain attitude); or (d) the manner in which the attribute influences performance (e.g., that certain category–response assignments lead to better performance than do other category–response assignments).

Can IAT effects actually be unaware in one or more of these four ways? First, IAT effects are obtained by instructing participants to categorize the relevant stimuli in certain ways. Therefore, participants must be aware of the four categories and the stimuli that are presented as instances of these categories. Second, there is some evidence that IAT effects can register attitudes even when participants do not know the origin of those attitudes. As mentioned above, Olson and Fazio (2001) created new attitudes by pairing neutral objects with liked or disliked objects and found that the IAT could register these attitudes even though participants were not aware of how the attitudes were created. Note that the participants could be made aware of the attitudes themselves because they could express these attitudes when asked to do so. Hence, the studies of Olson and Fazio do not demonstrate that IAT effects can capture unaware attitudes in the sense of attitudes that participants do not know they possess. Another observation that might be relevant in this context is that participants are sometimes poor in predicting their IAT performance and express surprise when informed about the meaning of their score on certain IATs (e.g., Mitchell et al., 2003; Nosek, Greenwald, & Banaji, 2007). However, it is unclear whether this observation means that the IAT picks up attitudes of which the participants are unaware or whether the IAT effect reflects other attributes such as extrapersonal knowledge or salience asymmetries. Therefore, at present, there is no strong evidence to support the conclusion that IAT effects can register attributes of which participants are unaware.

There is one published study that is relevant for the third and fourth ways in which IAT effects can be considered as unaware. Monteith et al. (2001) interviewed White participants about their experiences with a racial IAT. Up to 64% of the participants noticed that they were faster in the White-positive task than in the Black-positive task. Of the participants who noticed that they were faster in the White-positive task, 37% attributed this slower per-

formance to the fact that they apparently had a more negative attitude toward Black persons than toward White persons. These findings were confirmed in two recent unpublished studies showing that more than 80% of the participants who took part in a racial IAT could correctly describe the aim of the IAT (De Houwer & Moors, 2006; Popa-Roch, 2008, p. 118). De Houwer and Moors moreover found that the percentage of participants who were aware of the aim of the IAT was twice as large for a racial IAT (80%) as for an IAT designed to measure attitudes toward political parties (40%). Together, these results strongly suggest that a substantial part of the participants are aware of what IATs are supposed to measure and have a basic understanding of how IAT effects measure attributes. Hence, IAT effects typically do not seem to be unaware in this sense.

Many issues remain to be examined. For instance, it is not clear why the percentage of participants who are aware of the aim of an IAT depends on the categories featured in the IAT (De Houwer & Moors, 2006). It is also unclear whether awareness of the purpose of an IAT affects the magnitude or predictive validity of the IAT effects.

The presence of processing resources. An important feature of automaticity (and thus of implicitness) is whether a process can operate even when processing resources are scarce. This feature is examined most often by asking participants to perform a primary task that depends on the process under study while they perform a secondary task that deploys the available processing resources to a certain extent. A process is said to be efficient when the degree of load imposed by the secondary task does not impact on performance on the primary task (Moors & De Houwer, 2006). We know of only two studies in which the effect of mental load on IAT effects was examined. Devine, Plant, Amodio, Harmon-Jones, and Vance (2002, Study 3) failed to find an effect of a secondary task on IAT effects. In an unpublished study, Schmitz, Teige, Voss, and Klauer (2005) found that an increase in working memory load led to an increase in the magnitude of IAT effects but did not influence external correlations with self-reported attitudes. Hence, these initial results suggest that the translation of individual attitudes in IAT scores is efficient. However, more research is needed before firm conclusions can be drawn.

The availability of time. Moors and De Houwer (2006) pointed out that the minimal time needed for a process to operate is a central feature in the concept of automaticity both in its own right and because it can determine several other features. For instance, processes that require very little time to run to completion are most often difficult to control. In extreme cases, the process might occur so quickly that participants cannot become aware of the process or its input. With regard to the IAT, the impact of time on the validity of IAT effects could be examined by limiting the time that participants have available for responding to each stimulus. As far as we know, such studies have yet to be conducted.

Summary. All in all, there is relatively little research about the claim that IAT effects provide a measure of psychological attributes that can be qualified as implicit. Although participants seem to have less control over the IAT effects than over many other, more traditional measures, several studies indicate that IAT effects can at least sometimes and to a certain extent be controlled in a conscious manner. There is evidence showing that IAT effects are unaware in that they can capture attitudes whose origins are unknown, but other studies have demonstrated that participants are aware of the fact that the IAT aims to capture the to-be-measured attribute (e.g., racial attitudes) and how it does so (e.g., the difference in performance on the White-positive and Black-positive tasks of a racial IAT). Our review indicates that the question of whether IAT effects are actually implicit in some sense of the word has largely been neglected in past research. Only the impact of the goals to avoid or alter the expression of attributes has been examined in some detail in studies on faking. Other features of automaticity (and thus of implicitness) have not been addressed at all or have been examined in only a handful of studies. It should be noted that the fact that IAT effects can predict variance in criterion variables that cannot be explained on the basis of traditional (explicit) measures (e.g., see Asendorpf et al., 2002; Hofmann, Rauch, & Gawronski, 2007) does not provide evidence for the implicitness of the effects. It is not clear whether this incremental predictive validity is due to the implicit nature of the IAT effects or to the many other differences between IAT effects and traditional measures.

Affective Priming Effects

The What Criterion: What Attributes Cause Variations in Affective Priming Effects?

Attitudes. It is generally assumed that affective priming effects reflect the attitudes that participants have toward the object represented by the prime stimuli. For instance, attitudes toward Black persons can be estimated by examining the extent to which stimuli representing Black persons (e.g., photographs of the faces of Black persons or names typical of Black persons) facilitate responding to positive versus negative targets.[1] Whereas the relevant categories are made explicit in IAT studies, in affective priming studies, the categories that the prime stimuli are meant to instantiate are typically not made explicit in the instructions. Studies by Olson and Fazio (2003; see also De Houwer, 2001, 2003a) suggest that, because of this, affective priming effects are determined primarily by the attitudes toward the individual stimuli rather than by the attitude toward the category of which they are exemplars. The impact of the category, however, can be amplified by directing attention to the category (Olson & Fazio, 2003).

The hypothesis that affective priming effects can be caused by attitudes is supported by experimental, semiexperimental, and correlational studies. Many studies have confirmed that affective priming effects can pick up novel attitudes that have been created by pairing neutral stimuli with other, liked or disliked, stimuli (e.g., De Houwer, Hermans, & Eelen, 1998; see Hermans, Baeyens, & Eelen, 2003, for a review), even when participants do not appear to be aware of how the attitudes were acquired (Olson & Fazio, 2002). Experiments on the malleability of affective priming effects have shown that these effects can be influenced by a range of variables, such as the nature of the experimental context and

[1] Priming procedures have been used to examine attributes other than attitudes (e.g., Wittenbrink, Judd, & Park, 1997; see Wittenbrink, 2007, for a review). However, in these procedures, the targets differ not with regard to their affective meaning but with regard to nonaffective, semantic features. For instance, to examine the stereotype that women are more likely to study art than math, one can present faces of women as primes and ask participants to decide whether a target word refers to art or math.

instructions (see Blair, 2002, for a review). As with studies on the malleability of IAT effects, however, these results provide evidence for validity only if it can be demonstrated that the results are caused by changes in the to-be-measured attitudes.

