Rydell, R. J., & McConnell, A. R. (2010). Consistency and inconsistency in implicit social cognition: The case of implicit and explicit measures of attitudes. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 295-310). New York: Guilford.

CHAPTER 16

Consistency and Inconsistency in Implicit Social Cognition

The Case of Implicit and Explicit Measures of Attitudes

Robert J. Rydell and Allen R. McConnell

Everyone experiences mixed or conflicted feelings from time to time. Does the divorcée "get back on the horse" and start dating again even when the sting of a disastrous marriage still lingers? At the restaurant, how does a diabetic patron ponder whether or not to order the tasty cheesecake, knowing that it is laden with so much sugar, fat, and calories? In a faculty meeting, how does one balance a tenure decision involving a productive colleague who is a complete jerk to students and other faculty alike? So many everyday decisions are fraught with a cacophony of feelings. In social psychological research, we understand that holding mixed feelings about attitude objects (e.g., dating, cheesecake, a colleague) is attitudinal ambivalence. In fact, the study of the consequences of holding inconsistent attitudes has long been acknowledged as important in the field (Kaplan, 1972; Petty & Briñol, 2009). However, it is interesting that social psychological thinking about inconsistent attitudes has focused almost exclusively on the implications of holding conflicting attitudes that people can verbalize. Yet what happens when evaluative inconsistencies operate at different levels (i.e., implicit and explicit measures of attitudes)? In the current chapter, we explore how discrepancies between implicit and explicit measures of attitudes develop, how a variety of theoretical perspectives and psychological processes can account for them, and what important implications result from holding them.

When considering these issues, it is important to specify a number of constructs that, at times, seem poorly defined in the literature. For example, there is much debate and discord about the definition of an attitude (Gawronski, 2007). In general, attitudes are viewed as a perceiver's evaluation of some object. In other words, is the object liked or disliked? More formally, attitudes are "a psychological tendency that is expressed by evaluating a particular entity with some degree of favor or disfavor" (Eagly & Chaiken, 1993, p. 1). Yet our understanding of attitudes is complicated by the fact that attitudes are not directly observable, and thus their existence is only manifested by measuring people's responses, directly or indirectly (Eagly & Chaiken, 2007; Fazio, 2007). That is, attitudes have to be understood by the measures used to assess them, and as a result, attitude measures at best provide estimates of attitudes.[1]

As attitude theories develop, attitude conceptualization and measurement evolve as well. During the past decade, the explosion of research featuring implicit attitude measures has led research-

ers to reconsider attitude theory, leading some to retool existent ideas and others to advance new theories about attitudes. Although implicit measures of attitudes have been around for some time (e.g., Fazio, Sanbonmatsu, Powell, & Kardes, 1986), the proliferation of research using newer, indirect attitude measures (see Wittenbrink & Schwarz, 2007, for a review) has led many psychologists to reconsider their conceptualization of attitudes. One reason why many researchers have extended or modified their theoretical perspectives on attitudes is that there are a number of demonstrations that implicitly and explicitly measured attitudes are often only weakly related and can be responsive to different manipulations (see Gawronski & Bodenhausen, 2006, for reviews of divergence; see Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005, and Nosek, 2005, for reviews on correspondence). In addition, implicit attitude measures are sometimes altered quite easily with subtle manipulations (e.g., Karpinski & Hilton, 2001; Wittenbrink, Judd, & Park, 2001) and at other times they take Herculean effort to change (e.g., Gregg, Seibt, & Banaji, 2006; Rydell & McConnell, 2006). Although explicit attitude measures show equivalent variation in response to attitude change manipulations (e.g., Petty & Wegener, 1998), some conditions reveal distinct and dissociated effects on implicit and explicit attitude measures (see also Gawronski & Sritharan, Chapter 12, this volume). How do attitude theorists explain these divergent outcomes, and what are their consequences for cognition and behavior?

To address these issues, we first review the impact that such discrepancies have had on theorizing about attitudes. Specifically, we detail how major theories of attitudes have been modified, refocused, or created to accommodate implicit–explicit attitude measure discrepancies. Then we examine the processes that can impact implicit and explicit attitude measures differently, leading to discrepancy on these measures. Finally, we examine the consequences of discrepancies between implicit and explicit attitude measures.

THEORETICAL CONCEPTUALIZATIONS OF IMPLICIT–EXPLICIT ATTITUDE MEASURE DISCREPANCIES

To explore why and how consistency and inconsistency occurs for implicit and explicit attitude measures, we examine how several different theories explain these discrepancies. Understanding these theoretical treatments of consistency and discrepancy serves to highlight underlying similarities and differences between the theories' explanations for consistency and inconsistency between implicit and explicit attitude measures.

The MODE Model

Research on the motivation and opportunity as determinants (MODE) model (Fazio, 1995, 2007; Fazio & Olson, 2003) was the first to integrate findings from reaction time measures of implicitly measured attitudes and paper-and-pencil measures of explicitly measured attitudes. From the perspective of the MODE model, attitudes are simply object–evaluation associations that are stored in memory, which can vary in their strength of association. The stronger the associative link between the object and its evaluation in memory, the greater that attitude's accessibility. From the perspective of the MODE model, implicit attitude measures capture attitudes at an earlier time in their expression than do explicit measures, and they better reflect the associative strength of the attitude object and its evaluation in memory. Explicit attitude measures, on the other hand, can be inconsistent with implicit attitude measures because reporting explicit attitudes is itself a deliberative behavior, which can be impacted by deception, self-presentational concerns, or inconsistency of the accessible attitude with previously expressed beliefs or ideals (Fazio, 2007).

According to the MODE model, implicit and explicit attitudes diverge when people have the motivation and ability to engage in deception, self-presentation, or comparison of behavior and thought over time. That is, divergence should be seen when people monitor their explicit attitude expression and deploy the cognitive resources required to alter their expression. Convergence, however, should be seen when motivation or opportunity is low and attitudes accessed from memory guide the expression of evaluation on explicit attitude measures. In the MODE model, attitude change (i.e., the change of the association between the attitude object and its evaluation in memory) occurs following repeated pairing of the attitude object with counterattitudinal information (e.g., Rudman, Ashmore, & Gary, 2001; Rydell, McConnell, Strain, Claypool, & Hugenberg, 2007).

However, how does the MODE model account for quick implicit attitude change (e.g., Wittenbrink et al., 2001)? As Fazio (2007) points out, this quick change can be due to manipulations that impact the difficulty of the response-mapping task posed

by many implicit attitude measures (e.g., Han, Olson, & Fazio, 2006; Olson & Fazio, 2004) or to changes in the categorization of an attitude object (e.g., Smith, Fazio, & Cejka, 1996). For example, decreased prejudice toward African Americans in church as opposed to those in a jail occurs not because implicit attitudes have changed but because African Americans are not the same attitude object in jail as in church. As another example, motivational states such as hunger (Seibt, Häfner, & Deutsch, 2007) and thirst (Ferguson & Bargh, 2004) affect the construal of these attitude objects to be so different that, for instance, a cheesecake is not the same attitude object to a diabetic when he or she is hungry versus full (Fazio, 2007, p. 626). The MODE model assumes that studies showing quick change among implicitly measured attitudes reflect the relative influence of current concerns, context, or construal (categorization) on attitude object instantiation. This temporary change is adaptive in the moment but does not impact the attitude object evaluation held in memory.

