



Likes and Dislikes: A Social Cognitive Perspective on Attitudes

Melissa J. Ferguson & Jun Fukukura

The vast majority of articles and chapters about attitudes (including this one) introduce the topic by referencing the famous quote by Gordon Allport that attitudes are “the most distinctive and indispensable concept in contemporary social psychology” (Allport, 1935). Although it might seem a little suspicious that such a claim is offered by the very people who study the construct, it also happens to have some merit. Attitudes have occupied a central place in the annals of social psychological scholarship from the beginning of the last century up through today (e.g., for reviews, see Albarracín, Johnson, & Zanna, 2005; Eagly & Chaiken, 1993), and this is evident given, for example, the citation count of attitudes articles in any search through the literature.

A long history of scholarship denotes the persistent importance of, and interest in, the construct, but it also suggests that there is continued disagreement about issues. Indeed, there have been long-standing debates about many aspects of attitudes, including the most basic question of what they are. In the current review chapter, we offer a social cognitive perspective on “what they are” and describe the latest cutting-edge research and pressing theoretical questions. We try to shed light on what we now know to be reliably descriptive of attitudes, and also on the major outstanding questions.

We begin by commenting on the terminology in this area, and then describe some of the standard methods of measurement. With such

preliminaries out of the way, we turn then to what we consider to be the big questions about attitudes, including how they are generated, how they influence downstream processing and behavior, their stability and contextual dependence, and how they develop in the first place. Along the way we address some of the most central debates in the literature. A social-cognitive perspective means that we pay special attention to the social-cognitive literature on attitudes, which in turns consists of those articles that pay special attention to the cognitive (broadly defined) processes and mechanisms enabling what we talk about when we talk about attitudes. Describing all of the social-cognitive attitudes literature would be an encyclopedic endeavor and outside the scope of this chapter, however, and so we instead focus on studies that are especially illustrative of a theoretical or methodological question, debate, or issue.

WHAT DO WE TALK ABOUT WHEN WE TALK ABOUT ATTITUDES?

The definition of an attitude has – not surprisingly – fluctuated over the last 100 years (e.g., Allport, 1935; Doob, 1947; Osgood, Suci, & Tannenbaum, 1957; Sarnoff, 1960; M. B. Smith, Bruner, & White, 1956; Thurstone, 1931). In the 1950s and 1960s, researchers argued that attitudes should be

understood in terms of the tripartite model, whereby an attitude consists of affect toward the stimuli, beliefs about the object, and behaviors toward the objects (e.g., Rosenberg & Hovland, 1960). This eventually shifted to the current contemporary definition of an attitude as a “psychological tendency that is expressed by evaluating a particular entity with some degree of favor or disfavor” (Eagly & Chaiken, 1993). A great deal of the social-cognitive research on attitudes has assumed the more specific definition put forth by Fazio and colleagues that an attitude is a positive or negative summary evaluation of the corresponding stimulus (Fazio, 1986; Fazio, Chen, McDonel, & Sherman, 1982). This definition differs from the Eagly and Chaiken one by including assumptions about the cognitive architecture underlying an attitude: i.e., an attitude is the summary evaluation that is associatively linked in memory with the object of that evaluation. The association between an attitude and an object can fluctuate in accessibility, which raises numerous testable questions about the precursors to and downstream consequences of attitude accessibility. For example, accessibility of the link between the evaluation and corresponding stimulus relates to characteristics traditionally examined in the attitudes literature, including strength, spontaneous activation, complexity, and certainty (e.g., Fazio & Williams, 1986; see also Fazio, 1990; Fazio et al., 1986).

This social-cognitive definition still leaves room for lots of debate, however, and indeed, researchers have argued over whether an attitude is a hypothetical construct developed by psychologists (e.g., Eagly & Chaiken, 1993; Schwarz, 2007) or corresponds to a specifiable state in the brain (Fazio, 2007). This debate requires more consideration and discussion about how the brain might enable expressions of favor or disfavor, but to cut to the chase we argue that it comes down to whether one wants to equate an attitude as a person’s general tendency across time and situations to respond to a stimulus in a favorable or unfavorable manner (so, in this case, averaging over many distinct brain states, and thus, existing as an average, a hypothetical construct, or a latent variable) or as a particular response to a stimulus (which could be identified as a specific brain state). The implications of these views require a little more theoretical heavy lifting and so we return to the nuances of the definition of an attitude during our discussion of how attitudes are generated.

Attitude as a label

One recurring source of confusion surrounding the terminology in this area is the use of the terms

evaluation versus attitude. Are they the same? Although many researchers use these terms interchangeably, the term attitude carries with it considerable conceptual baggage, whereas the term evaluation does not. Attitudes have been frequently assumed to be stable, and, as mentioned, to reflect psychological tendencies (real or hypothetical) that could potentially influence behavior. Evaluations, on the other hand, have sometimes been understood as behavioral responses to stimuli (the expression of an attitude; e.g., Cunningham, Zelazo, Packer, & Van Bavel, 2007).

People’s likes and dislikes have been studied in various ways across the social sciences, and some terms seem to be confined to particular literatures, making thematic connections across disciplines difficult. The term *preferences* has been used in judgment and decision making and behavioral economics research and is probably more aligned with the common interpretation of the term *evaluation* as a behavioral manifestation of an underlying attitude or set of attitudes. It is also used most commonly to indicate *relative* preferences (or choices). The term *taste* also shows up in the behavioral economics literature and is meant to reflect someone’s strong likes and dislikes.

What is an attitude object?

The term attitude object is used throughout the attitudes literature to refer to the target being evaluated (e.g., Allport, 1935; Bargh et al., 1992; Fazio, 2001; Fazio et al., 1986; Sarnoff, 1960; M. B. Smith, et al., 1956; Thurstone, 1931). It is important to note that although the term object might imply a material, physical thing, it can refer to anything that can be discriminated in psychological experience and so can include for instance, abstract concepts, smells, sounds, and the contents of our mental life more generally (Eagly & Chaiken, 1993). This means that we can consider attitudes as our likes and dislikes toward, well, anything. One interesting area for future research might be to take this wide definitional latitude seriously and test how attitudes toward traditional stimuli (e.g., people, issues, material objects) compare with attitudes towards less traditional ones (e.g., behavioral intentions, goals, emotions; see Ferguson, 2007a) in terms of basic issues such as predictive validity.

HOW WE MEASURE ATTITUDES

The bulk of scholarship on attitudes over the past approximately 100 years has mostly employed

self-report measures wherein the respondent is asked to report her liking or disliking of a stimulus (see Krosnick, Judd, & Wittenbrink, 2005 for a review). There is commonsense appeal to finding out what someone likes or dislikes simply by asking them, and this kind of measure is still frequently used within social cognition research, and used almost exclusively in other social sciences. Self-report measures tend to consist of a Likert unipolar or bipolar scale (e.g., a scale of 1–11), and respondents are asked to circle the number, for instance, that best represents how they feel about the stimulus.

Because the reporting of how we feel about something or someone is generally under our control (e.g., Schwarz & Bohner, 2001), people can provide answers that depart from their actual feelings about the stimulus of interest, perhaps to please the experimenter, or perhaps to present themselves (to others, or to themselves) in a flattering light (see Orne, 1962; Rosenberg, 1969). This becomes especially likely (and thus problematic) when people are asked about other people, things, or issues that are socially stigmatized and tend to elicit normatively socially desirable responses (Dovidio, Mann, & Gaertner, 1989; Jones & Sigall, 1971; Katz & Hass, 1988; McCauley & Stitt, 1978; McConahay, 1986). Researchers' efforts to circumvent such strategic editing increased considerably in the 1970s and 1980s and have resulted in a major theoretical (as well as methodological) shift in terms of what an attitude is thought to be.

Self-report attitude measures can be considered "direct" measures (Bassili & Brown, 2005; Fazio & Olson, 2003; Ranganath, Smith, & Nosek, 2008) because they consist of asking the respondent directly about her or his attitude. Indirect measures, on the other hand, are those that consist of inferring a person's attitude from other sorts of data, whether behavioral or neural. Direct measures tend to be referred to as explicit measures in reference to explicit memory, which involves cases when one is asked to recall or recognize something. In contrast, indirect measures are frequently referred to as implicit under the assumption that such measures depend on implicit memory.

Although the adjective *explicit* seems to adequately fit self-report measures, the term *implicit* may be a misnomer, as one of us has argued previously (Ferguson, 2007b). *Implicit memory* is a term from cognitive psychology and refers to the influence of a memory on a response wherein the person cannot introspectively identify that memory (Roediger, 1990; Squire & Kandel, 1999; Tulving & Craik, 2000). That is, implicit memory in the cognitive sciences always refers to cases where the person is consciously unaware of the memory.

In social cognition work, however, there has been very little, if any (e.g., see Gawronski, LeBel, & Peters, 2007), evidence that implicitly measured attitudes are beyond the reach of one's awareness. Instead, they are usually called implicit because they are spontaneously evoked by the stimulus without the person intentionally evaluating that stimulus.

Many of the current social-cognitive questions concerning attitudes involve a comparison between directly versus indirectly measured attitudes (for a comprehensive review, see De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009), and we address this issue wherever appropriate. Throughout this paper, we use the terms *direct attitudes* and *indirect attitudes* as shorthand for directly versus indirectly measured attitudes. The use of these terms does not imply any specific assumptions regarding representational, process, or system differences between the two, beyond the issue of whether the person is asked (or not) about her or his attitudes.

Indirect measures are based on either behavioral responses or on specific regions of neural activation. Below we describe a few of the most commonly used indirect measures. More detailed descriptions of other indirect measures can be found elsewhere (see De Houwer, 2003; De Houwer & Eelen, 1998; Dovidio, Kawakami, Johnson, & Johnson, 1997; Koole, Dijksterhuis, & van Knippenberg, 2001; Niedenthal, 1990; Nosek & Banaji, 2001; Payne, Cheng, Govorun, & Stewart, 2005).

Evaluative priming

Evaluative priming is measured with a paradigm that was first developed by Fazio and colleagues (Fazio et al., 1986). The paradigm was developed to measure the degree to which people's attitudes are activated spontaneously from memory upon perception of the corresponding attitude objects. In this paradigm, which was modeled after priming work in semantic cognition (Logan, 1980; Meyer & Schvaneveldt, 1971; Neely, 1976, 1977; Posner & Snyder, 1975; Shiffrin & Schneider, 1977), participants were presented on a computer with a series of sequentially presented prime–target pairs of stimuli. The prime appears first for a fraction of a second, and the respondent is not asked to respond to it. The target then appears, and the respondent usually is asked to make an evaluative (Is this a good or bad word?) or lexical (Is this a real or nonsense word?) decision about it. Given certain parameters of the paradigm, it produces evaluative priming such that people are faster to respond to targets when the target and prime are

similarly (vs dissimilarly) valenced. The phenomenon of priming itself appears to be reliable and robust, and has been interpreted as evidence that people spontaneously evaluate the primes, which then influences the readiness with which people can respond to the positive versus negative targets. There are ongoing debates though about the generality of the effect across different kinds of stimuli (i.e., whether attitude strength moderates automatic attitude activation; Castelli, Zogmaister, Smith, & Arcuri, 2004; Chaiken & Bargh, 1993; Fazio, 1993; Krosnick & Schuman, 1988), the type of response task (e.g., Klauer & Musch, 2003; Wentura, 1999, 2000), and the underlying processes and mechanisms (e.g., Fazio, 2001; Ferguson & Bargh, 2003; Klauer & Musch, 2003; Klauer & Stern, 1992; Klinger, Burton, & Pitts, 2000; Wentura, 1999).

The existence of evaluative priming has enabled researchers to use the paradigm as an indirect measure of the evaluations of the prime stimuli (e.g., Fazio et al., 1995; Ferguson & Bargh, 2004; Wittenbrink, Judd, & Park, 1997, 2001). That is, it is possible to compare the degree to which a certain prime (e.g., a photo of a Black face) facilitates responses to the positive and negative targets as a function of different conditions, or as compared with other prime stimuli (e.g., a photo of a White face).

Implicit Association Test (IAT)

The Implicit Association Test (IAT; Greenwald, McGhee, & Schwarz, 1998) was developed by Greenwald and colleagues and captures the ease with which people can associate a particular category (e.g., women) with pleasant or unpleasant stimuli (for reviews, see Fazio & Olson, 2003; Greenwald, Poehlman, Uhlmann, & Banaji, 2009). The respondent is asked to complete two different sorting tasks simultaneously. For one sorting task, the respondent has to indicate whether each of a series of stimuli is pleasant or unpleasant by pressing one of two keys. For the other, the respondent has to categorize each of a series of stimuli (e.g., female and male names) as belonging to one possible category (female) or another (male). The critical part of this measure is that there are only 2 keys for the 4 possible responses, and so the responses are paired together, such that, for example, pleasant and female names share the same response key and unpleasant and male names share the same response key; thus, it is possible to compare how easily the respondent can perform this version of the sorting tasks compared with the reverse pairing. If the person can respond more easily when female names are paired with

pleasant stimuli (versus the reverse), then the inference is that female names are implicitly preferred over male names.