Many semiexperimental studies have shown that stimuli to which participants should have different attitudes indeed evoke different affective priming effects (see Fazio, 2001, and Klauer & Musch, 2003, for a review). On the other hand, there are few affective priming studies in which the semiexperimental known-group approach was adopted. One of these is a study of Otten and Wentura (1999) in which an affective priming measure of attitudes toward groups revealed that participants preferred the group to which they were (randomly) assigned.

Finally, affective priming effects have been found to correlate in an expected manner with several kinds of criterion variables, such as real-life behaviors (e.g., Fazio et al., 1995; Spalding & Hardin, 1999) and other measures of the attitudes under study (e.g., Degner, Wentura, Gniewosz, & Noack, 2007; Dunton & Fazio, 1997; Frings & Wentura, 2003; Spruyt, Hermans, De Houwer, Vandekerckhove, et al., 2007; Wentura, Kulfanek, & Greve, 2005). It should be noted, however, that correlations between affective priming effects and criterion variables are sometimes small or even absent (e.g., Banse, 1999, 2001; Bosson, Swann, & Pennebaker, 2000). In part, this fact seems to be related to the on-average-limited reliability of affective priming scores. That is, repeated administrations (split half or test–retest) of the same affective priming measure tend to correlate only to a limited extent or do not correlate at all (e.g., Banse, 1999, 2001; Bosson et al., 2000; but see Cunningham, Preacher, & Banaji, 2001). This low reliability could in part be due to the fact that the relevant category (i.e., the attitude object that is being examined) is typically not made explicit (see Olson & Fazio, 2003; De Houwer, 2009). There is accordingly little control over whether or how participants process and categorize the prime stimuli, and this lack of control probably results in a large amount of error variance.

Other attributes. Very few studies have examined whether attributes other than attitudes can cause variations in affective priming effects. There is some evidence that affective priming effects are less sensitive to extrapersonal knowledge than are IAT effects. For instance, Han et al. (2006) showed that an experimental manipulation of extrapersonal knowledge had an effect on a traditional IAT measure but did not have one on an affective priming measure or on a personalized IAT measure that was designed to minimize the impact of extrapersonal views. More indirect evidence comes from the observation that fewer (White and Black) participants appear to prefer White persons over Black persons when racial attitudes are assessed by affective priming effects rather than by a standard racial IAT (Olson & Fazio, 2004; see also Spruyt, Hermans, De Houwer, Vandekerckhove, et al., 2007), but this finding could be related to the lower reliability of the priming measure.

As far as we know, there is little if any evidence regarding the impact of salience, similarity, or cognitive abilities on affective priming effects. There is some evidence to suggest that affective priming effects become stronger when the salience of the primes increases (e.g., Klauer, Mierke, & Musch, 2003). Also, it has long been known that priming effects in general (i.e., differences in responding to targets as a function of the nature of primes) can be driven not only by the evaluative features of the primes and targets (as is the case in affective priming effects) but by nonevaluative features, such as semantic meaning (i.e., semantic priming; Lucas, 2000), co-occurrence associations (i.e., associative priming; Ratcliff, 1988), and even perceptual similarity (e.g., Pecher, Zeelenberg, & Raaijmakers, 1998). Although we do not know any study that has examined priming on the basis of the similarity between the salience of the prime and the target, it seems reasonable to assume that salience-based priming effects can be observed.

It is important to note the fact that priming effects can be based on a large variety of attributes does not threaten the claim that affective priming effects can be based on attitudes toward the primes. In many affective priming studies, a large variety of primes and targets was used, so that it is unlikely that the evaluative features of the stimuli were confounded with other, nonevaluative features. Also, affective priming effects have been observed even in studies that controlled for a large variety of nonevaluative features (i.e., semantic meaning, associative links, perceptual similarity, familiarity; see Hermans, Smeesters, De Houwer, & Eelen, 2002). However, when one uses affective priming as a tool for assessing real-life attitudes, there is often less opportunity to control for nonevaluative features of the primes and targets. For instance, when affective priming effects are used to measure racial attitudes, it is possible that, at least for some individuals, Black faces and negative words are more similar than are Black faces and positive words not only with regard to their valence but also with regard to their salience. Hence, it is possible that affective priming effects for Black faces (e.g., faster responses to negative words preceded by a Black face than to positive words preceded by a Black face) do not reflect negative attitudes toward Black persons but the fact that Black persons are more salient for the participant. It is surprising that such risks to the validity of affective priming effects as a measure of real-life attitudes are rarely acknowledged and have not yet been studied.

Summary. The claim that affective priming effects can capture attitudes is supported mainly by the results of experimental and semiexperimental studies with stimuli that evoke different attitudes. Evidence from known-group and correlational studies is somewhat limited. One should keep in mind that priming effects can be based not only on evaluative features of the stimuli but on a range of other features that might sometimes be confounded with evaluative features. This possibility poses a risk to the validity of affective priming effects as a measure of attitudes and should receive more attention in future research.

The How Criterion: By Which Processes Do Attitudes Cause Variations in Affective Priming Effects?

Spreading of activation. The first account of affective (and other) priming effects was formulated in terms of activation spreading through a semantic network (Collins & Loftus, 1975; Collins & Quillian, 1969). In the network, each concept is represented by a node. If two concepts are somehow similar in meaning (e.g., if they share a valence), the nodes representing these concepts are linked by an association through which activation can spread. Hence, if a prime stimulus is presented, this will activate not only the corresponding concept node but all other nodes with which it is connected. Assuming that the speed of responding to a target stimulus depends on the activation level of the concept node representing the target, a prime stimulus that is affectively related

to the target stimulus could speed up responding to the target by preactivating the concept representation of the target in memory (see Fazio, 2001, 2007). This spreading-of-activation account of priming effects has dominated thinking about priming so much that the term *priming* is often used to refer not to the priming effect (i.e., faster responses when targets are presented in the context of a related prime) but to the process of preactivating representations in memory as the result of spreading of activation.

Despite the popularity of this account, research suggests that spreading of activation plays at best only a minor role in the production of affective priming effects. Most important, a spreading-of-activation account leads to the prediction that primes should facilitate not only the evaluation of affectively related targets (i.e., responses based on the valence of the targets) but the processing of other (semantic) features of the targets. For instance, if a prime preactivates the concept node of an affectively related target, this preactivation should reduce the time needed to determine the semantic category of the target (e.g., animal or object).

Several studies have failed to confirm this prediction. For instance, De Houwer, Hermans, Rothermund, and Wentura (2002) failed to find affective priming of semantic categorization responses (i.e., does the target refer to an object or a person) but did find strong affective priming of evaluation responses (i.e., is the target positive or negative), even though the same stimuli were presented in the same way in both tasks. More recent studies did find affective priming of semantic categorization responses and other nonevaluative responses (e.g., naming, lexical decision) but only under certain conditions (e.g., De Houwer, Hermans, & Spruyt, 2001; Spruyt, De Houwer, Hermans, & Eelen, 2007; Spruyt, Hermans, De Houwer, & Eelen, 2002; Wentura, 2000). Nevertheless, the consensus remains that processes akin to spreading of activation play little or no role in standard affective priming tasks (i.e., tasks in which participants are asked to evaluate the targets; see Klauer & Musch, 2003, for a more detailed review of the evidence supporting this conclusion).