The Associative–Propositional Evaluation Model

The associative–propositional evaluation (APE) model (Gawronski & Bodenhausen, 2006) assumes that implicit and explicit attitude measures tap into two different types of evaluative processes: associative and propositional. In this model, associative processes determine affective reactions automatically activated when one encounters an attitude object (e.g., feelings of negativity when walking by an obnoxious colleague's office). The APE model proposes that implicit attitudes capture the process of pattern activation (see Smith, 1996), wherein a subset of the information associated with an attitude object is activated based on learning history, contextual information, motivational states, and other stimuli. Thus, the evaluative information activated in memory by the attitude object is the best fit between these factors and the information connected in memory at that moment. In another moment when these factors change, other associations with the attitude object will be active, changing implicit attitude measure responses. Because of rapid changes in pattern activation, implicit attitude measures are able to change quickly.

On the other hand, propositional processes underlie the expression of evaluations on explicit attitude measures. According to Gawronski and Bodenhausen (2006), "Evaluations resulting from propositional processes can be characterized as evaluations that are based on syllogistic inferences derived from any kind of propositional information that is considered relevant for a given judgment" (p. 694). These propositions can include information extracted from the outcome of associative processes (e.g., "I like X" or "I dislike X"; see Strack & Deutsch, 2004). All propositions are then subject to logical inference that assesses which propositions are valid and, therefore, used in explicit attitude measure responses. For example, if a department chair has an automatic negative reaction to the obnoxious but productive professor, implicit attitude measures will reflect this negativity. However, if the department chair also knows that this professor receives substantial amounts of grant support, has prestigious publications, and attracts high-quality graduate students, explicit attitudes toward this professor will likely be more positive. In this latter case, the propositional information about the utility of the obnoxious professor for the department will result in greater positivity being expressed on explicit measures of attitudes. In other words, even though the department chair's automatic affective reaction is negative, the proposition formed in response to this reaction will be invalidated because of its inconsistency with other meaningful information about the professor and will thus be rejected. This propositional process of validating or invalidating automatic evaluations is how the APE model explains consistency and inconsistency between implicit and explicit attitude measures (see also Gawronski, Strack, & Bodenhausen, 2009).

Thus, when the information automatically activated from memory is consistent with propositional information about the attitude object, implicit and explicit attitude measures should be consistent. However, when information automatically activated from memory is inconsistent (i.e., does not logically fit) with propositional information, implicit and explicit attitude measures should be discrepant. As such, the truth value (or validity) of the propositional output determines whether discrepancies occur (Gawronski & Bodenhausen, 2006). Discrepancies between implicit and explicit measures can also result from changing the propositions available or seen as applicable to the situation. Of course, consistency between implicit and explicit attitude measures also occurs when associative information changes to be in line with propositional information or propositional information changes to be in line with the output of associative processes. Thus, a change in either process can lead to consistency or inconsistency of implicit and explicit attitude measures; however,

the propositional process ultimately determines whether there is consistency or inconsistency between measures.

The Metacognitive Model

The metacognitive model (MCM; Petty, Briñol, & DeMarree, 2007; Petty, Tormala, Briñol, & Jarvis, 2006), like the MODE model, assumes that attitudes are evaluations of objects that are stored in memory. However, the MCM posits that attitude objects can (sometimes) be associated simultaneously with positive and negative valence. When attitude objects are not simultaneously associated with positivity and negativity, this model is very similar to the MODE model (e.g., Fazio, 1995, 2007). However, the MCM model goes beyond the MODE model because it assumes people tag their associations (e.g., true–false, confident–doubt, accept–reject, valid–invalid). These metacognitive tags are stored in memory to help people determine the validity or confidence with which the association is held. Because of its predictions about simultaneous linkages to positivity and negativity (bivalent associations) and its assumptions about metacognitive tags, the MCM makes several predictions about when consistency and inconsistency will arise. When these bivalent associations exist and one association is tagged as invalid, the metacognitive tags are not easily detected by implicit attitude measures. For example, the negativity automatically activated in response to the obnoxious professor could be tagged as invalid because of his or her productivity. In this situation, implicit attitude measures would reflect the automatic negativity and not the tag of its invalidity; however, explicit attitude measures would account for the invalidity tag. Therefore, unless the tag is highly accessible, it will not impact implicit measures because of the default assumption that associations are true (Gilbert, 1991). Only when people devote time and resources are these tags utilized; this need for elaboration in utilizing tags is exacerbated because metacognitive tags are assumed to be associated with the positivity or negativity in memory and not the attitude object itself.

According to the MCM, implicit and explicit attitude measures will be consistent when (1) only positive or negative information is associated with an attitude object in memory and that information is "tagged" or accepted as true; and (2) when a metacognitive tag automatically negates either the linkage to positivity or negativity and this same negation is made after some deliberation. According to the MCM, implicit and explicit attitude measures will diverge when positive or negative information is not properly tagged as "false" or this tag is not detected by implicit measures, but the tag is used when completing explicit measures (see Petty & Briñol, 2009; Petty et al., 2006).

The Systems of Evaluation Model

The systems of evaluation model (SEM; McConnell, Rydell, Strain, & Mackie, 2008; Rydell & McConnell, 2006; Rydell, McConell, Mackie, & Strain, 2006) assumes that implicit and explicit measures tap into different mental systems that correspond to an associative system (implicit attitude measures) or a rule-based system (explicit attitude measures). Specifically, we (Rydell & McConnell, 2006; Rydell et al., 2006) have proposed that there are two somewhat or partially independent systems of evaluation that differ both in what information they use and in how they act on it (Sloman, 1996; Smith & DeCoster, 2000; Strack & Deutsch, 2004). The associative system of evaluation is relevant to our understanding of how implicit attitudes form and function because implicit attitude measures are posited to follow the basic principles of similarity and association (Smith & DeCoster, 2000). The rule-based system fits with a conceptualization of explicit attitude measures as evaluations based on conscious deliberation or syllogistic reasoning, which can reveal quick but resource-dependent changes on explicit attitude measures (Fazio, 1995, 2007).