The IAT has been employed to examine a variety of different topics, including self-esteem (Greenwald & Farnham, 2000), prejudice (e.g., Ashburn-Nardo, Voils, & Monteith, 2001; Blair, Ma, & Lenton, 2001; McConnell & Leibold, 2001), social identity (Greenwald, Banaji, Rudman, Farnham, Nosek, & Mellott, 2002), and personality traits (e.g., Jordan, Spencer, & Zanna, 2003; Marsh, Johnson, & Scott-Sheldon, 2001). There is still ongoing debate about the exact mechanisms underlying the effect, as well as the boundary conditions and predictive validity of the measure (e.g., see Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005; Karpinski & Hilton, 2001; Mierke & Klauer, 2003; Nosek, Greenwald, & Banaji, 2005; Olson & Fazio, 2004).

Event-related potentials and functional magnetic resonance imaging

In addition to indirect attitude measures based on behavior, researchers have also inferred people's evaluative processes indirectly by examining what regions of the brain are active while viewing evaluative versus non-evaluative stimuli (e.g., Amodio, Harmon-Jones, & Devine, 2003; Cunningham, Johnson, Raye, Gatenby, Gore, & Banaji, 2004; Ito & Cacioppo, 2000; Phelps, O'Connor, Cunningham, Funayama, Gatenby, Gore, & Banaji, 2000). For instance, Ito and Cacioppo (2000) recorded event-related brain potentials (ERP) and showed that people are sensitive to the evaluative nature of stimuli even when they are not intentionally evaluating those stimuli. Participants were asked to decide whether various kinds of stimuli (e.g., a chocolate bar, a couple hugging) included people or not. The focal task was to report whether people were present or absent, and in this way did not involve any evaluative processing per se. The results showed that whenever a series of stimulus involved evaluative inconsistency (positive and negative stimuli) versus consistency (only positive or only negative stimuli), there was an increase in participants' electroencephalographic activity, suggesting greater cognitive effort.

Researchers have also used functional magnetic resonance imaging (fMRI) to identify the specific regions of the brain that become active during evaluation. For example, the amygdala seems to be particularly active when normatively negative stimuli are processed (e.g., LeDoux, 2000;

Phelps et al., 2000), and this happens even when the stimuli are processed outside of awareness (e.g., Cunningham, Raye, & Johnson, 2004; Morris, Ohman, & Dolan, 1998). Cunningham and colleagues (Cunningham et al., 2003) have also found increased amygdala activity in response to negative versus positive stimuli, both when participants are intentionally evaluating those stimuli and even when they are not. This work also shows that this negativity occurs especially in the very early milliseconds of processing negative stimuli (30 ms), compared with relatively later in the processing stream (525 ms).

Research also shows that irrespective of whether a person is intentionally evaluating stimuli, the perception of normatively negative versus positive stimuli leads to greater activity in the amygdala and right inferior prefrontal cortex (PFC; e.g., see Cunningham et al., 2003). When people are intentionally assessing the valence of stimuli, there is greater activity in the medial and ventrolateral PFC, especially when the stimuli are evaluatively complex (Cunningham et al., 2003; Cunningham, Raye, & Johnson, 2004). These neuroscience methods have provided information that converges with the behavioral measures to suggest that evaluation seems to be a pervasive, easily triggered process. However, it also pinpoints the regions of the brain that are involved in different kinds of evaluation, which can then inform and constrain theorizing about evaluation more generally (Cunningham, 2010).

HOW ARE ATTITUDES GENERATED?

There are numerous assumptions in the current attitudes literature about how attitudes are generated. These assumptions involve the issues of representation, process, and system, and we address each one here. Throughout this discussion, the importance in the literature of indirectly versus directly measured attitudes will be apparent. The two types of measures are commonly assumed to tap different representations, processes, or systems, or all three. Issues about how attitudes are generated in the brain are at the center of a social-cognitive perspective.

Representation

What is a representation? A popular definition in the social psychological literature is that mental representations reflect information stored in memory. But, it is important to note that this construct is the subject of a historically long-standing

and intense debate among cognitive scientists, with many varieties of theoretical stances that are not normally seen in the pages of social psychological journals. The only thing strongly agreed upon in the cognitive science literature about representation is that there is no agreement about representation (e.g., Barsalou, 2009; Dietrich & Markman, 2003; Dretske, 1995; Haugeland, 1991; Palmer, 1978; Markman & Dietrich, 2000). Scholars differ in their opinions about the content, format, and architectural nature of mental representations. So, the most basic nature of a representation is still open for debate, both conceptually and empirically.

Some of the questions from this debate have made their way into the literature on attitudes (as well as on other social psychological topics; see, e.g., Carlston & Smith, 1996; Smith & Conrey, 2007). The most common question is whether (or, when) representations are discrete (symbolic/amodal) versus continuous (distributed/modal). Discrete representations are often defined as non-overlapping representations that are symbolic and are separate from the brain's modal systems for perception, action, and internal states (i.e., interoception, introspection, and emotion; see Barsalou, 2009). They are usually assumed to symbolize meaningful information, such as concepts (see Dietrich & Markman, 2003 for more detailed discussions). On the other hand, distributed representations are considered to be modal in that they are generated using the same mechanisms as in the brain's modal systems. They are commonly defined as patterns of activation among (i.e., distributed across) many units of processing (e.g., neurons), wherein the units are not themselves meaningfully correspondent with information in the same way as are discrete representations. A popular analogy to illustrate distributed representations is the television monitor, consisting of many pixels that individually take on only simplistic variations (e.g., binary) of information such as color, but together interactively provide an astoundingly large (though not infinite, in this case) number of possible images.

The two types of representational formats each have their strengths and weaknesses. Distributed representations possess relatively more biological plausibility in that it is possible to conceive of how they would be implemented by the brain via neurons and populations codes. This strength of distributed representations – biological plausibility – is in turn one of the major weaknesses of discrete representations. There are few ideas about how discrete representations could be implemented in the brain (though, see e.g., Devaney, 2003).

Distributed representations have successfully explained a wealth of lower-order phenomena in,

for example, perception, language, and categorization (e.g., Elman et al., 1996; Grossberg et al., 1997; McClelland & Rogers, 2003; Seidenberg & McClelland, 1989; E. Smith, 2009; Spivey, 2007; Thelen & L. Smith, 1994). Network models assuming distributed representations have even been able to explain dissociation data that previously were thought to provide strong evidence for two qualitatively distinct processes (see Spivey, 2007). In the social psychological literature, a number of researchers have argued that distributed representations are implicated in attitude generation (Bassili & Brown, 2005; Ferguson, 2007b; Gawronski & Bodenhausen, 2006; Mitchell et al., 2003; Smith & Conrey, 2007; Smith & DeCoster, 2000; Strack & Deutsch, 2004). In particular, theories on attitudes almost uniformly hold that indirectly measured attitudes are reflective of distributed representations.

The strength of discrete representations, on the other hand, is precisely the weakness of distributed representations. Namely, multiple theorists have argued that basic functions of cognition would just not be possible without discrete representations (see Carey, 1985; Dietrich & Markman, 2003; Markman & Dietrich, 2000; Keil, 1989; Marcus, 2001; Spivey, 2007; cf. Thagard, 1991; Van Overwalle & Van Rooy, 2001). For example, Dietrich and Markman (2003) argue that categorization, reasoning, and decision making all require discrete representations. They argue that although distributed representations might be able to explain or reproduce data from lower-level processing, they would be unable to reproduce/explain data from higher-order cognition. And, in fact, there is very little evidence to dispute that claim. Although some social psychologists have started to apply models assuming distributed representations to higher-order cognition, this research is in its infancy. Critically, it is noteworthy that some of the fiercest advocates of distributed representations acknowledge that such a representational format may never be able to explain some higher-order cognition (Spivey, 2007, pp. 284–285). Attitude theorists have tended to imply (though usually not explicitly) that in addition to distributed representations being involved in attitude generation, discrete representations are involved as well. The assumption is that indirect attitudes draw on distributed representations, while direct attitudes draw on both distributed as well as discrete representations (e.g., see Gawronski & Bodenhausen, 2006; Smith & Conrey, 2007; Strack & Deutsch, 2004).

Some cognitive scientists concede that there are probably both types of these representations, along with other variations in format (Barsalou, 2009; Dietrich & Markman, 2003; Jilk, Lebiere, O'Reilly, & Anderson, 2008) and then the

question becomes *where* (in the brain) and *when* these representations play a role in cognition (broadly defined). For the attitudes literature then, these questions about discrete versus distributed representation need to be addressed. Although the current consensus is that indirectly measured attitudes reflect distributed representations, as noted, there is very little evidence for, and powerful philosophical arguments against, the notion that distributed representations could explain conscious, higher-order cognition. And, if indirect attitudes involve this sort of high-order cognition, then this poses a problem for theory concerning this issue. Do indirect attitudes involve high-order cognition? This depends on one's definition of "higher-order" cognition, but, at the least, there is considerable neural connectivity between limbic structures typically involved in immediate affective responding and cortical regions that are traditionally assumed to underlie higher-order cognition-like decision making and goal pursuit. Moreover, some indirect attitude measures are correlated with controlled processes (Klauer & Mierke, 2005; Klauer, Schmitz, Teige-Mocigemba, & Voss, 2010; Payne, 2005), which are traditionally assumed to be integrally involved in higher-order cognition.

For theorists who assume discrete representations, there is little theoretical development about how such representations could be biologically implemented in the brain, as noted. Furthermore, a major challenge to any theorist who assumes both types of representation is how they interact. There is some recent speculation that the mathematics involved in dynamical systems (which assume distributed representations) can handle discrete, or binary, decisions. This sub-field has been dubbed "symbolic dynamics," but its development is occurring outside of the social psychological literature (Devaney, 2003). In sum, any glance at the cognitive science literature on representation reveals some serious questions about representation and these debates challenge current assumptions about the types of representations underlying attitudes, or any social psychological construct or process.

Process

The definition of process depends on one's view of representations. For advocates of distributed representations, the representation is the process (e.g., see Conrey & Smith, 2007). A distributed representation is the pattern of activation across simple processing units that are interconnected by weighted functions. So, the activation of a representation from this perspective is the way in which

the representation is processed. In other words, it is not necessary to postulate any external, orthogonal processes to act on the distributed representations. But, for discrete theorists, process is something external to the representation. And, although process is rarely defined, when it is defined in the social psychological literature it usually is meant to imply a transformation of mental representations. Its meaning is more fleshed out when it is used to compare associative or rule-based processing – two modes of thought that have received a lot of attention in both cognitive and social psychology (e.g., see Chaiken & Trope, 1999). Associative processing means processing that is based on the statistical covariance (in space or time) of stimuli. That is, through Hebbian learning (Hebb, 1949), the association in memory between stimuli strengthens as those stimuli are experienced close together in space or time. This kind of computation is used for predicting weather systems and the stock market, for instance, and is invaluable in the generation of responses to stimuli that are likely given past experience. This kind of processing is usually characterized as relatively fast, effortless, non-conscious, uncontrollable, and spontaneous. Associative processing is very similar to automatic processing, and the two terms are often used synonymously in the social psychological literature (see Moors & De Houwer, 2006). This type of processing is largely assumed to recruit (or, be synonymous with) distributed representations. Research over the last two decades shows that attitudes can be activated in memory in response to a wide array of stimuli, and under conditions of limited processing, awareness, intention, and control (e.g., see Ferguson & Zayas, 2009; Wittenbrink & Schwarz, 2007). This sort of evidence supports the notion that indirect attitudes are produced through associative or automatic processing.

Associative processing is often contrasted with rule-based processing. This type of processing is assumed to follow rules, where rules are abstract statements about the logical relations between variables (see Sloman, 1996). Rule-based processing is assumed to actually proceed according to steps or procedures of rules, rather than just being able to be described by rules (associative processing can be described by rules even if it does not operate according to them). For instance, rule-based processing is assumed to characterize thinking about probability and logical reasoning. Rule-based processing is often characterized as slow, effortful, conscious, controllable, and intentional. It is often synonymous with “controlled” processing in the social psychological literature. This type of processing is assumed to recruit both distributed as well as discrete representations. Direct attitudes are assumed to be more flexible,

changeable, and context-dependent due to the controlled nature of the underlying processing (e.g., see Wilson et al., 2000).

Most theories state, tacitly imply, or are consistent with the notion that associative processing underlies indirect attitude measures, while both associative and rule-based processing underlie direct attitude measures (Gawronski & Bodenhausen, 2006; Smith & Conrey, 2007). And yet, the notion that these two types of attitude measures tap into distinct cognitive processes has already been challenged by work showing that multiple processes underlie any measure (Jacoby, 1991; Payne, 2005; Sherman, 2009; Sherman, Gawronski, Gonsalkorale, Hugenberg, Allen, & Groom, 2008). Work on the quad model, in particular, shows evidence for four distinct processes, which demonstrates that there are at least more than two and probably more than four (Sherman, 2009). There is still the question of what (multiple) processes underlie attitude generation during indirect versus direct measurement, and there are likely interesting differences. After all, the behavior captured by these two types of measurement often differs in terms of intentionality, speed, and effort (see De Houwer et al., 2009). The degree to which the measures differ in awareness is still an open empirical question (see Gawronski et al., 2007). But, these different characteristics suggest differences in the underlying processes and at this point it is unclear how many processes are operating, and when. At the least, the evidence suggests strongly that there are more than two, and that direct versus indirect measures do not map exclusively onto different processes.