Response activation. The available evidence strongly supports the hypothesis that affective priming effects in the evaluation task (i.e., is the target positive or negative) are due to the fact that the prime stimuli activate responses on the basis of their valence. Consider trials on which a positive target (e.g., the word *healthy*) is presented. Because the target is positive, participants need to give a positive response (e.g., say "good"). When the target is preceded by a positive prime (e.g., a White face for a person who likes White individuals), the positive valence of the prime will induce a tendency to give a positive response (e.g., say "good") and will thereby facilitate the selection of the positive response that needs to be given to the target. When the prime is negative (e.g., a Black face for someone who dislikes Black individuals) and the target is positive, the prime will induce a tendency to give a negative response and will thereby slow the selection of the correct (positive) response. The response activation account of affective priming thus implies that the prime influences the response selection process, whereas a spreading-of-activation account implies that the prime influences the processing of the target itself.

Many studies have found evidence for the assumption that affective priming effects arise at the response selection stage (see Klauer & Musch, 2003, for an extensive review, and Klauer, Musch, & Eder, 2005, for a more recent discussion). For instance, the important finding that affective priming effects occur in the evaluation task but not in a semantic categorization task (e.g., De Houwer et al., 2002; see above) is compatible with the fact that the valence of the positive and negative primes can induce a tendency to give positive and negative responses but not a tendency to give semantic categorization responses. Other strong evidence comes from Wentura (1999), who observed very specific aftereffects of affective priming trials with an incompatible prime and target. Responses on the trial after such an incompatible trial were slower when the valence of the correct response corresponded to the valence of the incompatible prime on the previous trial. This negative priming effect can be explained in the following manner: When the prime and target differ in valence, the incorrect response that is activated by the prime needs to be inhibited before the correct response can be selected. This inhibition carries over to the next trial and makes it harder to emit the previously inhibited response.

What implications does the response activation account have for hypotheses about the kinds of attributes that cause variations in affective priming effects? As we have discussed in the context of the response activation account of IAT effects, it is generally assumed that stimuli activate those responses to which they are similar in some respect (e.g., Kornblum & Lee, 1995). Hence, the response activation account of affective priming is compatible with the observation that priming effects can be induced by similarity not only with regard to valence but with regard to nonevaluative features, such as semantic meaning and salience. Because the impact of response conflicts on performance depends on the cognitive abilities of the participants to deal with the conflicts (e.g., Kane & Engle, 2003), one can predict on the basis of the response activation account that the cognitive abilities of participants will determine the magnitude of affective priming effects (see Klauer & Teige-Mocigemba, 2007, for evidence related to this prediction).

Summary. The available evidence allows for the conclusion that standard affective priming effects (i.e., those observed in tasks in which participants are asked to evaluate the targets) are due mainly to response activation processes. Priming effects by means of this mechanism can be caused not only by attitudes but by other attributes, including semantic meaning, salience, and cognitive abilities.

The Implicitness Criterion: In What Sense Do Affective Priming Effects Provide an Implicit Measure of Attributes?

The presence of proximal goals. In one of the early studies on affective priming, Hermans, De Houwer, and Eelen (1994, Experiment 1) observed significant affective priming effects even though participants were instructed to ignore the prime stimuli. This result suggested that the effects can occur in the presence of the goal to avoid an impact of the primes on performance (see also Klauer & Musch, 2003). In more recent studies, Teige-Mocigemba and Klauer (2008; also see Klauer & Teige-Mocigemba, 2007) did find evidence that participants can consciously control affective priming effects. In some conditions, participants were promised an extra monetary reward for fast and accurate responses to targets following specific primes. In other conditions, participants were explicitly instructed to fake certain attitudes. The affective priming

effects that were targeted by these instructions were found to be eliminated. The findings are remarkable, because the stimulus onset asynchrony (SOA) between prime and target was short (275 ms) and responses had to be emitted within a window of 800 ms. Degner (in press) also found evidence for successful control of affective priming effects but could eliminate control by imposing a response deadline of 600 ms. Although it is now clear that participants can consciously control affective priming effects, more research is needed about the conditions under which control is possible.

The presence of distal goals. In a standard affective priming task, participants are asked to evaluate the targets as good or bad; doing so requires them to adopt the goal to evaluate stimuli. This goal is distal in that it does not refer to the processes by which the attitude toward the prime causes variations in affective priming effects. Nevertheless, it is possible that the processes underlying affective priming operate only when participants have the distal goal to evaluate stimuli. Many studies have shown, however, that affective priming effects (faster responses when prime and target have the same valence than when they differ in valence) can also be found in tasks that do not require the participants to adopt the goal to evaluate stimuli. For instance, affective priming effects have (under certain conditions) been observed when participants are required to read or name the target (e.g., Bargh, Chaiken, Raymond, & Hymes, 1996; De Houwer et al., 2001; Spruyt et al., 2002), to determine the lexical status (e.g., Wentura, 2000) or semantic category of the target (e.g., Spruyt, De Houwer, et al., 2007), or to compare the prime and target with regard to a nonaffective feature, such as color (e.g., Klauer & Musch, 2002). Note, however, that this evidence is not entirely conclusive, because there never was a direct test of whether participants (implicitly) adopted the goal to evaluate stimuli. It would be good to assess this question in future studies, because it is possible that participants adopt the goal to evaluate stimuli even when it is not required by the task.

The presence of awareness. As we discussed earlier, a measure can be denoted as unaware in that it measures an attribute even when participants are unaware of (a) the stimuli that activate the attribute, (b) the origins of the attribute itself, (c) the fact that the attribute influences performance, or (d) the manner in which the attribute influences performance. Evidence suggests that at least some affective priming effects can be classified as unaware in the first two respects. First, several studies (e.g., Abrams, Klinger, & Greenwald, 2002; Draine & Greenwald, 1998; Klauer, Eder, Greenwald, & Abrams, 2007) have revealed affective priming effects even when the primes were presented subliminally (i.e., when participants were not aware of the presentations of the primes). Second, when novel attitudes are created in the lab, they can lead to affective priming effects even when participants are not aware of how the attitudes were acquired (e.g., Olson & Fazio, 2002). Note, however, that this fact does not imply that affective priming effects can register attributes of which participants are not aware. This issue remains to be examined.

We do not know of any study that examined whether participants were aware of the fact that their attitudes toward the prime stimuli influenced their performance or of the way in which the attitudes influenced performance. Of course, in studies on subliminal affective priming, participants are not aware of the prime stimuli and thus cannot be aware of the fact that the attitude toward the primes influences their responses (see Wittenbrink, 2007). It remains to be seen, however, whether participants are aware of the impact of the primes when the primes are presented supraliminally.

The presence of processing resources. Hermans, Crombez, and Eelen (2000) asked participants to perform an affective priming task while they recited a series of digits. They found that the magnitude of the affective priming effect was unaffected by the degree of mental load imposed by the secondary task. This finding suggests that the translation of the attitude in the priming effect is relatively independent of available processing resources and is thus efficient. Klauer and Teige-Mocigemba (2007) replicated this finding for participants who had larger-than-average working memory capacity (as measured on memory span tasks). For participants who had smaller-than-average working memory capacity, however, the priming effect became larger with increases in mental load. The latter finding is in line with the idea that participants engage in effortful processes in an attempt to minimize the impact of the primes on responding. When very few processing resources are available (e.g., when participants who have smaller-than-average working memory capacity are tested under high mental load), these effortful processes can no longer operate and result in stronger affective priming effects.