According to the SEM, implicit attitude measures are more sensitive to associative forms of information such as subliminal primes (Rydell et al., 2006) or associative cues (e.g., race, obesity, physical attractiveness; McConnell et al., 2008). For example, eating cheesecake brings about pleasurable feelings associated with rich, satisfying foods. On the other hand, explicit attitude measures are more sensitive to symbolic forms of information (quite often verbal in nature, although other forms exist as well, such as musical notation and mathematical symbols that exist in a framework governed by rules), which are often used in logic and symbolic reasoning. For instance, one can use reasoning and deduction to understand that eating cheesecake, because it can lead to high blood sugar, could lead to blindness or death.[2] However, these systems do interact when no (or a relatively small amount of) associative or verbal information is available. For instance, subliminal evaluative priming has been shown to impact explicit attitude measures (e.g., Olson & Fazio, 2001; Petty et al., 2006), and a large amount of verbal informa-

tion can influence implicit attitude measures (e.g., Rydell & McConnell, 2006; Rydell et al., 2007). Thus, the SEM predicts that implicit and explicit attitude measures will be consistent when the valence of associative information (e.g., subliminal primes, cues with strong evaluative associations) and verbal information is equivalent or when a considerable amount of verbal information contradicts the valence of previous associative learning. Implicit and explicit measures will be inconsistent when the valence of the associative information and verbal information becomes divergent through a variety of means, including different evaluative implications of rule-based (e.g., verbal) and associative-based (e.g., associative cues) information (e.g., McConnell et al., 2008; Rydell et al., 2006) or when a relatively small amount of new verbal information revises rule-based evaluations but is insufficient to alter association-based evaluations (e.g., Rydell et al., 2007)

One major difference between the SEM and the theories reviewed previously is that it posits that there are distinct systems as opposed to distinct processes that can lead to discrepancies between implicit and explicit attitude measures. In the SEM, systems refer to aspects of the mind that use particular types of information and operating principles, actively or passively, to render evaluations. Essentially, each system can be conceived of as an interrelated group of mechanisms that are dedicated to evaluation. Because the two systems differ in their inputs and style of information processing, their outputs can differ. A system approach assumes that two broad groups of processes and mechanisms operate and serve to disambiguate the attitude object, its features, and its value to (or probable impact on) the perceiver given current environmental constraints and perceiver goals. Instead of viewing these mechanisms as part of one process (e.g., activation of evaluations stored in memory—Fazio, 2007; Petty et al., 2007; pattern activation in a connectionist network—Gawronski & Bodenhausen, 2006), a systems approach assumes that, although related mechanisms operate on similar input, even similar mechanisms may produce inconsistent outputs.

The Constructivist Perspective

Finally, a constructivist perspective views all attitudes as momentarily constructed evaluations that are created each time an object is encountered (Schwarz, 2007; Schwarz & Bohner, 2001). As such, the term *attitude* does not refer to stable evaluative information stored in memory but rather to the process of constructing evaluations based on the information that is available at the time and the demands of the attitude measure. In this conceptualization, attitudes are epiphenomena and are a by-product of the construction process. Thus, the notion that people possess attitudes is, conceptually, inherently flawed.

According to a constructionist perspective, implicit and explicit attitude measures will converge when the same information is culled from memory to construct an attitude. This could result from context making accessible the same information or the measures themselves eliciting the same information (e.g., "What was my last experience with cheesecake?"). To the extent that the context or any momentary stimuli impact the information recruited from memory (e.g., a divorcée is at a romantic movie vs. talking to the ex-spouse), implicit and explicit attitude measures will be inconsistent with past measurements. In addition, if the methods used to assess implicit and explicit attitudes differ in what they access from memory or demand different responses based on framing (Schwarz, 1999), then implicit and explicit attitude measures will differ. From this perspective, the consistency of implicit and explicit attitude measures is a consequence of how attitudes are constructed in the moment, which can be impacted by how attitudes are assessed (cf. Fazio, 2007, for a strong challenge to this perspective). Finally, the timing of responses may lead to discrepancies between implicit and explicit attitudes according to a constructivist perspective. Implicit attitude measures assess quick responses, and explicit attitude measures might reflect different evaluations to the extent that they take longer to answer. That is, when completing explicit attitude measures, people could recruit additional information from memory that they were unable to retrieve when responding quickly on implicit attitude measures that might change their construal of the attitude object.

MECHANISMS OF CONSISTENCY AND INCONSISTENCY

In this section, we review several key mechanisms that have been studied in research examining discrepancies and consistencies between implicit and explicit attitude measures. The mechanisms are organized (in order) based on whether they tend to (1) have a larger impact on automatically activated evaluations or (2) be based on further elaboration and integration of automatically acti-

vated evaluations with other information relevant for rendering evaluations. By highlighting these mechanisms, we provide an overview of how the aforementioned attitude theories differ in their explanation of these different processes, processes that must be explained by a comprehensive theory of attitudes. In our assessment, we examine integration of counterattitudinal information, balanced identities, extrapersonal associations, dissonance and balance, and "downstream" processes. In addition, we discuss how the simultaneous automatic activation of positivity and negativity, which is an associative process, is accounted for by different models of attitudes.

Integration of Counterattitudinal Information

The most straightforward way that automatic affective reactions can be changed is by repeatedly encountering counterattitudinal information (e.g., Karpinski & Hilton, 2001; Rudman et al., 2001; Rydell & McConnell, 2006). For example, Rydell and colleagues (2007) had participants learn 100 positive pieces of information about a novel individual. After this initial learning phase, participants were presented with 0, 20, 40, 60, 80, or 100 pieces of additional, counterattitudinal information (i.e., negative information). Implicit attitude measures assessing evaluations of the novel individual became increasingly negative in proportion to the amount of counterattitudinal information encountered. Furthermore, Sherman and colleagues (2008) showed through multinomial modeling (see Sherman, Klauer, & Allen, Chapter 9, this volume) that changes in implicit attitude measures in this paradigm were due to changes in association between the attitude object and valence in memory as opposed to other possible processes (i.e., overcoming bias, discrimination of stimuli, guessing). Most models of attitudes predict that implicit attitude measures should change when enough counterattitudinal information is encountered (see Gawronski & Sritharan, Chapter 12, this volume). The MODE model, MCM, and SEM specify that this should occur through the alteration of object–evaluation links in memory. The APE model assumes that this change occurs because a larger subset of heterogeneous (i.e., positive and negative) information about the attitude object is stored in memory and, therefore, associations that are of the opposite valence of originally learned information should be more likely to be used (and connected to other attitude relevant information) in associative processes.

Balanced Identities

Research examining Greenwald and colleagues' (2002) unified model of social cognition provides considerable evidence that implicit attitude measures can be changed via balance principles. As an illustration, Greenwald and colleagues assessed female participants' self-esteem, gender identification, and attitudes toward women with implicit and explicit measures. The results showed that the interaction of any two of the implicit measures predicted the third implicit measure. Most relevant to the current discussion, women who had greater implicitly measured self-esteem and implicitly measured gender identification showed relatively more positive implicitly measured attitudes toward women. However, consider a situation where women have been exposed to extremely negative female exemplars. Based on past research, this exposure to negative female exemplars should make implicitly measured attitudes toward women relatively more negative (Dasgupta & Greenwald, 2001). However, a balanced identity perspective predicts that implicitly measured attitudes toward women should be relatively more positive to the extent that women have greater implicitly measured gender identification and greater implicitly measured self-esteem. Implicitly measured attitudes toward women could be influenced by the exemplar manipulation, but because this association is inconsistent with (or "pressured" by) greater implicitly measured self-esteem and implicitly measured gender identification, implicitly measured attitudes toward women should be relatively positive (see Greenwald et al., 2002, Principle 2). Presumably, negative evaluations of women would be not activated or may even be inhibited to maintain balance among implicit measures of self-esteem, gender identification, and attitudes toward women. It should be noted that there are two alternatives to these predictions about the impact of a negative female exemplar and inhibition. First, the exemplar manipulation could, in fact, lead to less positive implicit evaluations of women, but this change in the mean level of implicitly measured attitudes would not impact the correlations to gender identity and self-esteem. Second, self-esteem could be reduced via spreading activation from evaluations of women overall. This change in self-esteem would lead to balance but at a cost to the self (e.g., Walther, Nagengast, & Trasselli, 2005). To our knowledge, these interesting possibilities have not been examined empirically.