System

Although the constructs of process and system are often used interchangeably, they are different conceptually and empirically, and arguably exist at different levels of analysis (Keren & Schul, 2009). A system can be classified according to the information it acts on (input), the processes or rules of operation that transform that information, and the accompanying neural substrates (Schacter & Tulving, 1994; for alternative definitions see e.g., Bechtel, 2008; Sperber, 2005; Tooby & Cosmides, 2005).

The notion that there are two systems of cognition has a long tradition, and is closely related to the proposal of two different kinds of processes. There are numerous dual system models in the social psychological literature, and the characteristics and functions of these systems vary (Chaiken & Trope, 1999), which means that either there are considerably more than two systems, or the

characteristics of the systems are not correct (see Gilbert, 1999). There have been some recent critical papers on dual-system models in social psychology (Evans, 2008; Keren & Schul, 2009). Keren and Schul (2009) point out the conceptual and empirical vagueness with which dual systems tend to be defined and operationalized (and tested), and argue that the notion of duality in the mind more generally requires considerably more empirical and conceptual support than exists currently in the literature.

Dual systems have been proposed for reasoning (e.g., Kahneman & Frederic, 2002; Sloman, 1996), self-regulation (Metcalfe & Mischel, 1999), persuasion (Chaiken, 1980; Petty & Wegener, 1999), attitudes (Rydell, McConnell, Mackie, & Strain, 2006), and affect and emotion (e.g., Epstein, 1994), among others. The attitudes literature in particular though seems to rely most strongly on the assumption that there are separable systems of *memory*. There is an extensive literature on the assertion that explicit and implicit memory are separate systems, with considerable supporting evidence (McClelland, McNaughton, & O'Reilly, 1995; O'Keefe & Nadel, 1978; Schacter & Tulving, 1994; Sherry & Schacter, 1987; E. Smith & DeCoster, 2000; Tulving, 1983). There are also convincing arguments for the evolutionary development and need for separable memory systems. For example, Sherry and Schacter (1987) discuss how human and non-human animals have memory needs that are functionally incompatible. Birds, for example, need memory to enable frequently revised food cache locations as well as revision-impervious song learning. These two needs would seem to pose distinct, functionally incompatible memory capacities and may have prompted the development of separate memory systems. Sherry and Schacter also discuss how primates and humans likely have needs such as habit forming and episodic memory that would similarly seem incompatible. Although there is recently some work challenging the notion of separable memory systems (e.g., Berry, Shanks, & Henson, 2008), there is generally consensus behind the idea given empirical evidence and functional arguments.

Do attitudes differentially rely on these separable memory systems? The assumption seems to be, as the terms implicit and explicit attitudes suggest, that indirect attitudes tap implicit memory while explicit attitudes tap explicit memory. There are multiple purported differences between explicit and implicit memory in the cognitive literature, but the difference that emerges in the pages of the attitudes literature most commonly is the ease with which implicit versus explicit memories are learned and can be revised. Whereas explicit memory allows for fast learning, implicit memory allows for slow learning. The assumption that

indirect attitudes exhibit slow learning is frequently assumed in the literature. This is despite the fact that there are several notable exceptions to this assumption that show that indirect attitudes can form relatively quickly (Ashburn-Nardo et al., 2001; Castelli et al., 2004; De Houwer et al., 1998; Gregg, Seibt, & Banaji, 2006). For instance, Ashburn-Nardo et al. (2001) used a minimal group paradigm to assign participants to one of two different artists. Participants were told that they showed a preference for either the artist Quan or the artist Xanthie. They then completed an IAT testing their preference for others who prefer Quan or Xanthie with the understanding that names that contained a *q* reflected Quan fans while names with an *x* indicated Xanthie fans. Although participants developed their preference only moments earlier, and on the basis of novel information, they showed a significant preference for fans of the artist to which they had been assigned. This shows relatively fast learning. Even stronger evidence for fast learning comes from Gregg, Seibt, and Banaji (2006). In this paper, participants learned about two novel groups of people, one of which was described as good and the other evil. They learned this information either quickly and abstractly (e.g., read that one group was good and the other bad) or concretely (e.g., read lots of detailed information about the character of the two groups). They then completed an IAT to measure their preference for the two groups, and participants showed a significant preference for the good group, regardless of whether they had learned about the groups in an abstract or concrete manner. This again shows that indirect attitudes can respond to relatively fast learning. These examples strongly contradict the assumption in the current attitudes literature that indirect attitudes operate exclusively on the basis of a slow-learning implicit memory system.

However, Gregg et al. (2006) then tested whether newly formed indirect attitudes would be able to be revised with new (countervailing) information. Participants who had learned about the two novel groups then learned that the two groups actually had the opposite character. They learned this in either an abstract manner ("suppose the two groups were switched in character") or in a concrete manner (they read a long and detailed explanation for how the good group eventually turned bad, and how the bad group eventually turned good). They found that although direct attitudes tracked the instructions and were revised accordingly, indirect attitudes were largely resistant to change. Gregg et al. reasoned that indirect attitudes may be like perceptual defaults in that once an evaluation about a stimulus has been learned (perhaps quickly, or through conscious appraisal), it then becomes largely resistant to change thereafter.

This latter conclusion – that indirect attitudes are “stuck” once they have been formed – is consistent with the general assumption that they are unable to be revised quickly, and in line with assumptions about implicit memory underlying indirect attitudes. Still, these findings concerning the fast formation of indirect attitudes challenge the widely adopted view that indirect attitudes rely exclusively on implicit memory. The findings would seem to pose a challenge either for this claim, or for the claim that implicit memory is slow learning. The more parsimonious implication is that indirect attitudes do not rely exclusively on implicit memory.

In addition to assuming differences between indirect and direct attitudes in terms of the speed of learning, researchers have also argued that they differ in the types of information they are influenced by *during* learning (Rydell et al., 2006). Rydell and colleagues have argued that if indirect attitudes are reliant on implicit memory and associative processes, they should be especially sensitive to subliminally presented information that is activated in close proximity with the novel stimulus. Direct attitudes, on the other hand, should be more sensitive to verbal information that is described as being about the novel stimulus. They tested this by presenting participants with information about a novel target named Bob. They presented many stimuli pairs consisting of a picture of Bob and behavioral information about Bob. The behavioral information was positive or negative in valence, and was described as being characteristic or not characteristic of Bob. Immediately before each presentation of the picture of Bob, however, highly positive or negative words were also subliminally presented. The valence of the subliminal information was always in opposition to the valence of the verbally presented information, and Rydell et al. examined how this learning paradigm would influence participants’ indirect and direct attitudes toward Bob. They found that direct attitudes were sensitive to the verbally presented information, while indirect attitudes were sensitive to the subliminally presented information. They concluded that these two different types of attitude measures capture different attitudes in memory, form within different systems, and consist of different representations.

Although the results of the Rydell et al. paper are intriguing, they raise multiple questions. First, if learning evaluative information about novel stimuli is modularized in this way, then other findings from other lines of research are puzzling. For instance, research on mere exposure has found that different frequencies in the exposure to subliminally presented stimuli lead to differences on direct attitude measures. Similarly, research on subliminal evaluative conditioning has also found

effects on direct attitude measures. As for whether indirect attitudes are sensitive to verbally presented information, although there is little work on the formation of indirect attitudes, the previously described studies by Ashburn-Nardo et al. (2001) and Gregg et al. (2006) are two examples against it. Thus, it is not clear how the interpretation of the dissociation in the Rydell et al. paper can be squared with other work. It may be the case that indirect and direct attitudes differ in the kinds of information that will exert an influence under some circumstances, and these circumstances need to be identified. However, if indirect attitudes can be revised or formed by exposure to verbally presented information and conscious reasoning, for instance, it would imply that the two types of attitude measures do not map exclusively or consistently into distinct memory systems as is assumed in the attitudes literatures.

We have already described how indirect attitudes depend on multiple processes that have been characterized as controlled, and they also may be conscious (unlike implicit memory as it has been operationalized in cognitive science), and so there would seem to be multiple reasons to assume that indirect and direct attitudes do not map onto implicit and explicit memory systems. Given the preceding discussion, we now turn to some of the most central debates within the attitudes literature and argue that some of them depend on assumptions about representation, process, system, or some combination of these issues.

DISSOCIATION DATA ACCORDING TO MEASUREMENT

If indirect and direct attitudes each consist of discrete and distributed representations, each rely on an assortment of processes, and each potentially involve implicit as well as explicit memory, what exactly explains the fact that they occasionally diverge? There are multiple reasons why attitudes from these two different types of measure might be dissociated (see De Houwer et al., 2009; Nosek, 2005; Wittenbrink & Schwarz, 2007). First, even though the previous discussion suggested that they might not map exclusively onto different types of representations, processes, or systems, this does not mean that they involve the same proportion or extent of representations, processes, or systems. They may differ in the degree to which they rely on implicit versus explicit memory, or the degree to which they involve multiple automatic processes, for instance. So, the preceding discussion does not preclude the possibility that these measures are tapping different kinds of

informational formats (representations), operate according to different kinds of rules (sets of processes), and implicate different neural substrates (systems). This remains a wide-open empirical question. Any of these possibilities could potentially explain dissociation data.

Alternatively, researchers have identified reasons for divergence in data from these two types of measures that do not necessarily point to underlying (substantive) differences in type of representation, process, or system. There are differences in the structural fit of the two tasks, method, task, and instruction differences, response bias, and presentation norms (e.g., Olson, Fazio, & Hermann, 2007; Payne, Burkley, Stokes, 2008; Smith & Nosek, 2011). These would seem to be superficial reasons for dissociation, but it is important to note that some processes, for example, may be triggered by particular task constraints, and so differences in methodology do not necessarily preclude differences in underlying cognitive process (nor do they necessarily point to them).

One obvious difference between the measures is the time that elapses between the perception of the stimulus and the respondent's behavior toward that stimulus. In an indirect measure it is often a matter of a hundred milliseconds or less, whereas in direct measures it can be multiple seconds. This is a considerable difference in cognitive processing time, and it is important to appreciate how cognition can vary over such a time lapse. An object and its associated attitude are not activated in memory and then frozen – they likely undergo repeated iterations in processing that recruit other information, with the influence of some kinds of information waxing or waning, etc. Thus, one important difference is the point in time that the measure captures the dynamic processing stream that is constantly changing (e.g., Cunningham, Zelazo, Packer, & Van Bavel, 2007; Fazio, 1995; Lamme & Roelfsema, 2000; O'Reilly, 1998; Spivey, 2007). This is not to say that we cannot hold a piece of information in mind for a spell, to concentrate on it, or keep it in working memory. We can, but it is not easy (e.g., Smallwood & Schooler, 2006). The mind does not stand still (see Spivey, 2007). And, the processing that likely accompanies attitude measures is constantly dynamically evolving. This is consistent with the view that some of the same representations or processes underlying indirect measures may be exerting an influence also during explicit measurement (e.g., see Cunningham et al., 2007; Fazio, 1995; Wojnowicz, Ferguson, Dale, & Spivey, 2009). A complete disconnect between the representations and processing underlying each type of measure would of course suggest orthogonal constructs, and the data to date do not support this.

STABILITY VERSUS CONTEXTUAL DEPENDENCE

One persistent debate is the extent to which attitudes are trait-like and stable or constructed on the spot, and thus heavily dependent on the context. There are two different levels at which this issue can be analyzed: one level concerns questions of generation in memory, and the other level concerns observed data. In fact, conflating these two levels leads to all kinds of confusion. Let's begin with the level of observed data. Do attitudes seem to persist regardless of time and context? Does someone's liking of pizza, for instance, emerge regardless of whether she is hungry or bored? As we review below, there is a lot of evidence both for stability across time as well as for sensitivity according to the context.

How do researchers tend to interpret this kind of evidence? Their explanations depend on their theory about how attitudes are generated, or what attitudes reflect. One view is that attitudes are relatively stable, trait-like constructs (e.g., Dovidio, Kawakami, & Beach, 2001; Petty et al., 2006; Rudman, 2004; Wilson et al., 2000). Fazio and colleagues (Fazio, 2007), for instance, have argued that many stimuli have summary evaluations associated with them in memory, and that a summary can be formed from many kinds of sources (feelings, beliefs, behaviors). Once this summary evaluation is formed, it can be relatively stable, especially if the summary is well learned, or highly accessible. From this view, a stimulus would be evaluated in roughly the same way regardless of time and context. Any fluctuation in observed attitudes across time and situations would simply mean that the stimulus was categorized in different ways across time and situations. Thus, attitudes – especially strong attitudes that can be activated spontaneously – are generally stable and durable across time. Evidence for fluctuation across time or situation is simply evidence that the object of judgment, rather than the judgment of the object (Asch, 1948, p. 256), is changing. This view would seem to assume that attitudes are discrete representations. Or, in the parlance of social psychological terminology, attitudes would be considered as abstract, generalized prototypes (summaries) of evaluative knowledge concerning stimuli.