If the findings of Klauer and Teige-Mocigemba can be confirmed, they would thus offer support for two conclusions: First, the observation of priming effects even when mental load is high suggests that the processes by which the attitude toward the prime influences responding to the target can be automatic in the sense of efficient. Second, the increase in priming effects when fewer processing resources are available suggests that the effect of the prime on responding can be controlled to a certain extent, provided that sufficient processing resources are available. Note that the results of Klauer and Teige-Mocigemba do not reveal whether participants have conscious control of priming effects. In principle, it is possible that the effortful processes involved in controlling the magnitude of the priming effect are activated unconsciously. Whether control is conscious needs to be examined in studies in which participants are asked to report their conscious goals while they perform the task or in which conscious goals are manipulated (e.g., via faking instructions).

The availability of time. There is ample evidence showing that the processes by which the attitude toward the prime produces affective priming effects can operate very quickly and tend to dissipate very quickly over time. For instance, Klauer, Rossnagel, and Musch (1997; see also Hermans, De Houwer, & Eelen, 2001, and Spruyt, Hermans, De Houwer, Vandromme, & Eelen, 2007) found affective priming effects when the onset of the prime occurred 100 ms before (i.e., SOA of 100 ms) and even simultaneously with (i.e., SOA = 0 ms) the onset of the target. Whereas reliable affective priming effects have been observed with SOAs up to 300 ms (e.g., Fazio, Sanbonmatsu, Powell, & Kardes, 1986; Hermans et al., 1994) and when the prime was presented 100 ms after the target (SOA = −100 ms; e.g., Fockenberg, Koole, & Semin, 2006), there have been very few if any reports of reliable affective priming with SOAs larger than 300 ms or smaller than −100 ms (see Klauer & Musch, 2003; for a detailed account of why and how SOA influences priming effects, see Klauer, Teige-Mocigemba, & Spruyt, in press). Given that participants need about 600 ms to evaluate the valence of the target (e.g., Hermans et al., 1994), one can conclude that the prime has an impact on

responses to the target within a time frame starting at about 500 ms after prime onset (in case of an SOA = −100 ms) and ending at 900 ms after prime onset (in case of an SOA = 300 ms). If it is assumed that response execution takes about 200 ms, these estimates can be reduced to 300 ms and 700 ms, respectively.

Summary. Evidence suggests that affective priming effects can be implicit in that they are based on fast, relatively efficient processes (but see Klauer & Teige-Mocigemba, 2007) that can operate even when participants are unaware of the prime stimuli and the origins of the attitude toward the primes. The distal goal to evaluate stimuli in the environment does not seem to be necessary for affective priming effects to occur. Although there is some evidence that certain proximal and distal goals can modulate affective priming effects, the evidence on this specific issue is sparse.

Implications

So far, we have (a) specified the normative criteria that an ideal implicit measure should meet and (b) examined the extent to which IAT and affective priming effects meet the normative criteria. In this third section, we make explicit some of the implications of our work. We first discuss implications for future research on the validation and development of implicit measures. Afterward, we address implications for the use of implicit measures as a tool in research and psychological practice.

Implications for the Validation and Development of Implicit Measures

One of the main virtues of the normative analysis presented in this article is that it clarifies what researchers should aim for when developing implicit measures. When the issues that have already been examined are compared with those that should be examined according to the normative analysis, it becomes apparent what still needs to be done. In short, the normative analysis can guide future research. Our review of the literature on IAT effects and affective priming effects indeed revealed many important caveats in our knowledge about these measures. With regard to the what criterion, more research should be directed not only at uncovering which psychological attributes causally influence IAT and affective priming effects but at understanding the variables that determine the relative impact of those attributes. With regard to the how criterion, there is still debate about the processes underlying IAT and affective priming effects. It seems to be the case that IAT and affective priming effects can be produced by several processes. The relative contributions of these processes and the variables determining their impact have hardly been studied (see Conrey et al., 2005, and Klauer et al., 2007, for exceptions). With regard to the implicitness criterion, much of the work still needs to be done. This is a surprising conclusion, given that implicitness is exactly the feature that is supposed to set apart implicit measures from other measures.

The normative analysis can guide not only future research on IAT and affective priming effects but research on other implicit measures that have already been proposed or that will be proposed in the future. It would lead us too far afield to discuss the implications of the normative analysis separately for each implicit measure that is currently available. We will discuss only one of these measures, namely, scores on the Thematic Apperception Test (TAT; Morgan & Murray, 1935). The TAT is a projective test in which participants are asked to describe pictures of socially ambiguous scenes. On the basis of the content of their responses, scores can be derived that are assumed to reveal implicit motives, such as the need for achievement (e.g., McClelland, Koestner, & Weinberger, 1989). We choose this test because it differs substantially from the IAT and the affective priming task and because it was developed long before the term *implicit measure* was introduced. As such, it allows us to illustrate the width of application of our normative analysis.

From the perspective of the normative analysis, most of the research on the TAT has been directed at verifying the what criterion but little or no research has looked at the how and implicitness criteria. Most TAT studies were correlational in nature and were aimed at assessing whether TAT scores indeed reflect implicit motives (see Lilienfeld, Wood, & Garb, 2000, and McClelland et al., 1989, for opposing views). Very little attention has been given to verifying the how criterion (i.e., to examining the causal nature of the processes by which implicit motives influence the stories that participants produce in response to TAT pictures). The only exception of which we are aware is the work of Tuerlinckx, De Boeck, and Lens (2002), who formulated and tested three simple theories about the processes underlying responses during the TAT. In doing so, they produced important new insights into the reliability and construct validity of the measure. The study of Tuerlinckx et al. is a perfect illustration of Borsboom et al.'s (2004) argument that examining the processes underlying a measure is an essential part of validating a measure.

To the best of our knowledge, there has been little research about whether TAT scores meet the implicitness criterion (i.e., about whether the processes underlying the scores are automatic in a certain manner). It is generally assumed that participants are not aware of the psychological attributes that TAT scores reflect (McClelland et al., 1989), but there are few empirical data about this. We also do not know of any research on the impact of proximal or distal goals, processing resources, or time on TAT scores. Such research is necessary before TAT scores can be described as implicit measures, and it could reveal important information about how these scores come about. We would also like to highlight that, from the perspective of our normative analysis, TAT scores could in principle qualify as implicit measures. Neither the fact that the TAT was introduced before the term *implicit measures* came into use, nor the fact that TAT scores are derived from the content rather than the speed of responses (see Payne et al., 2005, for a measure that is based on the content of responses and that is generally considered to be implicit), is relevant for deciding whether a measure is implicit. The only thing that counts is whether there is empirical evidence to support the conclusion that the processes underlying TAT scores possess features of automaticity.

Implications for Using Implicit Measures as a Tool

Many researchers and practitioners would probably prefer not to wait for future improvements of implicit measures but would like to know now whether and how they should use existing implicit measures as a tool for understanding human behavior. The arguments and evidence that we present in this article clearly show that

the available implicit measures are not perfect. For most measures, it is not entirely clear what they measure, what processes produce the measure, and whether those processes are automatic in a certain manner. This does not mean, however, that the existing measures should not be used. On the contrary, many studies have demonstrated the usefulness of implicit measures.