How would models of attitudes explain the tendency for balance in the face of information that

has, in past research, changed implicit attitude measures? Based on the MODE model, balance could be achieved through changing categorization of women when implicitly measured attitudes toward women are pressured. Thus, women could be recategorized in some other fashion (e.g., high-achieving women), and this would produce balance among implicit measures. The APE model would explain this balance among implicit measures by changes in pattern activation in response to information that the self is positive and that the self is a woman. A likely output, given this set of inputs, is that implicitly measured attitudes toward women would be positive. Through processes like pattern completion, a pattern of activity would be "settled" on indicating that implicitly measured attitudes toward women are relatively positive through a passive process of constraint satisfaction. The SEM has a more difficult time explaining balanced identities. It would predict that the critical factor would be whether the exemplar manipulation constituted associative or verbal information. To the extent that this information is verbal, implicitly measured self-esteem and implicitly measured gender identification could impact implicitly measured attitudes toward women by influencing the valence of the evaluations that are activated in memory via facilitation of associations between positivity and women and inhibition of associations between negativity and women. However, if the negative information is associative in nature, implicitly measured attitudes toward women should be less positive and not balanced with implicitly measured self-esteem and gender identification. Thus, the SEM predicts that manipulations that are associative in nature can counteract the processes underlying balanced identities.

It is worth pointing out that the impact of balanced identities is less likely on explicit measures of self-esteem, gender identity, and attitudes toward women because balance on explicit measures can be achieved in many more ways (e.g., adding concepts, splitting the pressured concept) than on implicit measures, and it is, therefore, not a necessity that explicit attitude measures converge to maintain overall balance (e.g., Greenwald et al., 2002).

Extrapersonal Information

Another way associations could impact the consistency or discrepancy between implicit and explicit attitude measures is if an implicit attitude measure is contaminated by "extrapersonal associations" (Fazio, 2007; Fazio & Olson, 2003; Han et al., 2006; Karpinski & Hilton, 2001; Olson & Fazio, 2004). Extrapersonal associations are environmental associations that come from one's experiences in the greater culture or from other individuals rather than reflecting one's own attitude (Fazio, 2007). It has been suggested that implicit attitude measures, such as the Implicit Association Test (IAT), are affected by extrapersonal associations unrelated to one's personally held attitude, and thus, discrepancies result when implicit measures of attitudes are unduly influenced by extrapersonal information.

In an illustrative study, Han and colleagues (2006) had participants initially learn that one Pokemon character was objectively superior to another. Next, they exposed participants to a video that had children express beliefs about the Pokemon characters that were consistent or inconsistent with initial learning. The children's views were seen as silly and illogical when they were inconsistent with initial learning (i.e., they were rejected as a valid indicator of the object's positivity or negativity). Despite the perceived invalidity of the information given by the children, their information impacted a traditional IAT measure of implicit attitudes. Interestingly, information given by the children did not impact a personalized IAT (where the category labels were marked "I like" vs. "I dislike" instead of "pleasant" vs. "unpleasant"; Olson & Fazio, 2004) or an evaluative priming measure (Fazio et al., 1986). Because the latter two measures are assumed to measure personal attitudes rather than momentarily accessible knowledge, they were perceived to be better measures of attitudes assessing an individual's, and not societal, beliefs. The main point of this work is that implicit and explicit measures of attitudes can sometimes diverge because extrapersonal information impacts implicit attitude measures, but these extrapersonal associations are invalidated and not used when explicit measures are completed.

The role of extrapersonal information in implicit attitude measures is a point of debate in the literature (Conrey & Smith, 2007; Gawronski & Bodenhausen, 2006; Nosek, 2005; Nosek & Hansen, 2008). Indeed, from several perspectives (e.g., APE, MCM, SEM), people should not be able, at an associative level, to distinguish intrapersonal from extrapersonal knowledge. That task must occur via controlled processes in which one can assess the extent to which an evaluation is personal. Indeed, recent research has challenged the claim that implicit attitude measures, including the IAT, are influenced by cultural knowledge (Nosek & Hansen, 2008). Furthermore, the impact that cultural knowledge has on implicit attitude measures

is accounted for by its impact on explicit attitude measures. As these authors note, "Endorsement, especially in the context of implicit cognition, is irrelevant for information to be a measure of individual attitude and predict individual behavior. ... Associative representations reflect accumulated experience with attitude objects regardless of whether those experiences are accepted or rejected as true" (Nosek & Hansen, 2008, p. 549). Because introspective access to the causes or circumstances involving associations in memory is poor and may not even be stored in memory, extrapersonal influences should not impact implicit attitude measures.

If this is true, why did Han and colleagues (2006) show an impact of extrapersonal associations on the IAT if cultural knowledge does not impact this measure? One possibility is that people actually do tag evaluations as "mine" or "not mine." The MCM's process of metacognitive tagging could be used to tag either positivity or negativity stored in memory as "not mine." Under most circumstances, this tag would not impact implicit attitude measures. However, if the tag is relatively salient or strongly associated with evaluative information, it could potentially impact an implicit attitude measure (see Gawronski, Peters, & LeBel, 2008, for a similar argument where personal associations could be interpreted as the stored outcomes of propositional inferences in associative memory). Another, more likely possibility for Han and colleagues' results is based on the fact that implicit attitude measures are not process pure (e.g., Sherman et al., 2008) and operate through different cognitive mechanisms (e.g., De Houwer, 2003). For the personalized IAT and the traditional IAT, which share the same response-interference mechanisms, the differences are likely due to how the labels ("I like" vs. "pleasant") impact controlled processes that influence responding to the IAT. Moreover, because the paradigm used in the traditional IAT (unlike the personalized IAT) requires that participants make correct responses on each trial (and that they receive error feedback on trials in which incorrect responses are made), it is possible that additional processes are involved in the personalized IAT (e.g., overcoming bias; Sherman et al., 2008) that do not apply in the same way to the traditional IAT. Turning to explicit attitude measures, personal endorsement is important for determining people's attitudes; however, it remains to be seen whether this is true for implicit attitude measures. Further research will need to address this issue more directly to determine the relation between cultural or extrapersonal knowledge and implicit attitude measures (see Nosek & Hansen, 2008, for a review).

Cognitive Consistency

An important aspect of understanding consistencies and discrepancies between implicit and explicit attitude measures is elucidating the role of cognitive consistency on these measures. Research on cognitive consistency has adopted two frameworks regarding implicit and explicit attitude measures: dissonance theory and balance theory (which is distinct from and potentially broader than the balanced identities perspective outlined previously). Thus, we consider each of these areas of research separately.