The view that attitudes are instead constructed on the spot has been understood as a contrast to the traits/stability view (e.g., Bassili & Brown, 2005; Schwarz, 2007; Schwarz & Bohner, 2001; Tesser, 1978; Wilson & Hodges, 1992; see also Gawronski & Bodenhausen, 2006), and has also gained popularity over the years, for example in the preferences literature (for a review, see

Lichtenstein & Slovic, 2006). What does “constructed on the spot” mean exactly? The only logically possible definition is that for each stimulus, some computation (or integration) is performed across stored sources of evaluative knowledge and information. In this way, there are stored memories about what is good or bad, and these are combined in some way to predict a particular person’s particular response to a particular stimulus at a particular time. This view seems most harmonious with the notion of distributed representations (e.g., Ferguson, 2007b), or the notion that any given attitude reflects the online computation across many specific similar instances (i.e., exemplar models of cognition).

One important question for this perspective concerns the nature of the computation or integration (see Betsch, Plessner, Schwieren, & Gütig, 2001). Additionally, some have questioned how such a constructivist process could be economical, given that it would mean that a person would have to “reconstruct” something in order to know whether it is good or bad (see Fazio, 2007; Wilson et al., 2000). However, such “recomputation” does not necessarily need to be effortful, and, in fact, construing an object on the basis of its salient features and the attendant evaluative implications is selectively narrowing down all the available information about the object and presumably making its resulting evaluation more precise, and computationally tractable. Furthermore, there is a great deal of evidence that non-evaluative object knowledge is context-dependent (e.g., see Barsalou, 2008, 2009; Yeh & Barsalou, 2006), so why would evaluative knowledge be computed differently?

After a closer look, however, one might ask how these two views concerning stability and contextual dependence are actually different. There seems to be more agreement between them than might be evident at first glance. First, the two views agree that attitudes reflect stored knowledge. According to the trait/stable view, the observed attitude is the stored summary evaluation. According to the constructivist view, the observed attitude is computed from different sources of stored evaluative information. It is also true that both views assume that the context influences the attitude. The trait/stable view assumes that the context influences the interpretation (or construal, or categorization) of the stimulus, which then (neatly) determines the attitude that is activated. This still means that attitudes are dependent on the context; it is just that the context is influential at the level of object construal, and once the object has been categorized, the attitude presumably follows from that without further interference from the context. In the constructivist view, the attitudes themselves are dependent on the context.

Critically, this of course means that both views are also entirely consistent with stability and variation in observed data across time and situations. If the many elements comprising the context in which a stimulus is encountered are uniform in valence, then the observed attitude could be highly stable, even though it (and/or the stimulus) would still be constructed according to the nature of the context (see Gawronski, Rydell, Vervliet, & De Houwer, 2010). This means that any amount of evidence for contextual dependence cannot really adjudicate between these views.

Instead, the real point of difference is in what the observed attitude reflects. The trait/stable view maintains that the attitude reflects a stored summary evaluation, and the constructivist view argues that it reflects an integration of evaluative information across various sources. These views are suggestive of different types of cognitive architectural commitments. The trait view would seem to assume discrete representations (though see Fazio, 2007), as well as a separation in processing between object recognition or categorization, and attitude activation. And, the constructivist view would seem to be most consistent with distributed representations, although theoretically a computation could operate across discrete representations (see Ferguson, 2007b).

This discussion makes it apparent that the two views are not really that different in terms of the question of whether the context matters, and whether stored knowledge is implicated in attitudes. What is not clear are the assumptions underlying these two views concerning representation, process, and system. We have already identified the many open questions with respect to these issues. These two views therefore require conceptual discussion about representation, process, and system, and cannot be decided simply from evidence of contextual variability in observed attitudes. As Schwarz has noted, deciding whether context variability represents a latent variable of an attitude with noise, or different evaluative responses, is not possible to answer empirically.

Moderators of stability and context dependence of direct attitudes

What is the evidence regarding stability versus context dependence? Demonstrating attitude stability is complicated in that an empirical demonstration would require null results, which are more difficult to interpret than results that reflect a significant difference. Thus, often the most memorable and most cited research on attitudes is when relatively small changes in the context have strong effects on attitude change (e.g., Wittenbrink

et al., 2001). Although researchers may not be able to quantify the precise degree to which attitudes are stable versus context dependent, they can identify the moderators that predict when attitudes are more or less likely to be stable versus context dependent (see Petty & Krosnick, 1995).

Although attitude strength would seem to be a likely candidate as a predictor of attitude stability, there is some debate on the relationship between attitude strength and attitude stability. Researchers have identified about a dozen strength attributes of attitudes, including attitude importance, knowledge, elaboration, certainty, ambivalence, accessibility, intensity, extremity, structural consistency, (see Petty & Krosnick, 1995). Across 27 experiments conducted in national surveys, Krosnick and Schuman (1988) showed that none of three strength-related attitude properties (importance, intensity, and certainty) was found to reliably moderate the impact of question variations known to produce response effects, such as question order, wording, or form of questions.

In contrast, in a study done by different researchers, attitude importance, elaboration, certainty, extremity, ambivalence (reverse scored), and intensity, were each shown to each be correlated with resistance to attitude change (Lavine, Huff, Wagner, & Sweeney, 1998). When each attribute was examined individually, Lavine et al. found strong and significant context effects for target issues among participants whose attitudes were highly ambivalent, low in prior elaboration, low in attitude certainty, low in attitude extremity, and low in intensity.

Although it is not entirely clear why seemingly conflicting findings have emerged, there are some key differences in the methodology that may have led to the differing results. Visser, Bizer, and Krosnick (2006) have suggested that Lavine and his colleagues' findings are limited to a specific kind of question order context effect. In contrast, Lavine et al. (1998) have argued that Krosnick and his colleagues' findings may be particular to the attitude issue and that attitude strength measures must be sufficiently broad to demonstrate an effect on attitude stability. Taken together, the research indicates that particular response effects are affected by different attitude attributes and that these effects may also interact differently depending on the attitude object.

There may also be chronic individual differences in attitude stability. Individual differences can also moderate attitude stability. People high in the *need for affect*, the tendency to become involved in emotion-inducing situations (Maio & Esses, 2001), versus people high in the *need for cognition*, the tendency to engage in and enjoy effortful cognitive activity (Cacioppo & Petty, 1982), maintain different degrees of attitude stability in the face of different kinds of persuasive

messages. Huskinson and Haddock (2004) found that people who are high in the need for affect are more persuaded by affect-based appeals than are people high in the need for cognition. In contrast, those high in the need for cognition find cognitive, rather than affective, appeals more persuasive.

Moderators of stability and context dependence of indirect attitudes

One of the central questions guiding research on indirect attitudes over the past 10 or so years has been the extent to which such evaluations are stable across time and contexts. Initially, attitudes measured by implicit methods were thought to be a potentially stable measurements precisely because they were assumed to be independent of the context in which they were measured (e.g., Banaji, 2001; Bargh et al., 1992, 1996; Devine, 1989; Fazio et al., 1995; Greenwald et al., 1998; Wilson & Hodges, 1992) unlike the contextual influences inherent in direct attitude measurement (see Banaji, 2001; Fazio et al., 1995; Schwarz & Bohner, 2001). Because indirect measures were assumed to assess participants' evaluations without their awareness (cf. Gawronski et al., 2007), participants were assumed to be unable to strategically modify their responses (for a review see Ferguson, 2007b).

As data on direct and indirect attitude measures accumulated, evidence began to emerge that indirect attitudes are often weakly correlated or completely unrelated to direct attitudes (e.g., Fazio et al., 1995; Greenwald & Banaji, 1995; Karpinski & Hilton, 2001; though, see Cunningham, Preacher, & Banaji, 2001; McConnell & Leibold, 2001; Wittenbrink, Judd, & Park, 1997). Although some degree of disconnect between the two types of measures is expected, given the differences in the nature of the measures, a complete lack of correspondence worried researchers, and some questioned the construct validity of indirect attitudes and evaluations (see Banaji, 2001 for a discussion). If indirect measures are tapping people's "true" attitudes and preferences (Fazio et al., 1995), then they should at least partially correspond with related measures in some situations, in line with basic conventions regarding convergent and criterion validity. This concern provoked considerable research efforts at examining the stability and contextual independence of indirect attitudes and the relation between indirect and direct measures in general (e.g., for a review see Blair et al., 2001; Dovidio et al., 2001; Fazio & Olson, 2003; Nosek, 2005).

The findings suggest many contextual influences on indirect attitudes, contrary to the initial

assumptions of contextual independence. Specifically, findings suggest that the direction and strength of an indirect attitude toward a given object vary, depending on the type of recently activated, or repeatedly learned, object-relevant information (e.g., Dasgupta & Greenwald, 2001; Karpinski & Hilton, 2001; Lowery, Hardin, & Sinclair, 2001; Mitchell et al., 2003; Wittenbrink et al., 2001). For instance, researchers found that participants displayed significantly less negative indirect attitudes toward group members who are commonly targets of prejudice after being exposed to pro-elderly stimuli (Karpinski & Hilton, 2001) and exemplars of well-liked African Americans and disliked White people (Dasgupta & Greenwald, 2001).

Recent work has shown that indirect attitudes are also influenced by the goal that the perceiver is currently pursuing (e.g., Ferguson, 2008; Ferguson & Bargh, 2004; Lowery et al., 2001; Moore, Ferguson, & Chartrand, 2011; Moors & De Houwer, 2001; Sherman et al., 2003) and by the perceiver's chronic motivations (Maddux, Barden, Brewer, & Petty, 2005). For example, Lowery et al. (2001) found that participants who completed an IAT administered by a Black (vs White) experimenter exhibited significantly reduced negative attitudes toward Blacks, demonstrating that indirect attitude measures may be susceptible to social influence pressures. Ferguson and Bargh (2004) investigated how the extent to which a goal is completed affects indirect attitudes. They found that participants who were currently pursuing a goal (or not) completed an evaluative priming paradigm that measured their indirect attitudes toward objects that varied in their relevance to the goal. The results suggest that objects that were relevant to the goal were evaluated as most positive when the perceiver was still pursuing the goal versus had already completed it. For example, participants who were thirsty evaluated the highly thirst-relevant objects (e.g., water, juice) as more positive, than other objects (e.g., chair, table), compared to participants who had just quenched their thirst. These findings demonstrate that indirect evaluations are sometimes prospective with regard to the utility of the objects, as opposed to solely a function of recent experience with the objects.

People's chronic goals can also influence indirect attitudes. Maddux et al. (2005) demonstrated that the impact of contextual cues on participants' indirect attitudes toward Black people depended on participants' chronic motivation to avoid being prejudiced. Participants low in this motivation exhibited negative attitudes toward Blacks in contexts that were threatening (e.g., a prison cell) compared to non-threatening (e.g., a church). In contrast, participants high in the motivation to avoid being prejudiced actually showed less

negative evaluations of Blacks in the threatening context, compared with other participants overall and also with high-motivation participants in the non-threatening context. Interestingly, these participants' less negative attitudes resulted from an inhibition of negative information in the threatening condition. This work suggests that people's chronic motivations can determine the way in which they respond to contextual cues regarding the nature of the evaluated stimuli (see also Moskowitz, Gollwitzer, Wasel, & Schaal, 1999).

The research on both indirect and direct methods of attitude measurement have uncovered a variety of moderators that predict when attitudes are likely to be more stable or context dependent. Situational differences, such as recently activated or chronic exemplars and goals, as well as individual differences in people, can influence the extent to which attitudes will be influenced by contextual cues. The question is not whether attitudes are basically stable versus constructed, but rather *when* each is true.

CHANGING OUR LIKES AND DISLIKES

What does attitude "change" mean? Attitude change is typically meant to imply that some manipulation produces changes in the underlying representations. But, how can we infer this? On the one hand, evidence for the contextual dependence of attitudes (reviewed above) would seem to imply that attitudes "change" quite readily. But, this would be misleading because such change could result from the stimulus being categorized differently across contexts, and with resulting differences in evaluations, but without any (significant) accompanying change at the level of representations. So, how might we address this question?

It is important to note that recent research on memory suggests that each time a stimulus is encountered (even in highly familiar situations), there are corresponding changes in the relevant representations (e.g., see Nader, Schafe, & LeDoux, 2000). Every encounter with a stimulus changes (e.g., strengthens) the underlying associations (or, weights) with those representations of accompanying stimuli, and weakens those with absent stimuli. This would be true presumably for both discrete and distributed representations. In this way, every encounter with a given stimulus is a learning/changing instance.

Given this, what is meant by attitude change? Probably what is most often meant by change is when an attitude shifts toward recently over previously learned, and *countervailing*, information. In other words, change is when the intensity or direction of the attitude becomes more aligned

with recently learned information. (This is a slightly different version than the case where one learns evaluative information about a novel stimulus.) From this perspective, “slow” change is what happens every time we process a stimulus, and this could be understood in terms of the attitude incrementally aligning with recent over previous information. “Fast” change, on the other hand, would be an instance of an attitude shifting relatively more quickly toward recently acquired, countervailing information.