Most important, it has been demonstrated that implicit measures are at least sometimes related to behavioral variance that is not related to traditional, explicit measures. The evidence for this incremental predictive validity is strongest for IAT effects (see Greenwald et al., in press, for a review). Hence, IAT effects can already provide new and unique insights into behavior. Unlike most other currently available implicit measures, IAT effects are reliable enough to be used as a measure of individual differences (e.g., Bosson et al., 2000; Cunningham et al., 2001). Also, software and guidelines for implementing the IAT are readily available (e.g., Lane et al., 2007).

Nevertheless, we do advise some degree of caution when interpreting IAT effects and other currently available implicit measures, especially at the level of a single individual. As with most behavior, the responses from which implicit measures are derived are determined by a variety of factors. It is therefore risky to interpret an implicit measure as a pure index of one particular psychological attribute. One should also avoid drawing conclusions about the implicitness of a measure in the absence of detailed empirical evidence. Because the different features of automaticity do not necessarily co-occur, each automaticity feature needs to be examined separately. The general scientific principle of convergence can be followed in an attempt to overcome these problems. A conclusion can be drawn with greater confidence when different implicit measures support that conclusion.

As our knowledge of implicit measures increases, less caution will be needed when interpreting these measures. The more we know about the different psychological attributes that influence an implicit measure (what criterion), the processes by which psychological attributes produce the measure (how criterion), and the automaticity of the underlying processes (implicitness criterion), the more confident we can be in deciding what a particular measure actually means. Hence, by verifying whether measures meet the what, how, and implicitness criteria, we can gradually increase the overall quality of implicit measures as tools for studying human behavior.

We want to point out that implicit measures are not the only measures that need to be interpreted with caution. The what and how criteria apply to all measures, implicit or otherwise. There is probably not a single measure of psychological attributes that is perfect, in that it fully meets both criteria. It is not entirely clear what is captured by many traditional measures. Self-report measures, for instance, are known to be susceptible to the effects of many extraneous factors (e.g., social desirability, the precise wording of items, the sequence in which items are presented; see Schwarz, 1999, 2007, for a discussion of some of these factors). Also, little is known about the processes by which psychological attributes can influence self-reports.

Just as traditional measures have proven to be useful despite these imperfections, implicit measures can provide added value despite the caveats regarding our knowledge about these measures. The fact that a measure does not meet the normative criteria should not necessarily stop us from using it. We should always use the best available measures and interpret them in ways that are supported by the available evidence. Identifying the imperfections of a measure should, however, provide the impetus and direction for studying the measure further and improving it where possible. The normative criteria that were put forward in this article facilitate the detection of imperfections and gaps in our knowledge. As such, they can be of great value for the further development of implicit measures of psychological attributes.

General Discussion

Implicit measures of attitudes, stereotypes, and other psychological attributes have become popular in research disciplines as diverse as social, personality, clinical, consumer, and health psychology. Despite their widespread use, there is still much confusion about what implicit measures actually are. On the basis of the work of Borsboom (Borsboom, 2006; Borsboom et al., 2004) and De Houwer (De Houwer, 2006; De Houwer & Moors, 2007), an implicit measure can be defined as the outcome of a measurement procedure that results from automatic processes by which the to-be-measured attribute causally determines the outcome (see Figure 1B). From this definition, we have derived three normative criteria that an ideal implicit measure should meet: (a) The what criterion stipulates that we should know the attributes that causally produce variation in the measure. (b) The how criterion requires that the processes by which the to-be-measured attribute causes variations in the measure are known. (c) The implicitness criterion entails that the processes underlying a measure should be automatic. For each implicit measure, one can examine the extent to which it meets the three normative criteria.

The normative analysis put forward in this article provides a heuristic framework for past and future research on implicit measures. We have used this framework to review the literature on the two currently most popular implicit measures: IAT effects and affective priming effects. By doing so, we have clarified what is already known about these measures and, perhaps more important, what needs to be examined in future studies.

Acceptance of our normative analysis and heuristic framework depends on acceptance of the definition of the concept "implicit measures," from which the analysis and framework were derived. As is the case for all definitions, the definition of the concept "implicit measure" is a matter of convention and thus to a certain extent arbitrary. We cannot guarantee that everyone will agree with our definition, but we do believe that the work of Borsboom et al. (2004) and De Houwer (De Houwer, 2006; De Houwer & Moors, 2007) provides a solid conceptual basis for our definition. At the very least, it has the merit of being explicit. As such, our definition provides the conceptual basis for clarifying disagreements about the meaning of the term *implicit measures* and thus about the normative criteria that an ideal implicit measure should meet.

In line with the arguments of Borsboom (Borsboom, 2006; Borsboom et al., 2004), we have argued that experimental studies should be crucial in validation research. Experiments are the gold standard for establishing whether a psychological attribute causes variation in a measurement outcome and how it does so. Validation research is theoretical research. It should be directed at testing theories about which attributes causally determine measurement outcomes in which ways. Correlational studies can inform the

construction and evaluation of such theories, especially when they are conducted in a systematic manner (see Nosek & Smyth, 2007, for an example of such an approach).

Until now, we have largely ignored one piece of correlational evidence: the reliability of a measure. Often, reliability is considered to be a necessary condition for validity. This is, however, not entirely true. As argued by Borsboom et al. (2004) and Tuerlinckx et al. (2002), a measure can be valid (i.e., caused by the to-be-measured attribute) even when it is not reliable. Such a situation can, for instance, arise when the underlying psychological attribute does not remain stable over time or context. The presence of reliability also provides little information about what it is that the measure captures. Reliability is, however, an important determinant of the overall quality of a measure (e.g., Borsboom et al., 2004). For instance, when the aim is to predict future behavior, a measure is required that remains stable over time. Hence, it is important to continue to examine the reliability of measures.

We should also examine whether a measure is influenced by attributes other than the to-be-measured attribute. The extent to which a measurement outcome can be used to make an inference about a specific attribute of the person depends not only on whether that attribute causes variation in the outcome but on whether other attributes cause variation in the outcome. Put differently, interpreting a measure as indicative of an attribute requires not only that the attribute is a cause of the outcome but that it is the only systematic cause of the outcome. If other attributes can cause variations in the outcome independent of the to-be-measured attribute, one can never be sure whether a certain outcome reflects the to-be-measured attribute or another one (e.g., Fiedler et al., 2006). Research on the what and how criteria thus should not be restricted to the to-be-measured attribute but should examine whether and when other attributes can cause variations in the outcome. It is also important to realize that empirical research will not suffice to determine what a measure actually captures. Detailed conceptual analyses should be undertaken to examine the ontological status of the psychological attributes that are measured. The upper limit of what a measure can tell is determined by what is known about the attribute that the measure is assumed to capture.

Studies on the implicitness criterion also involve extensive and complicated research. Which features of automaticity apply to each implicit measure must be examined empirically. One could argue that only some of the automaticity features are truly relevant for determining the implicitness of measures. We do not commit to a position on this point, but we do offer the conceptual tools for making possible a debate about what features of automaticity are central for implicit measures. Moors and De Houwer (2006) recently defined in detail the various features of automaticity. Applying this analysis in the context of implicit measures (see also De Houwer & Moors, 2007) makes it clear in what ways a measure can be implicit. This process allows researchers to specify what they mean when they say that a measure is implicit and to debate the merits of the various possible features of automaticity and implicitness.