Dissonance Theory

In early research on dissonance theory and implicit and explicit attitude measures, researchers found that dissonance only led to changes in explicitly measured attitudes (e.g., Wilson, Lindsey, & Schooler, 2000). Gawronski and Strack (2004) showed that an induced compliance paradigm led to changes in explicit attitude measures when participants were given choice but not when no choice was given. Implicit attitude measures did not show any impact of choice, but they were positively correlated to explicit attitude measures in the no-choice condition and uncorrelated to explicit attitude measures in the high-choice condition. Thus, the dissonance created by "freely chosen" counterattitudinal behavior had an impact only on explicit attitude measures. The APE, MCM, and SEM models predict, consistent with what was found, that only explicit measures changed. The APE model comes to this conclusion because logical relations between propositions (e.g., "I disliked a boring task," "I told someone else the task was fun") can only be tested by propositional processes. If these propositions are logically inconsistent, explicit attitude measures changed to resolve this logical inconsistency. The MCM would predict that the automatically activated negativity to the task would be tagged as invalid and, therefore, explicit attitudes would show the impact of this tag, unless the tag were to become strongly associated with negativity (which is extremely unlikely in an induced compliance paradigm). The SEM would predict that dissonance would be dealt with by the rule-based system of evaluation, where the logical inference and consistency concerns can be assessed. The MODE model would predict one of two accounts for these results. First, implicit atti-

tude measures assess evaluation at an earlier point when dissonance arousal is not a concern, and thus implicit attitude measures do not change. Second, implicit attitude measures did not change because they were measured with a standard IAT (rather than the personalized IAT or affective priming). That is, the impact of extrapersonal associations may have concealed changes on implicit attitude measures in a forced-choice paradigm. A problem with this second account is that explicit attitude measures that are completed under time pressure (presumably an implicit attitude measure) do not reveal attitude change in a forced-choice paradigm (Wilson et al., 2000), and this null result cannot be explained by extrapersonal associations impacting implicit attitude measurement.

However, there is evidence that in other dissonance paradigms implicit attitudes can be changed. In a postdecisional attitude change study, Gawronski, Bodenhausen, and Becker (2007) showed that implicit attitude measures can change. Specifically, participants chose between one of two equally attractive pictures. In the usual spreading of alternatives study, the dissonance created by the positive features of the unselected picture lead to a devaluing of the picture and the features on which it was seen as positive. In the Gawronski and colleagues study, implicit attitude measures for each picture were assessed before and after participants chose. The results showed that implicitly measured attitudes were more positive toward the chosen picture and more negative toward the rejected picture, suggesting that choice decisions led to changes in implicit attitude measures. Given that such attitude changes are typically explained by the operation of postdecisional dissonance, this outcome seems inconsistent with the models discussed previously. However, a second study found that these differences were not due to dissonance but to self-anchoring. Specifically, in a second study, the chosen picture became more strongly associated with the self than the rejected picture. Importantly, implicit measures of self-esteem were positively correlated to implicit attitude measures toward the chosen picture (Greenwald et al., 2002). Thus, dissonance paradigms can lead to changes in implicit attitude measures via different mechanisms than dissonance reduction that are not always propositional in nature.

How do these postdecisional attitude change studies fit with the models we discussed in the first section? All of the models would assume that as the attitude object becomes more associated with the self, and the self is viewed positively either through spreading activation or pattern activation, the attitude object would be more positive on implicit measures (see also Greenwald et al., 2002). The strength of this association and the positivity of the self should be the main determinants of implicitly measured attitude change. Especially because this association was just created in an experimental setting and does not have a long history of coactivation, the models' predictions do not differ.

Balance Theory

In a similar vein to studies on balanced identities, research has examined the impact of simple balance principles on attitudes assessed by implicit and explicit measures (e.g., Gawronski, Walther, & Blank, 2005). This research showed that when participants learned positive or negative information about a novel individual and then learned whether this individual liked or disliked a second, novel individual, balance was observed on implicit and explicit attitude measures. Take the example of learning that person *A* is positive. If person *A* acted positively and participants learned that person *A* liked person *B*, person *B* was assessed positively on implicit and explicit attitude measures. If person *A* acted positively and participants learned that person *A* disliked person *B*, person *B* was assessed negatively on implicit and explicit attitude measures. Interestingly, when participants learned about the relationship between person *A* and person *B* (whether they like or dislike one another) and then learned about person *A*, this balance effect did not occur on either implicit or explicit attitude measures. As stated by Gawronski and colleagues (2005), these findings "suggest that cognitive balance influences the encoding of social information rather than the retroactive construal of evaluative judgments" (p. 625).

This finding is interesting because Gawronski and Strack (2004) showed that cognitive dissonance led to "retroactive construal of evaluative judgments" or changes on an explicit attitude measure after a freely chosen counterattitudinal behavior. Why was this change absent in the balance paradigm? Gawronski and colleagues (2005) propose that this difference could be due to the fact that "source valence [evaluation of person *A*] and observed sentiment [person *A*'s feeling toward person *B*] may be stored independently in memory when source valence is encoded after observed sentiments" (p. 625). This explanation assumes that when source valence is encoded after observed sentiments, it is less likely that both pieces of information will be retrieved from memory and

associated with one another. Thus, similar to research on the sleeper effect, the relation between person *A* and person *B* is "lost" or unused when followed by the qualifying information about the valence of person *A* (Kumkale & Albarracín, 2004). If the relation between person *A* and person *B* is not accessible, there may be no inconsistency altogether because inconsistency at the explicit level requires that the relevant information is accessible and subjectively valid in order to produce inconsistency (see Gawronski, Peters, & Strack, 2008). Nevertheless, this lack of balance on explicit measures, unlike that seen in the balanced identity work, seems much more problematic for theories of attitudes. In this research, there are not, for instance, additional cognitions available to resolve the imbalance on explicit attitude measures.

"Downstream" Consequences

Implicit and explicit attitudes can also be discrepant because additional information, as opposed to automatic affective reactions, affects responding to explicit attitude measures. Indeed, there is much work showing that implicit and explicit attitudes are more likely to diverge when people have the opportunity and ability to change or "correct" the information activated automatically (Fazio & Olson, 2003). People can change their explicitly measured attitude because of self-presentational concerns, social norms, and motives to hold certain beliefs (e.g., Dunton & Fazio, 1997). All of the models that have been reviewed assume that explicit attitudes can be changed with motivation and ability. As Fazio (2007) noted, these downstream consequences can be a mixture of automatically activated evaluations and "corrective" processes, or they can be based solely on controlled processing. However, the APE model (Gawronski & Bodenhausen, 2006) assumes that these changes are all based on controlled processing of propositional pieces of information about the attitude object because the automatically activated affective reaction is transformed into a proposition itself (see Strack & Deutsch, 2004). Thus, there are differences between models as to whether mixed processes can lead to changes in explicit attitudes, but these differences are hard to distinguish empirically. For instance, according to the APE model, cognitive dissonance can be reduced only via propositional processes (Gawronski & Strack, 2004). In other models, when implicit and explicit attitudes themselves are inconsistent, this should lead to tension or ambivalence because the associative and rule-based systems are discordant (Rydell, McConnell, & Mackie, 2008) or because positivity and negativity are concurrently activated from memory (Petty & Briñol, 2009).