Most of the work on attitude change has focused on direct attitudes with the assumption that attitudes are relatively stable. Recently Gawronski and Bodenhausen (2006) presented a model for how direct and indirect attitudes can be expected to change. The model is called the associative–propositional evaluation (APE) model, and, as its name implies, it assumes that attitude change can be understood in terms of associative and propositional processes. Whereas attitudes activated by associative processes are independent of whether the person endorses those attitudes (i.e., their subjective “truth value”), attitudes based on propositional processes have undergone an assessment (relying on capacity and motivation) for their truth value and endorsement. Gawronski and Bodenhausen (2006) outline the circumstances in which both direct and indirect attitudes should show change. For example, they note that the two types of measures might show asymmetric change in that indirect attitudes are changed without any change in direct attitudes, as well as the reverse. The former might happen when, for example, people undergo evaluative conditioning of which they are aware (e.g., Karpinski & Hilton, 2001). Although the evaluative conditioning changes the kinds of evaluative information tapped in an indirect measure, during direct measurement people can reject that newly learned information, given their awareness of the potentially biasing nature of that paradigm. Direct attitudes, on the other hand, might show change when people have recently learned propositional information about the stimulus that is not well-learned enough to show up on indirect measures (Gregg et al., 2006). We now briefly comment on social-cognitive work on direct versus indirect attitude change.

Direct attitude change

Changing direct attitudes has long been a traditional topic of interest within the attitudes literature, and generally falls under the umbrella of persuasion research. There is a voluminous literature that addresses the circumstances under which people’s self-reported attitudes can be

influenced by persuasive appeals. Although reviewing this literature is beyond the purview of the current chapter, we note that much of this work is consistent with a social-cognitive perspective, and the debates in this literature recently have been social cognitive in nature. For example, one debate has been whether persuasion should be understood in terms of a dual-process model versus a uni-process model. The bulk of the research has assumed that people can respond to persuasive appeals in one of two ways (e.g., Chaiken, 1980, 1987; Chaiken & Stangor, 1987; Petty & Cacioppo, 1986). In a systematic or central manner, the person carefully and deliberately analyzes the message-relevant information and is influenced by the strength of the message. The other route is called heuristic or peripheral, and occurs when people are responding with considerably less effort and attention. In this situation, a person might be influenced by cues that are unrelated to the strength of the message, such as the attractiveness or expertise of the spokesperson, the amount of information provided, or even the fluency of the information. This work relies on dual-process assumptions that were described in an earlier section of this current chapter. People are assumed to follow a central or systematic route whenever they are motivated to be accurate and have sufficient cognitive resources. A considerable amount of research shows that people seem to be influenced by superficial cues when they are unmotivated or cognitively taxed, and are influenced by message-relevant information when they are motivated and have the resources.

In contrast with this perspective is the uni-model, by Kruglanski and colleagues (Kruglanski, Erb, Chun, Pierro, & Mannetti, 2003; Kruglanski, Thompson, & Spiegel, 1999). They argue that people actually use syllogistic reasoning in response to any persuasive appeal and so the process is the same regardless of the constraints of the situation and appeal. Differences emerge in terms of what people use for their reasoning process according to the placement and salience of the information in the appeal. Kruglanski and colleagues have argued that previous persuasion research has introduced a confound, wherein cues that have been identified as indicative of the heuristic or peripheral route have been easier to process (e.g., presented early in an appeal), whereas information that would be indicative of the central or systematic route has been more deeply embedded in the appeal.

Indirect attitude change

In terms of changing indirect attitudes, there has been less empirical activity. Much of this work

has used the learning paradigm of evaluative conditioning to try to alter (not just change the categorization of) existing indirect attitudes. The phenomenon of evaluative conditioning has received increased empirical attention recently from social psychologists (e.g., Hofmann, De Houwer, Perugini, Baeyens, & Crombez, 2010; Jones et al., 2009; Walther, Nagengast, & Trasselli, 2005). There are still various questions about the precise mechanisms and boundaries of evaluative conditioning (see Hofmann et al., 2010; Jones et al., 2009); however, multiple papers provide evidence for the effect itself. For instance, although there is evidence that evaluative conditioning emerges when the primes are presented subliminally or when participants report no awareness of the prime–target pairings (Aarts, Custers, & Holland, 2007; Aarts, Custers, & Marien, 2008; Baeyens, Eelen, & Van der Bergh, 1990; Custers & Aarts, 2005; Dijksterhuis, 2004; Jones et al., 2009; Jones, Pelham, Carvalho, & Mirenberg, 2004; Olson & Fazio, 2001, 2002, 2006; Ruys & Stapel, 2009; Walther, 2002; Walther & Nagengast, 2006), there is also evidence that it does not emerge when participants do not have at least some awareness (Dawson et al., 2007; Field, 2000; Pleyers et al., 2007; Stahl, Unkelbach, & Corneille, 2009). This is an ongoing debate.

Of most relevance to this topic of attitude change (vs formation), evaluative conditioning has emerged for familiar (vs novel) objects (Aarts et al., 2007; Custers & Aarts, 2005; Dijksterhuis, 2004; Karpinski & Hilton, 2001; Kawakami, Dovidio, Moll, Hermsen, & Russin, 2000; Olson & Fazio, 2006; Ruys & Stapel, 2009; Rudman, Asmore, & Gary, 2001), and can last until at least 2 days (Olson & Fazio, 2006). For example, Olson and Fazio (2006) found that participants who had received positive conditioning of African American faces later showed significantly less prejudice toward the group. This work is also consistent with recent research showing that participants who repeatedly associated approach movements toward a given stimulus later showed more positive indirect attitudes toward it (e.g., Kawakami, Phills, Steele, & Dovidio, 2007). Given the close correspondence between arm movements and evaluation (e.g., Chen & Bargh, 1999), these findings speak to how indirect attitudes can be changed.

THE FORMATION OF LIKES AND DISLIKES

The formation of attitudes has been an understudied topic in the attitudes literature, as noted repeatedly (Albarracín et al., 2005; Eagly &

Chaiken, 1993; Fazio & Olson, 2003). Whereas the topic of changing attitudes concerns cases where new (often countervailing) information is provided about familiar stimuli, the formation of attitudes concerns cases where information is provided about novel or unfamiliar stimuli. Most of the work has been conducted with direct attitudes, but some recent work has examined the formation of indirect attitudes. These two areas of research on development are briefly described below.

Direct attitudes

Although some early research focused on unconscious processes underlying the formation of attitudes (e.g., Staats & Staats, 1958), the majority of research until the past few years has focused on the formation of attitudes that people can verbalize. The subjective–expected–utility approach to decision making, first articulated by the statistician Leonard Savage (1954), became popular among social psychologists attempting to explain attitude formation (e.g., Fishbein & Ajzen, 1974; Rosenberg, 1956; Wyer, 1973). This approach assumes that people are basically rational – people form beliefs based on information they receive, form attitudes based on those beliefs, and choose actions based on those attitudes and the expected probabilities of the outcomes. Later approaches to direct attitude formation focused on shortcut strategies (e.g., Chaiken, 1980; Petty & Cacioppo, 1986). These approaches assumed that people, if they are not rational, at least want to be rational in that they are motivated to hold correct beliefs and attitudes and act accordingly. However, it is also assumed that people are motivated to conserve cognitive resources. Thus, people will maintain attitudes based on superficial (peripheral) cues unless they are sufficiently motivated to seek out more in-depth information and have the cognitive resources to process and seek out information more rigorously. For example, the quality of the arguments has a greater impact on persuasion under conditions of high- than low-issue involvement (Petty & Cacioppo, 1979) and for individuals high than low in need for cognition (Cacioppo, Petty, & Morris, 1983). Conversely, peripheral cues such as the attractiveness of the communicator have a greater impact on persuasion under conditions of low than high involvement (Chaiken, 1980; Petty, Cacioppo, & Goldman, 1981).

An especially interesting examination of the formation of direct attitudes was by Betsch et al. (2001). In this paper, the authors tested whether direct attitudes could form during exposure to only superficially processed evaluative information. Participants were asked to watch a series of ads

on a computer, and while they were doing so, information about the share value of novel stocks appeared at the top of the screen. Betsch et al. wanted to examine not only whether participants would form evaluations of the stocks based on such subtle exposure but also how the pieces of evaluative information would be combined. Although some research (e.g., Anderson, 1983) suggests that pieces of evaluative information would be averaged, other work from animal learning suggests that it might be summed. Betsch et al. found evidence that participants reported attitudes toward the novel stocks according to the overall summation of the positive information. This research is an especially interesting demonstration of not only how direct attitudes can be formed with minimal processing of the evaluative information but also of how they are formed in terms of process.

Indirect attitudes

What are the circumstances in which people form attitudes that can be measured indirectly? In other words, when and how are people able to form attitudes toward stimuli that can be activated unintentionally, rapidly, and perhaps at times even without awareness? Recent work has taken various different strategies to try to instill indirect attitudes.

Almost all of the work on the formation of indirect attitudes has provided a large number of trials in which the novel object is paired repeatedly with evaluative information. For example, Olson and Fazio (2001) developed an evaluative conditioning paradigm and examined whether it would lead to the formation of indirect attitudes. In this study, participants were presented with various stimuli randomly on a computer screen, and the stimuli consisted of images as well as words. Their task was to press a key whenever they saw a specific target stimulus. The conditioned stimulus (CS) was a Pokemon character (participants were unfamiliar with these stimuli), and the unconditioned stimuli (US) were highly positive (e.g., ice cream sundae) or negative (e.g., cockroach) stimuli. The CS was presented simultaneously with the US, and they were never the target stimulus that participants were monitoring. Participants did not report any awareness of the covariation, and yet still showed a significant preference for the positively conditioned stimulus both on direct as well as indirect measures (see also De Houwer, Baeyens, & Field, 2005; De Houwer, Thomas, & Baeyens, 2001; Martin & Levy, 1978; Walther, Nagengast, & Trasselli, 2005).

As another example of the formation of indirect attitudes through an extensive amount of learning, earlier in this chapter we described research by Rydell and colleagues (Rydell et al., 2006) that showed newly formed indirect attitudes toward a novel target. Participants were presented with many (100) trials in which the novel stimulus (Bob) was paired with positive or negative subliminally presented information as well as positive or negative verbally presented information. Participants' indirect attitudes developed in line with the valence of the subliminally presented information.

The methodological strategy of trying to form indirect attitudes through extensive new pairings or information is consistent with the notion that such attitude rely on implicit memory which only enables slow learning. However, as we have argued already, there are examples of relatively fast formation of indirect attitudes by Ashburn-Nardo et al. (2001) and Gregg et al. (2006). There is also research by De Houwer et al. (1998) and Castelli et al. (2004) showing that novel stimuli that have been only briefly and recently classified as positive or negative can themselves lead to consistent evaluations on indirect measures. For example, De Houwer et al. (1998) found that non-words that had been briefly paired with highly positive or negative words were themselves indirectly evaluated in line with those positive or negative classifications. And, Castelli and colleagues (2004) showed that recently learned evaluative information about people determined participants' indirect attitudes toward those people. Participants viewed a series of novel people, and these people were described simply as being either child molesters or child counselors. The classification of each person was brief, and yet minutes later, participants indirectly evaluated those novel targets in line with the classification. This work, too, shows that relatively novel stimuli can become evaluatively "stamped" relatively quickly and easily, and in such a way that allows for later unintentional and rapid evaluation.

The idea that stimuli can quickly take on an evaluative connotation is consistent with animal learning literature on single-trial learning. A wide variety of animals and insects, including worms, flies, monkeys, and humans, seem able to develop fear responses after a single experience with a threatening stimulus (see LeDoux, 2000). Rats who are electrically shocked in a specific cage, for instance, later show fear responses to the cage. It is interesting to note that humans as well as non-human animals can retain and express learned preferences even without any accompanying explicit memory about the stimulus or its evaluative meaning (e.g., Squire, 1992; Squire & Kandel, 1999). For instance, if a person with anterograde

amnesia has an unpleasant experience with a new acquaintance, the person will show evidence for the implicit memory of that evaluative experience even when he or she of course cannot retain any explicit awareness for the acquaintance.

WHAT ATTITUDES PREDICT

For decades, researchers have puzzled over the weak relationship between attitudes and behavior (e.g., LaPiere, 1934; Thurstone, 1928; Wicker, 1969). The oft-cited paper of Wicker (1969) showed that a meta-analysis of 45 studies yielded a mean correlation of 0.15 between attitudes and behaviors. More recently, researchers have become more optimistic about the consistency between the two variables. In a meta-analysis of 88 studies, Kraus (1995) found that the average attitude-behavior correlation was 0.38, with over half of the studies showing a correlation of above 0.30.

Efforts to explain the attitude-behavior relationship are of two kinds: moderator variable research and methodological research. The consistency (and lack thereof) between attitudes and behaviors can be explained by moderating variables that boost or weaken the relationship. These moderators can be situational, socio-cognitive, or personality variables. Alternatively, other research focuses on showing how methodological flaws yield a certain relationship, while improvements can lead to a stronger relationship. Specifically, indirect attitude researchers have argued that indirect measures of attitudes do not have the demand effects and other weaknesses that explicit measures have, and demonstrate a stronger attitude-behavior relationship (Greenwald, McGhee, & Schwarz, 1998).

In terms of theoretical work, one of the most influential theories concerning predictions about attitudes and behavior is Fazio's MODE model (Motivation and Opportunity as Determinants of the attitude-behavior relation; Fazio, 1986, 1990; Olson & Fazio, 2009). As previously discussed, according to Fazio, an attitude is a learned association in memory between an object and a positive or negative evaluation of that object. Consistent with other dual processing theories (e.g., Chaiken & Trope, 1999), the model claims that attitudes can be activated either in a controlled, deliberative manner or an automatic, spontaneous manner.