Rather than try to select one feature or set of features as defining for the implicitness of measures, one could examine as many features as possible in an attempt to understand more fully the conditions under which the to-be-measured attribute causally influences the measure. One could then try to match different measures with different real-life behaviors in terms of the extent to which those measures and behaviors are influenced by the same attribute under the same set of conditions. In line with the idea of transfer-appropriate processing (e.g., Roediger, 1990), one could argue that the more similar a measure is to a behavior in this respect, the more the measure will be able to predict the behavior. For instance, real-life, attitude-driven behavior that occurs when people do not have the conscious goal to evaluate stimuli in the environment (e.g., buying products under time pressure) might be related most to measurement outcomes that occur in the absence of a conscious evaluation goal. This approach entails that one should study in detail not only the conditions under which the attribute influences the measure but also the conditions under which the attribute influences the to-be-predicted behavior.

Given the quantity and complexity of the research involved in verifying whether and in what sense a measure can be regarded as an implicit measure, one might choose to adopt an apparently more simple, pragmatic approach in which various measures are simply related to various behaviors without much consideration for conceptual or theoretical issues. Measures that predict a particular behavior can be considered useful even if it is not known what attribute the measures actually capture, how they do so, or which features of automaticity apply. On the one hand, we do agree that the practical use of implicit measures should not await a full evaluation in terms of the three normative criteria put forward in this article. Despite the important gaps in our knowledge about implicit measures, these measures could help researchers predict and understand certain behaviors. On the other hand, a purely pragmatic approach does have serious limitations. First, in the absence of strong empirical evidence, one should refrain from making statements about how a measure should be interpreted, how it works, or whether it is implicit. Second, without a basic level of theoretical understanding of the measures, there is little ground for predicting when a measure will be related to which kind of behavior. Progress in obtaining evidence for relations between measures and behavior will thus proceed slowly and haphazardly. Likewise, there will be little guidance for attempts to improve the quality of the measures. In the end, a purely pragmatic approach might be less efficient than a conscientious conceptual and theoretical approach to understanding implicit measures of psychological attributes. We hope that the normative analysis put forward in this article will be of help to researchers who choose to adopt this difficult but necessary approach.

References

Abrams, R. L., Klinger, M. R., & Greenwald, A. G. (2002). Subliminal words activate semantic categories (not automated motor responses). *Psychonomic Bulletin & Review, 9,* 100–106.

Anastasi, A. (1958). *Differential psychology* (3rd ed.). New York: Macmillan.

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

Back, M. D., Schmukle, S. C., & Egloff, B. (2005). Measuring task-switching ability in the Implicit Association Test. *Experimental Psychology, 52,* 167–179.

Banaji, M. R. (2001). Implicit attitudes can be measured. In H. L. Roediger & J. S. Nairne (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 117–150). Washington, DC: American Psychological Association.

Banse, R. (1999). Automatic evaluation of self and significant others: Affective priming in close relationships. *Journal of Social and Personal Relationships, 16,* 805–824.

Banse, R. (2001). Affective priming with liked and disliked persons: Prime visibility determines congruency and incongruency effects. *Cognition & Emotion, 15,* 501–520.

Banse, R., Seise, J., & Zerbes, N. (2001). Implicit attitudes toward homosexuality: Reliability, validity, and controllability of the IAT. *Zeitschrift für Experimentelle Psychologie, 48,* 145–160.

Bargh, J. A. (1992). The ecology of automaticity. Toward establishing the conditions needed to produce automatic processing effects. *American Journal of Psychology, 105,* 181–199.

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

Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review, 6,* 242–261.

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

Blanton, H., & Jaccard, J. (2006). Arbitrary metrics in psychology. *American Psychologist, 61,* 27–41.

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

Borsboom, D. (2006). The attack of the psychometricians. *Psychometrika, 71,* 425–440.

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

Bosson, J. K., Swann, W. B., & Pennebaker, J. W. (2000). Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited? *Journal of Personality and Social Psychology, 79,* 631–643.

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

Cai, H. J., Sriram, N., Greenwald, A. G., & McFarland, S. G. (2004). The Implicit Association Test's D measure can minimize a cognitive skill confound: Comment on McFarland and Crouch (2002). *Social Cognition, 22,* 673–684.

Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review, 82,* 407–428.

Collins, A. M., & Quillian, M. R. (1969). Retrieval from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8,* 240–248.

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.

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

Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science, 12,* 163–170.

Czellar, S. (2006). Self-presentational effects in the Implicit Association Test. *Journal of Consumer Psychology, 16,* 92–100.

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

Degner, J. (in press). On the (un)controllability of affective priming: Strategic manipulation is feasible but can possibly be prevented. *Cognition & Emotion.*

Degner, J., Wentura, D., Gniewosz, B., & Noack, P. (2007). Hostility-related prejudice against Turks in adolescents: Masked affective priming allows for a differentiation of automatic prejudice. *Basic and Applied Social Psychology, 29,* 245–256.

De Houwer, J. (2000). [Testing the response criterion account of the Implicit Association Test]. Unpublished raw data.

De Houwer, J. (2001). A structural and process analysis of the Implicit Association Test. *Journal of Experimental Social Psychology, 37,* 443–451.

De Houwer, J. (2003a). The extrinsic affective Simon task. *Experimental Psychology, 50,* 77–85.

De Houwer, J. (2003b). A structural analysis of indirect measures of attitudes. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 219–244). Mahwah, NJ: Erlbaum.

De Houwer, J. (2006). What are implicit measures and why are we using them. In R. W. Wiers & A. W. Stacy (Eds.), *The handbook of implicit cognition and addiction* (pp. 11–28). Thousand Oaks, CA: Sage.

De Houwer, J. (2009). Comparing measures of attitudes at the procedural and functional level. In R. Petty, R. H. Fazio, & P. Brinol (Eds.), *Attitudes: Insights from the new implicit measures*. New York: Psychology Press.

De Houwer, J. (in press). Implicit measures: Similarities and differences. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition.* New York: Guilford.

De Houwer, J., Beckers, T., & Moors, A. (2007). Novel attitudes can be faked on the Implicit Association Test. *Journal of Experimental Social Psychology, 43,* 972–978.

De Houwer, J., Geldof, T., & De Bruycker, E. (2005). The Implicit Association Test as a general measure of similarity. *Canadian Journal of Experimental Psychology, 59,* 228–239.

De Houwer, J., Hermans, D., & Eelen, P. (1998). Affective and identity priming with episodically associated stimuli. *Cognition & Emotion, 12,* 145–169.

De Houwer, J., Hermans, D., Rothermund, K., & Wentura, D. (2002). Affective priming of semantic categorization responses. *Cognition & Emotion, 16,* 643–666.

De Houwer, J., Hermans, D., & Spruyt, A. (2001). Affective priming of pronunciation responses: Effects of target degradation. *Journal of Experimental Social Psychology, 37,* 85–91.

De Houwer, J., & Moors, A. (2006). [Awareness of the aim of the Implicit Association Test]. Unpublished raw data.

De Houwer, J., & Moors, A. (2007). How to define and examine the implicitness of implicit measures. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes: Procedures and controversies* (pp. 179–194). New York: Guilford Press.