Regardless, most models of attitudes assume that downstream consequences occur through very similar deliberative processes (e.g., syllogistic reasoning, logic) and need motivation and ability to be instantiated. For instance, the MODE model and the SEM predict that implicit and explicit attitude measures can diverge only through elaborative information processing and that greater divergence would be associated with more elaborative information processing (e.g., Hofmann et al., 2005). Although controlled thought is likely necessary for discrepancies between implicit and explicit attitude measures, the type of thought matters (Gawronski & Bodenhausen, 2006). If thoughts that are contradictory to automatically activated affective reactions are considered during conscious deliberations, then implicit and explicit attitudes should diverge with more elaborative processing. However, if thoughts supporting the automatically activated affective reaction are utilized during elaborative processing, implicit and explicit attitude measures could show greater convergence as more elaboration occurs.

Because the impact of downstream consequences on implicit attitude measures has been extensively examined elsewhere (Fazio, 1995; Gawronski & Bodenhausen, 2006; Sloman, 1996), we do not discuss them further. However, it should be noted that perceivers could use an almost infinite number of strategies to modulate automatically activated evaluations of an attitude object when responding to explicit attitude measures. Future work should address when certain strategies are used and how they impact the consistency between implicit and explicit attitude measures and the cognitive and behavioral outcomes of these discrepancies.

SIMULTANEOUS AUTOMATIC ACTIVATION OF POSITIVITY AND NEGATIVITY

Not all inconsistency has to come from differences in implicit and explicit attitude measures or differences in explicit attitude measures. Namely, de Liver, van der Pligt, and Wigboldus (2007) showed that both positive and negative information can be simultaneously activated from memory and impact implicit attitude measures. Specifically, de Liver and colleagues had participants generate objects that they believed were positive, negative, neutral, or ambivalent (i.e., both positive and neg-

ative). Their work showed (1) strong associations between positivity and objects that were described as positive; (2) strong associations between negativity and objects that were described as negative; (3) no associations between positivity or negativity for objects that were described as neutral; and (4) strong associations between the objects and both positivity and negativity when objects were described as ambivalent. Thus, it seems that a single attitude object can simultaneously activate both positivity and negativity in memory (see also Petty et al., 2007).

How is this finding explained with the models of attitudes outlined previously? As discussed earlier, the MCM (Petty et al., 2007) explicitly makes this assumption about attitude representation and indeed uses this simultaneous activation along with metacognitive tagging to explain discrepancies between implicit and explicit attitude measures and as the basis for implicit ambivalence (i.e., implicit ambivalence is caused by this simultaneous activation along with a metacognitive tag negating the relation between either positivity or negativity and the attitude object in memory). According to the APE model (Gawronski & Bodenhausen, 2006), simultaneous, spontaneous activation of positivity and negativity in response to an attitude object can occur when both positive and negative information is stored in memory, activated from memory, and is the output of pattern activation and constraint satisfaction. That is, if the pattern of activation that best fits the current inputs from information in memory, the environment, and motivation is indeed mixed in valence, then implicit attitude measures will be able to detect this pattern of activation.

It is interesting to think about how the output of associative processes is transferred into propositional information when both positivity and negativity are automatically activated. It could be that two inconsistent propositions are created (i.e., "I like X" and "I dislike X") or that a single proposition is created that includes both positivity and negativity ("I like and dislike X"; "I am unsure how I feel about X"; or even "X makes me feel uncomfortable"). Research on this question may be important for incorporating the MCM and APE model. The MODE model (Fazio, 2007) asserts that activating both positive and negative information can occur when attitudes are less accessible and evaluations need to be constructed. However, when attitudes are more accessible or have already been summarized in memory, then there should not be the automatic activation of both positivity and negativity in response to an attitude object. As Fazio (2007) states, "The essence of more reasoned forms of attitude development is an integration of any such conflicting information into a summary evaluation ... the summary evaluation effectively resolves the ambivalence" (p. 627). Of course, positive or negative summary evaluations can be automatically activated from memory depending on how an object is categorized (Smith et al., 1996). Perhaps asking participants to think of attitudes for which they are ambivalent leads to categorization of the attitude object into both a category for which there is positivity and a category for which there is negativity. In this way, the MODE model could explain the findings of de Liver and colleagues (2007). The SEM (Rydell & McConnell, 2006) would explain the simultaneous spontaneous activation of positivity and negativity in response to an attitude object that can occur when heterogeneous associative information is stored in memory and the situation and attitude object activates both this positive and negative information at the same time.

CONSEQUENCES OF DISCREPANCIES ON IMPLICIT AND EXPLICIT ATTITUDE MEASURES

Research and theorizing on when and how implicit and explicit attitude measures differ or converge is extremely helpful in defining and understanding the attitude concept. However, another key question is, what are the consequences of implicit and explicit attitude measures diverging? The only research examining the psychological consequences of divergent implicit and explicit attitude measures (outside of work on self-esteem; see Zeigler-Hill & Jordan, Chapter 21, this volume) or beliefs has shown that increased discrepancies lead to greater implicit ambivalence (a stronger association between the attitude object and doubt in memory) and increased information processing of attitude-relevant information (Petty et al., 2006; see Petty & Briñol, 2009, for an extensive review). This research showed that once attitudes formed they were not completely replaced when attitudes changed and increased implicit ambivalence accompanied attitude change. In conditions in which attitudes had changed and implicit and explicit measures were divergent, information processing increased. Briñol, Petty, and Wheeler (2006) showed that the greater the discrepancy between standardized measures of explicit and implicit self-beliefs (e.g., one's own shyness), the more

extensive the processing of persuasive messages related to the domain of discrepancy (e.g., arguments favoring shyness). This result, however, did not occur when information available for subsequent information processing was unrelated to the attitude object. As discrepancies between implicit and explicit attitude measures increased, people were motivated to consider carefully subsequently presented relevant information. Yet why does this outcome occur, and what phenomenology occurs with this increased information processing?

Rydell and colleagues (2008) showed that when explicit–implicit attitude discrepancies were greater, more negative arousal (i.e., dissonance; Festinger, 1957) was evoked. This negative arousal, in turn, induced greater information processing of attitude object-relevant information. The results are consistent with the idea that discrepancies between implicit and explicit attitude measures can create aversive feelings that people attempt to reduce through learning more about the attitude object. This research also showed that a previously used measure of implicit ambivalence (i.e., IAT measure used by Petty et al., 2006) could not account for increased information processing when implicit and explicit attitude measures diverged, but reports of dissonance arousal could.

However, both Rydell and colleagues (2008) and Petty and Briñol (2009) note that inconsistent cognitions (here attitudinal information assessed by implicit and explicit measures) are uncomfortable and lead to increased information processing. The dissonance referred to in Rydell and colleagues was similar to Festinger's (1957) conceptualization: negative arousal felt in response to two inconsistent cognitions. Dissonance theory now recognizes several conditions that must be met for inconsistency to lead to dissonance (e.g., Cooper & Fazio, 1984). At this point, the theoretical conceptualizations of a dissonance or an ambivalence view are relatively similar because ambivalence and dissonance as conceptualized by Rydell and colleagues (consistent with Festinger, 1957) are simply inconsistent cognitions that are uncomfortable, and people are motivated to avoid such feelings (see Petty & Briñol, 2009, for the same prediction). Although the implicit ambivalence IAT measure used by Petty and colleagues (2006) did not predict the extent of information processing for these inconsistencies, the main underlying point of both these lines of research is the same and can probably best be described as felt ambivalence, a subjective state of discomfort or tension that is due to inconsistent evaluations, because it does not require all of the conditions necessary for cognitive dissonance to occur.