Based on the MODE model, indirect attitudes will lead to behaviors consistent with the attitude depending on two moderators: motivation and opportunity. Automatically activated attitudes are the starting point of the attitude-to-behavior process.

When there is a lack of sufficient motivation and/or cognitive capacity to process information (opportunity), overt judgments and behaviors are hypothesized to reflect the automatically activated attitude. In contrast, sufficient motivational factors and/or opportunity can intercept the automatic attitude, such that the subsequent behavior is less influenced by the automatically activated attitude. Alternatively, in situations when there is no motivation and/or opportunity, nor chronically accessible automatic attitudes, the behavior is more likely to be determined by salient cues associated with the attitude object. According to the MODE model, automatic or spontaneous activation is reflective of strong attitudes. When a strong link has been established in memory, it is more likely to be automatically activated, and thus, is more chronically accessible.

Therefore, the MODE model predicts that attitudes should be good predictors of behaviors when they are readily accessible from memory. In contrast, attitudes should be relatively poor predictors of behaviors when they are not readily accessible. Indeed, research testing this hypothesis has shown that attitude accessibility (as operationalized by response latency), predicts attitude-behavior consistency in domains such as voting (Fazio & Williams, 1986) and prejudiced behavior (Fazio et al., 1995).

Direct attitudes and behavior consistency

Several moderators between direct attitudes and behavior consistency have been identified in the literature, such as attitude accessibility, certainty, and temporal stability. Because there are several excellent recent reviews and meta-analyses on the relationship between direct attitudes and behavior consistency (see Cooke & Sheeran, 2004; Crano & Prislin, 2006; Glasman & Albarracín, 2006), they will not be further discussed in this chapter.

Implicit attitudes and behavior consistency

More recent research has focused on the extent to which attitude-behavior consistency can be predicted by implicit measures of attitudes, such as the Implicit Association Test (IAT, Greenwald et al., 1998). The IAT and other implicit measures of attitudes have received much attention, in part, because they have been shown to be resistant to self-presentational concerns that can mask

personally or socially undesirable attitudes (Greenwald et al., 1998).

The IAT has also been demonstrated to predict behaviors, in some cases better than explicit measures of attitudes. In a review of 32 studies on attitudes toward African-American–White interracial behavior, predictive validity of IAT measures significantly exceeded that of self-report measures (Greenwald, Poehlman, Uhlmann, & Banaji, 2009). Although an impressive amount of research has utilized the IAT and other implicit measures of attitudes to predict behaviors, some researchers have questioned the predictive validity of implicit measures (e.g., Blanton, Jaccard, Klick, Mellers, Mitchell, & Tetlock, 2009).

Implicit measures of attitudes appear to be better at predicting behaviors that are uncontrolled and/or not overt compared to more controlled behaviors. For example, Vanman, Saltz, Nathan, and Warren (2004) showed that IAT measures of racial bias did not correspond with discrimination against African Americans, as measured by participants not choosing the African-American candidate among two other equally strong candidates to win a fellowship. However, studies which measured less overt behaviors have demonstrated that the IAT is a significant predictor. McConnell and Leibold (2001) showed that indirect measures of racial bias predicted a greater discrepancy in how the White and the African-American confederates rated their interaction with the participant, such that the White confederate rated the interaction more positively, while the African-American confederate rated the interaction more negatively. Indirect measures of racial bias correlated with negative nonverbal behaviors while direct measures of prejudice did not. Similarly, indirect measures of racial bias, but not direct measures of prejudice, predicted the distance White participants placed their chair from an African-American confederate's belongings (Amodio & Devine, 2006).

Other implicit measures of prejudice have been shown to relate to intergroup behavior. Fazio, Jackson, Dunton, and Williams (1995) developed a supraliminal affective priming measure which was shown to predict the warmth and interest a participant displayed toward an African-American confederate while responses to the Modern Racism Scale did not. Using a subliminal affective priming task, Dovidio, Kawakami, Johnson and Johnson (1997) demonstrated that indirect measures of racial bias predicted negative nonverbal behaviors toward an African-American confederate but did not predict deliberative race-related responses.

The most powerful demonstrations in science tend to be those that show a real-world effect. After demonstrating that roommate relationships of randomly paired interracial freshmen are more

likely to dissolve than randomly paired White freshmen (Study 1), Towles-Schwen and Fazio (2006) found that the indirect measure of racial bias of the White roommate (via a superluminal affective priming procedure) predicted roommate longevity between interracial roommates (Study 2). Meanwhile, the White roommate's explicit motivation to control for his or her prejudice did not predict the relationship outcome.

In an international study of impressive scale, Nosek et al. (2009) used the IAT to measure people's indirect attitudes toward women, men, and science from a sample of almost a half a million people in 34 countries. The researchers found that the more strongly people of a certain country associated men with science, the larger the gender gap of math and science achievement among 8th graders of that country. Direct (self-report) measures of attitudes toward men, women, and science did not provide additional predictive validity of the achievement gap. Taken together, these studies suggest that indirect measures of attitudes serve as powerful predictors of important real-world behavioral outcomes, particularly in contexts such as race, gender, or harmful habits, where responses on more direct measures may be intentionally or unintentionally biased.

CONCLUSION

In this chapter, we have provided a social cognitive perspective on the construct of attitudes. We have reviewed some of the most widely agreed upon characteristics about the construct, including that they predict judgment and behavior, can vary across time and situation, and can form relatively easily. We have also discussed some of the more contentious and ongoing conceptual questions in this area such as how to think about attitude generation according to the issues of representation, process, and system. Throughout the chapter we have noted the contemporary emphasis on differences between indirectly and directly measured attitudes, and the possible basis of such differences. It seems safe to say that the current state of social cognitive interest in and work on attitudes speaks well for the continued quoting of Allport's (1935) famous stance on attitudes.

REFERENCES

- Aarts, H., Custers, R., & Holland, R. W. (2007). The nonconscious cessation of goal pursuit: When goals and negative affect are coactivated. *Journal of Personality and Social Psychology, 92*, 165–178.

- Aarts, H., Custers, R., & Marien, H. (2008). Preparing and motivating behavior outside of awareness. *Science*, *319*, 1639.
- Albarracín, D., Johnson, B. T., & Zanna, M. P. (2005). *The handbook of attitudes*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Allport, G. W. (1935). Attitudes. In C. Murchison (Ed.), *Handbook of social psychology* (pp. 798–844). Worcester, MA: Clark University Press.
- Amodio, D. M., & Devine, P. G. (2006). Stereotyping and evaluation in implicit race bias: Evidence for independent constructs and unique effects on behavior. *Journal of Personality and Social Psychology*, *91*, 652–661.
- Amodio, D. M., Harmon-Jones, E., & Devine, P. G. (2003). Individual differences in the activation and control of affective race bias as assessed by startle eyeblink responses and self-report. *Journal of Personality and Social Psychology*, *84*, 738–753.
- Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Arkes, H. R., & Tetlock, P. E. (2004). Attributions of implicit prejudice, or "Would Jesse Jackson 'fail' the Implicit Association Test?". *Psychological Inquiry*, *15*, 257–278.
- Ashburn-Nardo, L., Voils, C. I., & Monteith, M. J. (2001). Implicit associations as the seeds of intergroup bias: How easily do they take root? *Journal of Personality and Social Psychology*, *81*, 789–799.
- Asch, S. E. (1948). The doctrine of suggestion, prestige, and imitation in social psychology. *Psychological Review*, *55*, 250–276.
- Baeyens, F., Eelen, P., & Van den Bergh, O. (1990). Contingency awareness in evaluative conditioning: A case for unaware affective-evaluative learning. *Cognition and Emotion*, *4*(1), 3–18.
- Banaji, M. R. (2001). Implicit attitudes can be measured. In H. L. Roediger, J. S. Nairne, I. Neither, & A. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 117–150). Washington, DC: American Psychological Association.
- Bargh, J. A. (1994). The Four Horsemen of automaticity: Awareness, efficiency, intention, and control in social cognition. In R. S. Wyer, Jr., & T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed., pp. 1–40). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bargh, J. A., Chaiken, S., Govevender, R., & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal of Personality and Social Psychology*, *62*, 893–912.
- Bargh, J. A., & Ferguson, M. J. (2000). Beyond behaviorism: On the automaticity of higher mental processes. *Psychological Bulletin*, *126*, 925–945.
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–645.
- Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society of London: Biological Sciences*, *364*, 1281–1289.
- Bassili, J. N., & Brown, R. D. (2005). Implicit and explicit attitudes: Research, challenges, and theory. In D. Albarracín, B. T. Johnson & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. 543–574). Mahwah, NJ: Lawrence Erlbaum Associates.
- Bechtel, W. (2008). *Mental mechanisms: Philosophical perspectives on cognitive neuroscience*. London: Routledge.
- Berry, C. J., Shanks, D. R., & Henson, R. N. A. (2008). A unitary signal-detection model of implicit and explicit memory. *Trends in Cognitive Sciences*, *12*, 367–373.
- Betsch, T., Plessner, H., Schwieren, C., & Gütig, R. (2001). I like it but I don't know why: A value-account approach to implicit attitude formation. *Personality and Social Psychology Bulletin*, *27*, 242–253.
- Blair, I. V., Ma, J. E., & Lenton, A. P. (2001). Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. *Journal of Personality and Social Psychology*, *81*, 828–841.
- Blanton, H., Jaccard, J., Klick, J., Mellers, B., Mitchell, G., & Tetlock, P. E. (2009). Strong claims and weak evidence: Reassessing the predictive validity of the IAT. *Journal of Applied Psychology*, *94*, 567–582.
- Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*, 116–131.
- Cacioppo, J. T., Petty, R. E., & Morris, K. (1983). Effects of need for cognition on message evaluation, recall, and persuasion. *Journal of Personality and Social Psychology*, *45*, 805–818.
- Carlston, D. E., & Smith, E. R. (1996). Principles of mental representation. In E. T. Higgins, & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 184–210). New York: Guilford Press.
- Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: Bradford Books, MIT Press.
- Castelli, L., Zogmaister, C., Smith, E. R., Arcuri, L. (2004). On the automatic evaluation of social exemplars. *Journal of Personality and Social Psychology*, *86*, 373–387.
- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality & Social Psychology*, *39*(5), 752–766.
- Chaiken, S. (1987). The heuristic model of persuasion. In Zanna, M. P., Olson, J. M., & Herman, C. P. (Eds.), *Social influence: The Ontario Symposium* (Vol. 5, pp. 3–39). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Chaiken, S., & Bargh, J. A. (1993). Occurrence versus moderation of the automatic attitude activation effect: Reply to Fazio. *Journal of Personality & Social Psychology*, *64*, 759–765.
- Chaiken, S., & Stangor, Ch. (1987). Attitudes and attitude change. *Annual Review of Psychology*, *38*, 575–630.
- Chaiken, S., & Trope, Y. (1999). *Dual-process theories in social psychology*. New York: Guilford Press.
- Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach and avoid the stimulus. *Personality and Social Psychology Bulletin*, *25*, 215–224.
- Conroy, F. R., & Smith, E. R. (2007). Attitude representation: Attitudes as patterns in a distributed, connectionist representational system. *Social Cognition*, *25*, 739–758.