Devine, P. G., Plant, E. A., Amodio, D. M., Harmon-Jones, E., & Vance, S. L. (2002). The regulation of explicit and implicit race bias: The role of motivations to respond without prejudice. *Journal of Personality and Social Psychology, 82,* 835–848.

Draine, S. C., & Greenwald, A. G. (1998). Replicable unconscious semantic priming. *Journal of Experimental Psychology: General, 127,* 286–303.

Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin, 23,* 316–326.

Eder, A., & 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.

Egloff, B., & Schmukle, S. C. (2002). Predictive validity of an Implicit Association Test for assessing anxiety. *Journal of Personality and Social Psychology, 83,* 1441–1455.

Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and individual differences: A natural science approach.* New York: Plenum.

Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition & Emotion, 15,* 115–141.

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

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

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.

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

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

Fockenberg, D. A., Koole, S. L., & Semin, G. R. (2006). Backward affective priming: Even when the prime is late, people still evaluate. *Journal of Experimental Social Psychology, 42,* 799–806.

Frings, C., & Wentura, D. (2003). Who is watching "Big Brother"? TV consumption predicted by masked affective priming. *European Journal of Social Psychology, 33,* 779–791.

Gawronski, B. (Ed.). (2007). What is an attitude? [Special issue]. *Social Cognition, 25*(5).

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., 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.

Gawronksi, 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.

Gemar, M. C., Segal, Z. V., Sagrati, S., & Kennedy, S. J. (2001). Mood-induced changes on the Implicit Association Test in recovered depressed patients. *Journal of Abnormal Psychology, 110,* 282–289.

Govan, C. L., & Williams, K. D. (2004). Changing the affective valence of the stimulus items influences the IAT by redefining the category labels. *Journal of Experimental Social Psychology, 40,* 357–365.

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., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74,* 1464–1480.

Greenwald, A. G., & Nosek, B. A. (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., Nosek, B. A., Banaji, M. R., & Klauer, K. C. (2005). Validity of the salience asymmetry interpretation of the Implicit Association Test: Comment on Rothermund and Wentura (2004). *Journal of Experimental Psychology: General, 134,* 420–425.

Greenwald, A. G., Nosek, B. A., & Sriram, N. (2006). Consequential validity of the Implicit Association Test: Comment on Blanton and Jaccard (2006). *American Psychologist, 61,* 56–61.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E., & Banaji, M. R. (in press). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology.*

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.

Hermans, D., Baeyens, F., & Eelen, P. (2003). On the acquisition and activation of evaluative information in memory: The study of evaluative learning. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 139–168). Mahwah, NJ: Erlbaum.

Hermans, D., Crombez, G., & Eelen, P. (2000). Automatic attitude activation and efficiency: The fourth horseman of automaticity. *Psychologica Belgica, 40,* 3–22.

Hermans, D., De Houwer, J., & Eelen, P. (1994). The affective priming effect: Automatic activation of evaluative information in memory. *Cognition & Emotion, 8,* 515–533.

Hermans, D., De Houwer, J., & Eelen, P. (2001). A time course analysis of the affective priming effect. *Cognition & Emotion, 15,* 143–165.

Hermans, D., Smeesters, D., De Houwer, J., & Eelen, P. (2002). Affective priming for associatively unrelated primes and targets. *Psychologica Belgica, 42,* 191–212.

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

Hofmann, W., Rauch, W., & Gawronski, B. (2007). And deplete us not into temptation: Automatic attitudes, dietary restraint, and self-regulatory resources as determinants of eating behavior. *Journal of Experimental Social Psychology, 43,* 497–504.

Houben, K., & Wiers, R. W. (2006). A test of the salience asymmetry interpretation of the Alcohol-IAT. *Experimental Psychology, 53,* 292–300.

Hummert, M. L., Garstka, T. A., O'Brien, L. T., Greenwald, A. G., & Mellott, D. S. (2002). Using the Implicit Association Test to measure age differences in implicit social cognitions. *Psychology and Aging, 17,* 482–495.

Kane, M. J., & Engle, R. W. (2003). Working memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General, 132,* 47–70.

Karpinski, A., & Hilton, J. L. (2001). Attitudes and the Implicit Association Test. *Journal of Personality and Social Psychology, 81,* 774–778.

Kim, D. Y. (2003). Voluntary controllability of the Implicit Association Test (IAT). *Social Psychology Quarterly, 66,* 83–96.

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

Kinoshita, S., & Peek-O'Leary, M. (2006). Two bases of the compatibility effect in the Implicit Association Test (IAT). *Quarterly Journal of Experimental Psychology, 56,* 2102–2120.

Klauer, K. C., Eder, A. B., Greenwald, A. G., & Abrams, R. L. (2007). Priming of semantic classifications by novel subliminal prime words. *Consciousness and Cognition, 16,* 63–83.

Klauer, K. C., & Mierke, J. (2005). Task-set inertia, attitude accessibility, and compatibility-order effects: New evidence for a task-set switching account of the Implicit Association Test effect. *Personality and Social Psychology Bulletin, 31,* 208–217.

Klauer, K. C., Mierke, J., & Musch, J. (2003). The positivity proportion effect: A list context effect in masked affective priming. *Memory & Cognition, 31,* 953–967.

Klauer, K. C., & Musch, J. (2002). Goal-dependent and goal-independent effects of irrelevant evaluations. *Personality and Social Psychology Bulletin, 28,* 802–814.

Klauer, K. C., & Musch, J. (2003). Affective priming: Findings and theories. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 7–50). Mahwah, NJ: Erlbaum.

Klauer, K. C., Musch, J., & Eder, A. (2005). Priming of semantic classifications: Late and response-related, or earlier and more central? *Psychonomic Bulletin & Review, 12,* 897–903.

Klauer, K. C., Rossnagel, R., & Musch, J. (1997). List–context effects in evaluative priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 246–255.

Klauer, K. C., & Teige-Mocigemba, S. (2007). Controllability and resource dependence in automatic evaluation. *Journal of Experimental Social Psychology, 43,* 648–655.

Klauer, K. C., Teige-Mocigemba, S., & Spruyt, A. (in press). Contrast effects in spontaneous evaluations: A psychophysical account. *Journal of Personality and Social Psychology.*

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.

Kornblum, S., & Lee, J.-W. (1995). Stimulus–response compatibility with relevant and irrelevant stimulus dimensions that do and do not overlap with the response. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 855–875.

Lane, K. A., Banaji, M. R., Nosek, B. A., & Greenwald, A. G. (2007). Understanding and using the Implicit Association Test: 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.

Lasaga, M. I., & Garner, W. R. (1983). Effect of line orientation on various information-processing tasks. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 215–225.

Lilienfeld, S. O., Wood, J. M., & Garb, H. N. (2000). The scientific status of projective techniques. *Psychological Science in the Public Interest, 1,* 27–66.

Lowery, B. S., Hardin, C. D., & Sinclair, S. (2001). Social influence effects on automatic racial prejudice. *Journal of Personality and Social Psychology, 81,* 842–855.

Lucas, M. (2000). Semantic priming without association: A meta-analytic review *Psychonomic Bulletin & Review, 7,* 618–630.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109,* 163–203.

Maison, D., Greenwald, A. G., & Bruin, R. H. (2004). Predictive validity of the Implicit Association Test in studies of brands, consumer attitudes and behavior. *Journal of Consumer Psychology, 14,* 405–416.

McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Bulletin, 96,* 690–702.