The next step for this research is to examine which behaviors that are engaged to reduce felt ambivalence are effective in resolving the discrepancy between implicit and explicit attitude measures and actually reducing felt ambivalence. It seems prudent to look at self-regulatory models to examine when this effort will and will not be successful. For instance, in terms of convergence between measures, models of self-regulation suggest that the velocity (or speed) with which inconsistencies can be resolved reduces negative affect (e.g., Carver & Scheier, 2002). Therefore, it is likely that changes in explicit attitude measures may be more likely in many circumstances. Because explicit attitude measures generally show change more quickly (e.g., Rydell et al., 2007) and implicit attitude measures may change because of contextual moderation of activated associations or changes in categorization (which may be ephemeral), explicit attitude change may be the most likely route to provide faster relief from discomfort (i.e., consonant cognitions). However, changes in the associations stored in memory between an attitude object and valence (whether this is through an associative network or the amount of strongly weighted information available for pattern activation) may be more likely to reduce the frequency of felt ambivalence in the future. Thus, moderation of implicit measures of attitudes is possible too. Regardless, the fewer self-regulatory resources needed to reconcile discrepancies between implicit and explicit attitude measures, the more resources that can be applied to other issues that people face. Indeed, the experience of discomfort caused by discrepancies between the mental residue revealed by implicit and explicit attitude measures may be functional in that it serves as a signal that action is needed and self-regulatory resources should be brought to bear on the attitude object (e.g., Harmon-Jones & Harmon-Jones, 2002).

Furthermore, whether implicitly or explicitly measured attitudes change in response to discrepancies may be related to the accessibility of the attitude. If an attitude is highly accessible because of recent or frequent use (Fazio, 1995) and the attitude object elicits discrepancy between implicit and explicit attitude measures, it may be more imperative to alter evaluations to attain convergence between measures because feelings of discomfort should be more accessible as well. In addition, to the extent that there is a positive relation between implicit and explicit attitude measure discrepan-

cies and discomfort (see Rydell et al., 2008), highly accessible attitudes that are widely discrepant on implicit and explicit measures should be especially pronounced, attracting greater attentional resources in the service of reducing attitude discrepancy. And it may be that lasting changes on implicit attitude measures may be more likely to reduce discomfort over a longer period of time than changes on explicit attitude measures (on average) because changing the information automatically activated in memory, in many theories the "seed" of the attitude concept (Fazio, 2007; Gawronski & Bodenhausen, 2006), should be more likely to reduce discrepancies consistently.

CONCLUSIONS

Understanding discrepancies and consistencies between explicit and implicit attitudes has been at the forefront of contemporary social cognition research. Delineating when implicit and explicit attitude measures are impacted similarly or differently by a manipulation not only is important for models of attitudes but speaks to meaningful debates regarding (1) the processes through which automatic and controlled cognition diverge (e.g., Gawronski & Bodenhausen, 2006), (2) the consequences of their divergence (Briñol et al., 2006), and (3) exactly how to conceptualize attitude representation (Fazio, 2007; Schwarz, 2007). From a functional perspective, it seems that having inconsistent scores on implicit and explicit attitude measures is aversive and induces one to expend cognitive resources in the service of understanding the attitude object better, suggesting that such circumstances merit the individual's attention and thus have psychological value. How people get to that point is, at present, not fully understood but is necessary for correctly predicting behavior (see Perugini, Richetin, & Zogmaister, Chapter 14, this volume).

People experience conflicted feelings toward the self, others, groups, and consumer products. Social psychologists have long understood and studied this phenomenon by examining discrepancies between explicit attitude measures under the auspices of attitude ambivalence. Although this approach has yielded many insights (Petty & Briñol, 2009, for a review), the current chapter examined how discrepancies between implicit and explicit attitude measures occur (in terms of both attitude models and cognitive processes) and what their impact is on cognition and behavior.

Let us return to the examples that we opened the chapter with: the divorcée thinking about dating again, the diabetic patron deciding about whether to order the cheesecake, the faculty member making a tenure decision about an obnoxious but productive junior colleague. Each of these examples could be examined by looking at ambivalent attitudes (e.g., does the sweetness of the cheesecake outweigh its fat and sugar content?). However, we believe that by examining discrepancies between implicit and explicit measures, social psychologists gain further insight into how people experience and render decisions. How does a spontaneous negative reaction to an obnoxious junior colleague and a more positive evaluation based on consideration of several positive features (e.g., grant support, prestigious publications) occur? Different models of attitudes answer this question very differently. What processes bring these measures closer together or pull them farther apart and why? Again, the answer depends on the model of attitudes adopted. Finally, what is the psychological impact of these discrepancies for the perceiver and for making decision strategies? When do these discrepancies help versus hinder action or impact versus not impact behavior? Understanding attitudes has proven difficult time and time again. However, we believe that work on discrepancies between implicit and explicit attitudes can provide useful insights for understanding why a divorcée starts dating someone who is obviously wrong for her or why diabetics find it difficult to resist the foods that wreak havoc on their health. It is our hope that by understanding how these discrepancies operate we can examine how to manage them in ways that promote beneficial outcomes.

ACKNOWLEDGMENTS

The research described in this chapter was supported by Grant No. BCS 0601148 from the National Science Foundation and by the Lewis Endowed Professorship.

NOTES

1. In this chapter we use the terms *implicit attitude* and *explicit attitude* to refer to measures of attitudes that are indirect (e.g., infer evaluations from response latency tasks that involve the attitude object or from evaluations of novel stimuli associated with the attitude object) or direct (e.g., asking people for their evaluations of an attitude object on a scale), respectively. Here, implicit at-

titude measures are conceptualized as measures that better assess evaluations activated without the intent of the perceiver, whose expression is difficult to control, and are less affected by higher order cognitive processes. Explicit attitude measures are conceptualized as measures that better assess evaluations whose expression can be controlled by the perceiver and for which people can utilize higher order cognitive processes. These relatively broad definitions are used in the current work to facilitate comparisons between and among the different models of attitudes described in this chapter that examine consistency and inconsistency between implicit and explicit attitude measures (yet disagree about, e.g., the antecedents of these measures). It is also important to note that implicit and explicit measures are not process pure (e.g., Sherman et al., 2008), meaning that a strict dichotomy between processes and operations assessed by implicit and explicit attitude measures is not warranted.

2. In this model, the input into the rule-based system is referred to as "verbal," at least in part, because the way social psychologists usually present information in attitudes research is in terms of language (there are some exceptions in research on evaluative conditioning; e.g., Olson & Fazio, 2001). We emphasize that what makes this system distinct is its rule-based underpinnings, not its ties to language. For example, other forms of rules operate in music (e.g., musical notation) and math (e.g., laws of algebra or trigonometry) that are not purely verbal. Although these examples are certainly verbalizable, they are symbolic and rule governed.