- Cooke, R., & Sheeran, P. (2004). Moderation of cognition–intention and cognition–behaviour relations: A meta-analysis of properties of variables from the theory of planned behaviour. *British Journal of Social Psychology, 43*, 159–186.
- Crano, W. D., & Prislin, R. (2006). Attitudes and persuasion. *Annual Review of Psychology, 57*, 345–374.
- Cunningham, W. A. (2010). In defense of brain mapping in social and affective neuroscience. *Social Cognition, 28*, 716–721.
- Cunningham, W. A., Johnson, M. K., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2003). Neural components of social evaluation. *Journal of Personality and Social Psychology, 85*, 639–649.
- Cunningham, W. A., Johnson, M. K., Raye, C. L., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2004). Separable neural components in the processing of Black and White faces. *Psychological Science, 15*, 806–813.
- Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science, 12*, 163–170.
- Cunningham, W. A., Raye, C. L., & Johnson, M. K. (2004). Implicit and explicit evaluation: fMRI correlates of valence, emotional intensity, and control in the processing of attitudes. *Journal of Cognitive Neuroscience, 16*, 1717–1729.
- Cunningham, W. A., Zelazo, P. D., Packer, D. J., & Van Bavel, J. J. (2007). The iterative reprocessing model: A multilevel framework for attitudes and evaluation. *Social Cognition, 25*, 736–760.
- Custers, R., & Aarts, H. (2005). Positive affect as implicit motivator: On the nonconscious operation of behavioral goals. *Journal of Personality and Social Psychology, 89*(2), 129–142.
- 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.
- Dawson, M. E., Rissling, A. J., Schell, A. M., & Wilcox, R. (2007). Under what conditions can human affective conditioning occur without contingency awareness? Test of the evaluative conditioning paradigm. *Emotion, 7*, 755–766.
- De Houwer, J. (2003). The extrinsic affective Simon task. *Experimental Psychology, 50*, 77–85.
- De Houwer, J., Baeyens, F., & Field, A. P. (2005). Associative learning of likes and dislikes: Some current controversies and possible ways forward. *Cognition and Emotion, 19*, 161–174.
- De Houwer, J., & Eelen, P. (1998). An affective variant of the Simon paradigm. *Cognition and Emotion, 8*, 45–61.
- De Houwer, J., Hermans, D., & Eelen, P. (1998). Affective and identity priming with episodically associated stimuli. *Cognition and Emotion, 12*, 145–169.
- De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin, 135*, 347–368.
- De Houwer, J., Thomas, S., & Baeyens, F. (2001). Associative learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin, 127*, 853–869.
- Devaney, Robert L. (2003). *An introduction to chaotic dynamical systems* (2nd ed.). Boulder, CO: Westview Press.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56*, 5–18.
- Dietrich, E., & Markman, A. B. (2003). Discrete thoughts: Why cognition must use discrete representations. *Mind & Language, 18*, 95–119.
- Dijksterhuis, A. (2004). Think different: The merits of unconscious thought in preference developments and decision making. *Journal of Personality and Social Psychology, 87*, 586–598.
- Doob, L.W. (1947). The behavior of attitudes. *Psychological Review, 51*, 135–156.
- Dovidio, J. F., Kawakami, K., & Beach, K. R. (2001). Implicit and explicit attitudes: Examination of the relationship between measures of intergroup bias. In R. Brown & S.L. Gaertner (Eds.), *Blackwell handbook of social psychology* (Vol. 4, *Intergroup Relations*, pp. 175–197). Oxford, UK: Blackwell.
- Dovidio, J. F., Kawakami, K., Johnson, C., & Johnson, B. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology, 33*, 510–540.
- Dovidio, J. F., Mann, J., & Gaertner, S. L. (1989). Resistance to affirmative action: The implications of aversive racism. In F. A. Blanchard & F. J. Crosby (Eds.), *Affirmative action in perspective* (pp. 83–102). New York: Springer-Verlag.
- Dretske, Fred, F. (1995). *Naturalizing the mind*. Cambridge, MA: MIT Press.
- Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich.
- Elman et al., (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.
- Epstein, S. (1994). Integration of the cognitive and psychodynamic unconscious. *American Psychologist, 49*, 709–724.
- Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgement and social cognition. *Annual Review of Psychology, 59*, 255–278.
- Fazio, R. H. (1986). How do attitudes guide behavior? In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (pp. 204–243). New York: Guilford Press.
- Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 75–109). New York: Academic Press.
- Fazio, R. H. (1993). Variability in the likelihood of automatic attitude activation: Data re-analysis and commentary on Bargh, Chaiken, Gøvdender, and Pratto (1992). *Journal of Personality and Social Psychology, 64*, 753–758, 764–765.
- 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). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition and Emotion, 14*, 1–27.
- Fazio, R. H. (2007). Attitudes as object–evaluation associations of varying strength. *Social Cognition, 25*, 603–637.
- Fazio, R. H., Chen, J., McDonel, E. C., & Sherman, S. J. (1982). Attitude accessibility, attitude–behavior consistency and the strength of the object–evaluation association. *Journal of Experimental Social Psychology, 18*, 339–357.
- 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.
- Fazio, R. H., & Williams, C. J. (1986). Attitude accessibility as a moderator of the attitude–perception and attitude–behavior relations: An investigation of the 1984 presidential election. *Journal of Personality and Social Psychology, 51*, 505–514.
- Fazio, R. H., & Zanna, M. P. (1978). Attitudinal qualities relating to the strength of the attitude–behavior relationship. *Journal of Experimental Social Psychology, 14*, 398–408.
- Ferguson, M. J. (2007a). On the automatic evaluation of end-states. *Journal of Personality and Social Psychology, 92*, 596–611.
- Ferguson, M. J. (2007b). The automaticity of evaluation. Invited chapter in J. A. Bargh (Ed.), *Social psychology and the unconscious: The automaticity of higher mental processes* (pp. 219–264). New York: Psychology Press.
- Ferguson, M. J. (2008). On becoming ready to pursue a goal you don't know you have: Effects of nonconscious goals on evaluative readiness. *Journal of Personality and Social Psychology, 25*, 557–571.
- Ferguson, M. J., & Bargh, J. A. (2003). The constructive nature of automatic evaluation. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 169–188). NJ: Erlbaum.
- Ferguson, M. J., & Bargh, J. A. (2004). Liking is for doing: Effects of goal pursuit on automatic evaluation. *Journal of Personality and Social Psychology, 88*, 557–572.
- Ferguson, M. J. & Zayas, V. (2009). Nonconscious evaluation. *Current Directions in Psychological Science, 18*, 362–366.
- Field, A. P. (2000). Evaluative conditioning is Pavlovian conditioning: Issues of definition, measurement, and the theoretical importance of contingency awareness. *Consciousness and Cognition, 9*(1), 41–49.
- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.
- 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., LeBel, E. P., & Peters, K. R. (2007). What do implicit measures tell us? Scrutinizing the validity of three common assumptions. *Perspectives on Psychological Science, 2*, 181–193.
- Gawronski, B., Peters, K. R., Brochu, P. M., & Strack, F. (2008). Understanding the relations between different forms of racial prejudice: A cognitive consistency perspective. *Personality and Social Psychology Bulletin, 34*, 648–665.
- Gawronski, B., Rydell, R. J., Vervliet, B., & De Houwer, J. (2010). Generalization versus contextualization in automatic evaluation. *Journal of Experimental Psychology: General, 139*, 683–701.
- Gilbert, D. T. (1999). What the mind's not. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology*. New York: Guilford Press.
- Glasman, L. R., & Albarracín, D. (2006). Forming attitudes that predict future behavior: A meta-analysis of the attitude–behavior relation. *Psychological Bulletin, 132*, 778–822.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*, 4–27.
- Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109*, 3–25.
- Greenwald, A. G., & Farnham, S. D. (2000). Using the Implicit Association Test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology, 79*, 1022–1038.
- Greenwald, A. G., McGhee, D. E., & Schwarz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74*, 1464–1480.
- Greenwald, A. G., Poehlman, T. A., Uhlmann, E., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology, 97*, 17–41.
- 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.
- Grossberg, S., Boardman, I., & Cohen, M. A. (1997). Neural dynamics of variable-rate speech categorization. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 481–503.
- Hamann S., & Mao, H. (2002). Positive and negative emotional verbal stimuli elicit activity in the left amygdala. *Neuroreport, 13*, 15–19.
- Haugeland, J. (1991). Semantic engines: An introduction to mind design. In J. Haugeland (Ed.), *Mind dDesign: Philosophy, psychology, artificial intelligence* (pp. 1–34). Cambridge, MA: MIT Press.

- Hebb, D.O. (1949). *The organization of behavior*. New York: John Wiley & Sons.
- Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (2010). Evaluative conditioning in humans: A meta-analysis. *Psychological Bulletin*, *136*, 390–421.
- 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.
- Huskinson, T. L. H., & Haddock, G. (2004). Assessing individual differences in attitude structure: Variance in the chronic reliance on affective and cognitive information. *Journal of Experimental Social Psychology*, *40*, 82–90.
- Ito, T. A., & Cacioppo, J. T. (2000). Electrophysiological evidence of implicit and explicit categorization processes. *Journal of Experimental Social Psychology*, *36*, 660–676.
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*, 513–541.
- Jilk, D. J., Lebiere, C., O'Reilly, R. C., & Anderson, J. R. (2008). SAL: An explicitly pluralistic cognitive architecture. *Journal of Experimental and Theoretical Artificial Intelligence*, *20*, 197–218.
- Jones, E. E., & Sigall, H. (1971). The bogus pipeline: A new paradigm for measuring affect and attitude. *Psychological Bulletin*, *76*, 349–364.
- Jordan, C. H., Spencer, S. J., & Zanna, M. P. (2003). "I love me ... I love me not": Implicit self-esteem, explicit self-esteem, and defensiveness. In S. J. Spencer, S. Fein, M. P. Zanna, & J. M. Olson (Eds.), *Motivated social perception: The Ontario Symposium* (Vol. 9, pp. 117–145). Mahwah, NJ: Lawrence Erlbaum Associates.
- Jones, C. R., Fazio, R. H., & Olson, M. A. (2009). Implicit misattribution as a mechanism underlying evaluative conditioning. *Journal of Personality and Social Psychology*, *96*, 933–948.
- Jones, J. T., Pelham, B. W., Carvallo, M., & Mirenberg, M. C. (2004). How do I love thee? Let me count the Js: Implicit egotism and interpersonal attraction. *Journal of Personality and Social Psychology*, *87*(5), 665–683.
- Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, *58*, 697–720.
- Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp.49–81). New York: Cambridge University Press.
- Karpinski, A., & Hilton, J. L. (2001). Attitudes and the implicit association test. *Journal of Personality and Social Psychology*, *81*, 774–788.
- Katz, I., & Hass, R. G. (1988). Racial ambivalence and American value conflict: Correlational and priming studies of dual cognitive structures. *Journal of Personality and Social Psychology*, *55*, 893–905.
- Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S., & Russin, A. (2000). Just say no (to stereotyping): Effects of training in the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology*, *78*, 871–888.
- Kawakami, K., Phills, C. E., Steele, J. R. & Dovidio, J. F. (2007). (Close) Distance makes the heart grow fonder: Improving implicit racial attitudes and interracial interactions through approach behaviors. *Journal of Personality and Social Psychology*, *92*, 957–971.
- Keil, F. (1989). *Concepts, kinds and cognitive development*. Cambridge, MA: MIT Press.
- Keren, G., & Schul, Y. (2009) Two is not always better than one: A critical evaluation of two-system theories. *Perspectives on Psychological Science*, *4*, 533–550.
- Klauer, K. C., & Mierke, J. (2005). Task-set inertia, attitude accessibility, and compatibility-order effects: New evidence for a task-set switching account of the IAT effect. *Personality and Social Psychology Bulletin*, *31*, 208–217.
- Klauer, K. C., & Musch, J. (2003). Affective priming: Findings and theories. In K. C. Klauer, & J. Musch (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 7–50). Mahwah, NJ: Lawrence Erlbaum Associates.
- Klauer, K. C., Schmitz, F., Teige-Mocigemba, S., & Voss, A. (2010). Understanding the role of executive control in the Implicit Association Test: Why flexible people have small IAT effects. *Quarterly Journal of Experimental Psychology*, *63*, 595–619.
- Klauer, K. C., & Stern, E. (1992). How evaluations guide memory-based judgments: A two-process model. *Journal of Experimental Social Psychology*, *28*, 186–206.
- Klinger, M. R., Burton, P. C., & Pitts, G. S. (2000). Mechanisms of unconscious priming: I. Response competition, not spreading activation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 441–455.
- Koole, S.K., Dijksterhuis, A., & van Knippenberg, A. (2001). What's in a name: Implicit self-esteem. *Journal of Personality and Social Psychology*, *80*, 614–627.
- Kraus, S. J. (1995). Attitudes and the prediction of behavior: A meta-analysis of the empirical literature. *Personality and Social Psychology Bulletin*, *21*, pp. 58–75.
- Krosnick, J. A., Judd, C. M., & Wittenbrink, B. (2005). The measurement of attitudes. In D. Albarracín, B. T. Johnson, & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. 21–76). Mahwah, NJ: Lawrence Erlbaum Associates.
- Krosnick, J. A., & Schuman, H. (1988). Attitude intensity, importance, and certainty and susceptibility to response effects. *Journal of Personality and Social Psychology*, *54*, 940–952.
- Kruglanski, A.W., Erb, H. P., Chun, W. Y., Pierro, A., & Mannetti, L. (2003). A parametric model of human judgment: Integrating the dual-mode frameworks in social cognition from a singular perspective. In J. Forgas, Von Hippel, W., & K. Williams (Eds.), *Responding to the social world: Explicit and implicit processes in social judgments and decisions*. Cambridge, UK: Cambridge University Press.
- Kruglanski, A. W., Thompson, E. P., & Spiegel, S. (1999). Separate or equal? Bimodal notions of persuasion and a single-process "unimodel". In S. Chaiken & Y. Trope (Eds.),