McFarland, S. G., & Crouch, Z. (2002). A cognitive skill confound on the Implicit Association Test. *Social Cognition, 20,* 483–510.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review, 100,* 254–278.

Meiran, N., Chorev, Z., & Sapir, A. (2000). Component processes in task switching. *Cognitive Psychology, 41,* 211–253.

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

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

Mischel, W., & Shoda, Y. (1995). A cognitive–affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review, 102,* 246–268.

Mitchell, J. P., Nosek, B. A., Banaji, M. R. (2003). Contextual variations in implicit evaluation. *Journal of Experimental Psychology: General, 132,* 455–469.

Monteith, M. J., Voils, C. I., & Ashburn-Nardo, L. (2001). Taking a look underground: Detecting, interpreting, and reacting to implicit racial bias. *Social Cognition, 19,* 395–417.

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

Morgan, C. D., & Murray, H. A. (1935). A method of investigating fantasies: The Thematic Apperception Test. *Archives of Neurology and Psychiatry, 34,* 289–306.

Moskowitz, G. B., Salomon, A. R., & Taylor, C. M. (2000). Preconsciously controlling stereotyping: Implicitly activated egalitarian goals prevent the activation of stereotypes. *Social Cognition, 18,* 151–177.

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

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration website. *Group Dynamics, 6,* 101–115.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (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., Greenwald, A. G., & Banaji, M. R. (2007). The Implicit Association Test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Social psychology and the unconscious: The automaticity of higher mental processes* (pp. 265–292). New York: Psychology Press.

Nosek, B. A., & Hansen, J. J. (2008a). The associations in our heads belong to us: Searching for attitudes and knowledge in implicit evaluation. *Cognition & Emotion, 22,* 553–594.

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

Nosek, B. A., & Smyth, F. L. (2007). A multitrait–multimethod validation of the Implicit Association Test: Implicit and explicit attitudes are related but distinct constructs. *Experimental Psychology, 54,* 14–29.

Olson, M. A., & Fazio, R. H. (2001). Implicit attitude formation through classical conditioning. *Psychological Science, 12,* 413–417.

Olson, M. A., & Fazio, R. H. (2002). Implicit acquisition and manifestation of classically conditioned attitudes. *Social Cognition, 20,* 89–104.

Olson, M. A., & Fazio, R. H. (2003). Relations between implicit measures of prejudice: What are we measuring? *Psychological Science, 14,* 636–639.

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

Otten, S., & Wentura, D. (1999). About the impact of automaticity in the minimal group paradigm: Evidence from affective priming tasks. *European Journal of Social Psychology, 29,* 1049–1071.

Payne, B. K., Burkley, M. A., & Stokes, M. B. (2008). Why do implicit and explicit attitude tests diverge? The role of structural fit. *Journal of Personality and Social Psychology, 94,* 16–31.

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

Pecher, D., Zeelenberg, R., & Raaijmakers, J. G. W. (1998). Does pizza prime coin? Perceptual priming in lexical decision and pronunciation. *Journal of Memory and Language, 38,* 401–418.

Popa-Roch, M.-A. (2008). *Validation expérimentale d'une mesure implic-*

ite du prejudge: Le test d'associations implicates. Unpublished doctoral dissertation, Université Pierre Mendès, Grenoble, France.

Proctor, R. W., & Cho, Y. S. (2006). Polarity correspondence: A general principle for performance of speeded binary classification tasks. *Psychological Bulletin, 132,* 416–442.

Ratcliff, R. (1988). A retrieval theory of priming in memory. *Psychological Review, 95,* 385–408.

Roediger, H. L. (1990). Implicit memory: Retention without remembering. *American Psychologist, 45,* 1043–1056.

Rogers, R., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124,* 207–231.

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

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

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

Schmitz, F., Teige, S., Voss, A., & Klauer, K. C. (2005, June). *Working memory load in the IAT.* Paper presented at the Fifth Workshop on Implicit Representations and Personality, Berlin, Germany.

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.

Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist, 54,* 93–105.

Schwarz, N. (2007). Cognitive aspects of survey methodology. *Applied Cognitive Psychology, 21,* 277–287.

Sekaquaptewa, D., Espinoza, P., Thompson, M., Vargas, P., & von Hippel, W. (2003). Stereotypic explanatory bias: Implicit stereotyping as a predictor of discrimination. *Journal of Experimental Social Psychology, 39,* 75–82.

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

Spalding, L. R., & Hardin, C. D. (1999). Unconscious unease and self-handicapping: Behavioral consequences of individual differences in implicit and explicit self-esteem. *Psychological Science, 10*(6), 535–539.

Spruyt, A., De Houwer, J., Hermans, D., & Eelen, P. (2007). Affective priming of non-affective semantic categorization responses. *Experimental Psychology, 54,* 44–53.

Spruyt, A., Hermans, D., De Houwer, J., & Eelen, P. (2002). On the nature of the affective priming effect: Affective priming of naming responses. *Social Cognition, 20,* 225–254.

Spruyt, A., Hermans, D., De Houwer, J., Vandekerckhove, J., & Eelen, P. (2007). On the predictive validity of indirect attitude measures: Prediction of consumer choice behavior on the basis of affective priming in the picture-picture naming task. *Journal of Experimental Social Psychology, 43,* 599–610.

Spruyt, A., Hermans, D., De Houwer, J., Vandromme, H., & Eelen, P. (2007). On the nature of the affective priming effect: Effects of stimulus onset asynchrony and relatedness proportion in naming and evaluative categorization. *Memory & Cognition, 35,* 95–106.

Steffens, M. (2004). Is the Implicit Association Test immune to faking? *Experimental Psychology, 51,* 165–179.

Strayer, D. L., & Kramer, A. F. (1994). Strategies and automaticity: I. Basic findings and conceptual framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 318–341.

Teige-Mocigemba, S., & Klauer, K. C. (2008). "Automatic" evaluation? Strategic effects on affective priming. *Journal of Experimental Social Psychology, 44,* 1414–1417.

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

Thurstone, L. L. (1928). Attitudes can be measured. *American Journal of Sociology, 33,* 529–554.

Tuerlinckx, F., De Boeck, P., & Lens, W. (2002). Measuring needs with the Thematic Apperception Test: A psychometric study. *Journal of Personality and Social Psychology, 82,* 448–461.

Wentura, D. (1999). Activation and inhibition of affective information: Evidence for negative priming in the evaluation task. *Cognition & Emotion, 13,* 65–91.

Wentura, D. (2000). Dissociative affective and associative priming effects in the lexical decision task: Yes versus no responses to word targets reveal evaluative judgment tendencies. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 456–469.

Wentura, D., Kulfanek, M., & Greve, W. (2005). Masked affective priming by name letters: Evidence for a correspondence of explicit and implicit self-esteem. *Journal of Experimental Social Psychology, 41,* 654–663.

Wentura, D., & Rothermund, K. (2007). Paradigms we live by: A plea for more basic research on the IAT. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes* (pp. 195–215). 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.

Wittenbrink, B. (2007). Measuring attitudes through priming. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes* (pp. 17–58). New York: Guilford Press.

Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology, 72,* 262–274.

Wittenbrink, B., Judd, C. M., & Park, B. (2001). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality and Social Psychology, 81,* 815–827.

Received May 19, 2008
Revision received September 3, 2008
Accepted September 8, 2008 ■