REFERENCES

Briñol, P., Petty, R. E., & Wheeler, S. C. (2006). Discrepancies between explicit and implicit self-concepts: Consequences for information processing. *Journal of Personality and Social Psychology, 91,* 154–170.

Carver, C. S., & Scheier, M. F. (2002). Control processes and self-organization as complementary principles underlying behavior. *Personality and Social Psychology Review, 6,* 304–315.

Conrey, F. R., & Smith, E. R. (2007). Attitude representation: Attitudes as patterns in a distributed, connectionist representational system. *Social Cognition, 25,* 718–735.

Cooper, J., & Fazio, R. H. (1984). A new look at dissonance theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 17, pp. 229–266). Orlando, FL: Academic Press.

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

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

de Liver, Y., van der Pligt, J., & Wigboldus, D. (2007). Positive and negative associations underlying ambivalent attitudes. *Journal of Experimental Social Psychology, 43,* 319–326.

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.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes.* Fort Worth, TX: Harcourt, Brace, Jovanovich.

Eagly, A. H., & Chaiken, S. (2007). The advantages of an inclusive definition of attitude. *Social Cognition, 25,* 582–602.

Fazio, R. H. (1995). Attitudes as object-evaluation associations: Determinants, consequences, and correlates of attitude accessibility. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 247–282). Mahwah, NJ: Erlbaum.

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

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and uses. *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.

Ferguson, M. J., & Bargh, J. A. (2004). Liking is for doing: The effects of goal pursuit on automatic evaluation. *Journal of Personality and Social Psychology, 87,* 557–572.

Festinger, L. (1957). *A theory of cognitive dissonance.* Palo Alto, CA: Stanford University Press.

Gawronski, B. (2007). Attitudes can be measured! But what is an attitude? *Social Cognition, 25,* 573–581.

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., Bodenhausen, G. V., & Becker, A. P. (2007). I like it, because I like myself: Associative self-anchoring and post-decisional change of implicit evaluations. *Journal of Experimental Social Psychology, 43,* 221–232.

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

Gawronski, B., Peters, K. R., & Strack, F. (2008). Cross-cultural differences vs. universality in cognitive dissonance: A conceptual reanalysis. In R. M. Sorrentino & S. Yamaguchi (Eds.), *Handbook of motivation and cognition across cultures* (pp. 297–314). New York: Elsevier.

Gawronski, B., & Strack, F. (2004). On the propositional nature of cognitive consistency: Dissonance

changes implicit but not explicit attitudes. *Journal of Experimental Social Psychology, 40,* 535–542.

Gawronski, B., Strack, F., & Bodenhausen, G. V. (2009). Attitudes and cognitive consistency: The role of associative and propositional processes. In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.), *Attitudes: Insights from the new implicit measures* (pp. 86–117). New York: Psychology Press.

Gawronski, B., Walther, E., & Blank, H. (2005). Cognitive consistency and the formation of interpersonal attitudes: Cognitive balance affects the encoding of social information. *Journal of Experimental Social Psychology, 41,* 618–626.

Gilbert, D. T. (1991). How mental systems believe. *American Psychologist, 46,* 107–119.

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.

Gregg, A. P., Seibt, B., & Banaji, M. R. (2006). Easier done than undone: Asymmetry in the malleability of implicit preferences. *Journal of Personality and Social Psychology, 90,* 1–20.

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.

Harmon-Jones, E., & Harmon-Jones, C. (2002). Testing the action-based model of cognitive dissonance: The effect of action orientation on post-decisional attitudes. *Personality and Social Psychology Bulletin, 28,* 711–723.

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.

Kaplan, M. F. (1972). Interpersonal attraction as a function of relatedness of similar and dissimilar attitudes. *Journal of Experimental Research in Personality, 6,* 17–21.

Karpinski, A., & Hilton, J. L. (2001). Attitudes and the Implicit Association Test. *Journal of Personality and Social Psychology, 81,* 774–788.

Kumkale, G. T., & Albarracín, D. (2004). The sleeper effect in persuasion: A meta-analytic review. *Psychological Bulletin, 130,* 143–172.

McConnell, A. R., Rydell, R. J., Strain, L. M., & Mackie, D. M. (2008). Forming implicit and explicit attitudes toward individuals: Social group association cues. *Journal of Personality and Social Psychology, 94,* 792–807.

Nosek, B. A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General, 134,* 565–584.

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

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

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

Petty, R. E., & Briñol, P. (2009). Implicit ambivalence: A meta-cognitive approach. In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.), *Attitudes: Insights from the new implicit measures* (pp. 119–164). New York: Psychology Press.

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

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

Petty, R. E., & Wegener, D. T. (1998). Attitude change: Multiple roles for persuasion variables. In D. T. Gilbert, S. T., Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 1, pp. 323–390). New York: McGraw-Hill.

Rudman, L. A., Ashmore, R. D., & Gary, M. L. (2001). "Unlearning" automatic biases: The malleability of implicit prejudice stereotypes. *Journal of Personality and Social Psychology, 81,* 856–868.

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

Rydell, R. J., McConnell, A. R., & Mackie, D. M. (2008). Consequences of discrepant explicit and implicit attitudes: Cognitive dissonance and increased information processing. *Journal of Experimental Social Psychology, 44,* 1526–1532.

Rydell, R. J., McConnell, A. R., Mackie, D. M., & Strain, L. M. (2006). Of two minds: Forming and changing valence inconsistent attitudes. *Psychological Science, 17,* 954–958.

Rydell, R. J., McConnell, A. R., Strain, L. M., Claypool, H. M., & Hugenberg, K. (2007). Implicit and explicit attitudes respond differently to increasing amounts of counterattitudinal information. *European Journal of Social Psychology, 37,* 867–878.

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

Schwarz, N. (2007). Attitude construction: Evaluation in context. *Social Cognition, 25,* 638–655.

Schwarz, N., & Bohner, G. (2001). The construction of attitudes. In A. Tesser & N. Schwarz (Eds.), *Blackwell handbook of social psychology: Intraindividual processes* (pp. 436–457). Malden, MA: Blackwell.

Seibt, B., Häfner, M., & Deutsch, R. (2007). Prepared to eat: How immediate affective and motivational responses to food cues are influenced by food deprivation. *European Journal of Social Psychology, 37,* 359–379.

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.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119,* 3–22.

Smith, E. R. (1996). What do connectionism and social psychology offer each other? *Journal of Personality and Social Psychology, 70,* 893–912.

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

Smith, E. R., Fazio, R. H., & Cejka, M. A. (1996). Accessible attitudes influence categorization of multiply categorizable objects. *Journal of Personality and Social Psychology, 71,* 888–898.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8,* 220–247.

Walther, E., Nagengast, B., & Trasselli, C. (2005). Evaluative conditioning in social psychology: Facts and speculations. *Cognition and Emotion, 19,* 175–196.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review, 107,* 101–126.

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

Wittenbrink, B., & Schwarz, N. (Eds.). (2007). *Implicit measures of attitudes.* New York: Guilford Press.