- Dual process models in social cognition: A source book.* New York: Guilford Press.
- Lamme, V. A. F., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neuroscience*, *23*, 571–579.
- LaPiere, R. T. (1934). Attitudes vs. actions. *Social Forces*, *13*, 230–237.
- Lavine, H., Huff, J.W., Wagner, S. H., & Sweeney, D. (1998). The moderating influence of attitude strength on the susceptibility to context effects in attitude surveys. *Journal of Personality and Social Psychology*, *75*, 359–373.
- LeDoux, J. (1996). *The emotional brain*. New York: Touchstone.
- LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review Neuroscience*, *23*, 155–184.
- Levey, A. B., & Martin, I. (1975). Classical conditioning of human "evaluative" responses. *Behaviour Research and Therapy*, *13*, 221–226.
- Lewin, K. (1935). *A dynamic theory of personality*. New York: McGraw-Hill.
- Liberzon I., Phan, K. L., Decker, L. R., & Taylor, S. F. (2003). Extended amygdala and emotional salience: A PET investigation of positive and negative affect. *Neuropsychopharmacology*, *28*, 726–733.
- Lichtenstein, S., & Slovic, P. (2006). *The construction of preference*. New York: Cambridge University Press.
- Logan, G. D. (1980). Attention and automaticity in Stroop and priming tasks: Theory and data. *Cognitive Psychology*, *12*, 523–553.
- Lowery, B. S., Hardin, C. D., & Sinclair, S. (2001). Social influence on automatic racial prejudice. *Journal of Personality and Social Psychology*, *81*, 842–855.
- Maddux, W. W., Barden, J., Brewer, M. B., & Petty, R. E. (2005). Saying no to negativity: The effects of context and motivation to control prejudice on automatic evaluative responses. *Journal of Experimental Social Psychology*, *41*, 19–35.
- Maio, G. R., & Esses, V. M. (2001). The need for affect: Individual differences in the motivation to approach or avoid emotions. *Journal of Personality*, *69*, 583–615.
- Marcus, G. F. (2001). *The algebraic mind: Integrating connectionism and cognitive science*. Cambridge, MA: MIT Press.
- Markman, A. B., & Dietrich, E. (2000). In defense of representation. *Cognitive Psychology*, *40*, 138–171.
- Marsh, K. L., Johnson, B. T., & Scott-Sheldon, L. A. J. (2001). Heart versus reason in condom use: Implicit versus explicit attitudinal predictors of sexual behavior. *Zeitschrift für Experimentelle Psychologie*, *48*, 161–175.
- Martin, I., & Levey, A. B. (1978). Evaluative conditioning. *Advances in Behaviour Research and Therapy*, *1*, 57–102.
- McCausley, C., & Stitt, C. L. (1978). An individual and quantitative measure of stereotypes. *Journal of Personality and Social Psychology*, *36*, 929–940.
- McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*, 419–457.
- McClelland, J. L., & Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience*, *4*, 310–322.
- McConahay, J. (1986). Modern racism, ambivalence, and the Modern Racism Scale. In J. Dovidio (Ed.), *Prejudice, discrimination, and racism* (pp. 91–125). San Diego, CA: Academic Press.
- McConnell, A. R., & Leibold, J. M. (2001). Relations among the Implicit Association Test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology*, *37*, 435–442.
- Metcalfe, J., & Mischel, W. (1999). A hot/cool system analysis of delay of gratification: Dynamics of willpower. *Psychological Review*, *106*, 3–19.
- Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*, 227–234.
- Mierke, J., & Klauer, K. C. (2003). Method-specific variance in the Implicit Association Test. *Journal of Personality and Social Psychology*, *85*, 1180–1192.
- Mitchell, J. P., Nosek, B. A., & Banaji, M. R. (2003). Contextual variations in implicit evaluation. *Journal of Experimental Psychology: General*, *132*, 455–469.
- Moore, S. G., Ferguson, M. J., & Chartrand, T. L. (2011). Affect in the aftermath: How goal pursuit influences implicit evaluations. *Cognition and Emotion*, *25*, 453–465.
- Moors, A., & De Houwer, J. (2001). Automatic appraisal of motivational valence: Motivational affective priming and Simon effects. *Cognition and Emotion*, *15*, 749–766.
- Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin*, *132*, 297–326.
- Morris, J. S., Ohman, A., & Dolan, R. J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature*, *393*, 467–470.
- Moskowitz, G. B., Gollwitzer, P. M., Wasel, W., & Schaal, B. (1999). Preconscious control of stereotype activation through chronic egalitarian goals. *Journal of Personality and Social Psychology*, *77*, 167–184.
- Musch, J., & Klauer, K. C. (2003). *The psychology of evaluation: Affective processes in cognition and emotion*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Nader, K., Schafe, G. E., & LeDoux, J. E. (2000). The labile nature of consolidation theory. *Nature Review Neuroscience*, *1*(3), 216–219.
- Neely, J. H. (1976). Semantic priming and retrieval from lexical memory: Evidence for facilitatory and inhibitory processes. *Memory and Cognition*, *4*, 648–654.
- Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, *106*, 225–254.
- Niedenthal, P. M. (1990). Implicit perception of affective information. *Journal of Experimental Social Psychology*, *26*, 505–527.
- Nosek, B. A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General*.

- Nosek, B. A., & Banaji, M. R. (2001). The go/no-go association task. *Social Cognition, 19*(6), 625–666.
- Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = Me, Me = Female, therefore Math is not equal to me. *Journal of Personality and Social Psychology, 83*, 44–59.
- 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., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., et al. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences, 106*, 10593–10597.
- O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford, UK: Oxford University Press.
- 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*, 89–103.
- Olson, M. A., & Fazio, R. H. (2004). Reducing the influence of extra-personal associations on the Implicit Association Test: Personalizing the IAT. *Journal of Personality and Social Psychology, 86*, 653–667.
- Olson, M. A., & Fazio, R. H. (2006). Reducing automatically-activated racial prejudice through implicit evaluative conditioning. *Personality and Social Psychology Bulletin, 32*, 421–433.
- Olson, M. A., & Fazio, R. H. (2009). Implicit and explicit measures of attitudes: The perspective of the MODE model. In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.), *Attitudes: Insights from the new implicit measures*. New York: Psychology Press.
- Olson, M. A., Fazio, R. H., & Hermann, A. D. (2007). Reporting tendencies underlie discrepancies between implicit and explicit measures of self-esteem. *Psychological Science, 18*, 287–291.
- Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist, 17*, 776–783.
- O'Reilly, R. C. (1998). Six principles for biologically-based computational models of cortical cognition. *Trends in Cognitive Sciences, 2*, 455–462.
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Chicago, IL: University of Illinois Press.
- Palmer, S. E. (1978). Fundamental aspects of cognitive representation. In E. Rosch & B. L. Lloyd (Eds.), *Cognition and categorization* (pp. 259–302). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Payne, B. K. (2005). Conceptualizing control in social cognition: How executive control modulates the expression of automatic stereotyping. *Journal of Personality and Social Psychology, 89*, 488–503.
- Payne, B. K., Burkley, M., & 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: Attitude misattribution as implicit measurement. *Journal of Personality and Social Psychology, 89*, 277–293.
- Payne, B. K., Govorun, O., & Arbuckle, N. L. (2008). Automatic attitudes and alcohol: Does implicit liking predict drinking? *Cognition & Emotion, 22*, 238–271.
- Petty, R. E., & Cacioppo, J. T. (1979). Issue involvement can increase or decrease persuasion by enhancing message-relevant cognitive processes. *Journal of Personality and Social Psychology, 37*, 1915–1926.
- Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. New York: Springer-Verlag.
- Petty, R. E., Cacioppo, J. T., & Goldman, R. (1981). Personal involvement as a determinant of argument-based persuasion. *Journal of Personality and Social Psychology, 41*, 847–855.
- Petty, R. E., & Krosnick, J. A. (1995). *Attitude strength: Antecedents and consequences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- 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. (1999). The Elaboration Likelihood Model: Current status and controversies. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 41–72). New York: Guilford Press.
- Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., et al. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience, 12*, 729–738.
- Phelps, E. A., O'Connor, K. J., Gatenby, J. C., Grillon, C., Gore, J. C., & Davis, M. (2001). Activation of the human amygdala to a cognitive representation of fear. *Nature Neuroscience, 4*, 437–441.
- Pleyers, G., Corneille, O., Luminet, O., & Yzerbyt, V. (2007). Aware and (dis)liking: Item-based analyses reveal that valence acquisition via evaluative conditioning emerges only when there is contingency awareness. *Journal of Experimental Psychology: Learning, Memory & Cognition, 33*, 130–144.
- Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 55–85). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Ranganath, K. A., Smith, C. T., & Nosek, B. A. (2008). Distinguishing automatic and controlled components of attitudes from direct and indirect measurement methods. *Journal of Experimental Social Psychology, 44*, 386–396.
- Roediger, H. L. (1990). Implicit memory: Retention without remembering. *American Psychologist, 45*(9), 1043–1056.
- Rosenberg, M. J. (1956). Cognitive structure and attitudinal affect. *Journal of Abnormal and Social Psychology, 53*, 367–372.

- Rosenberg, M. J. (1969). The conditions and consequences of evaluation apprehension. In R. Rosenthal & R. L. Rosnow (Eds.), *Artifact in behavioral research* (pp. 279–349). New York: Academic Press.
- Rosenberg, M. J., & Hovland, C. I. (1960). Cognitive, affective, and behavioural components of attitudes. In C. I. Hovland & M. J. Rosenberg (Eds.), *Attitude organisation and change: An analysis of consistency among attitude components* (pp. 1–14). New Haven, CT: Yale University Press.
- Rudman, L. A. (2004). Sources of implicit attitudes. *Current Directions in Psychological Science*, 13(2), 80–83.
- Rudman, L. A., Ashmore, R. D., & Gary, M. L. (2001). "Unlearning" automatic biases: The malleability of implicit stereotypes and prejudice. *Journal of Personality and Social Psychology*, 81, 856–868.
- Ruys, K. I., & Stapel, D. A. (2009). Learning to like or dislike by association: No need for contingency awareness. *Journal of Experimental Social Psychology*, 45, 1277–1280.
- 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.
- Salancik, G. R., & Conway, M. (1975). Attitude inferences from salient and relevant cognitive content about behavior. *Journal of Personality and Social Psychology*, 32, 829–840.
- Sarnoff, I. (1960). Psychoanalytic theory and social attitudes. *Public Opinion Quarterly*, 24, 251–279.
- Schacter, D. L., & Tulving, E. (Eds.) (1994). *Memory systems 1994*. Cambridge, MA: MIT Press.
- Schuman, H., & Johnson, M. P. (1976). Attitudes and behavior. *Annual Review of Sociology*, 2, 161–207.
- Schwarz, N. (2007). Attitude construction: Evaluation in context. *Social Cognition*, 25, 638–656.
- Schwarz, N., & Bohner, G. (2001). The construction of attitudes. In A. Tesser & N. Schwarz (Eds.), *Blackwell handbook of social psychology: Intraindividual processes* (Vol. 1, pp. 436–457). Oxford, UK: Blackwell.
- Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96, 523–568.
- Sherman, J. W. (2009). Controlled influences on implicit measures: Confronting the myth of process-purity and taming the cognitive monster. In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.), *Attitudes: Insights from the new wave of implicit measures* (pp. 391–426). Hillsdale, NJ: Lawrence Erlbaum Associates.
- 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.
- Sherman, J. W., & Payne, B. K. (2009). Separating automatic and controlled components of implicit measures: Process dissociation in social cognition. In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.), *Attitudes: Insights from the new wave of implicit measures*. New York: Psychology Press.
- Sherry, D. F., & Schacter, D. L. (1987). The evolution of multiple memory systems. *Psychological Review*, 94, 439–454.
- Shiffrin, R. M., & Dumais, S. T. (1981). The development of automatism. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp.111–140). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, 84, 127–190.
- Sia, T. L., Lord, C. G., Blessum, K., Ratcliff, C. D., & Lepper, M. R. (1997). Is a rose always a rose? The role of social category exemplar change in attitude stability and attitude-behavior consistency. *Journal of Personality and Social Psychology*, 72, 501–514.
- Slooman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, 119, 3–22.
- Smallwood, J., & Schooler, J.W. (2006). The restless mind. *Psychological Bulletin*, 132(6), 946–958.
- Smith, C. T., & Nosek, B. A. (2011). Affective focus increases the concordance between implicit and explicit attitudes. *Social Psychology*, 42, 300–313.
- Smith, E. R. (1992). The role of exemplars in social judgment. In L. L. Martin and A. Tesser (Eds.), *The construction of social judgment*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- 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. (1997). Preconscious automaticity in a modular connectionist system. In R. S. Wyer (Ed.), *Advances in social cognition* (Vol. 10, pp. 181–202). Mahway, NJ: Lawrence Erlbaum Associates.
- Smith, E. R. (2009). Distributed connectionist models in social psychology. *Social and Personality Psychology Compass*, 3, 64–76.
- Smith, E. R., & Conrey, F. R. (2007). Agent-based modeling: A new approach for theory-building in social psychology. *Personality and Social Psychology Review*, 11, 87–104.
- Smith, E. R., & DeCoster, J. (1999). Associative and rule-based processing: A connectionist interpretation of dual-process models. In S. Chaiken, & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 323–336). New York: Guilford Press.
- 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 evaluations influence categorization of multiply categorizable objects. *Journal of Personality and Social Psychology*, 71, 888–898.
- Smith, E. R., & Lerner, M. (1986). Development of automatism of social judgments. *Journal of Personality and Social Psychology*, 50, 246–259.
- Smith, E. R., & Zarate, M. A. (1992). Exemplar-based model of social judgment. *Psychological Review*, 99, 3–21.
- Smith, M. B., Bruner, J. S., & White, R. W. (1956). *Opinions and personality*. New York: Wiley.
- Sperber, D. (2005). Modularity and relevance: How can a massively modular mind be flexible and context-sensitive? P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind: Structure and content*.

- Spivey, M. J. (2007). *The continuity of mind*. New York: Oxford University Press.
- Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review*, *99*, 195–231.
- Squire, L. R., & Kandel, E. R. (1999). *Memory: From mind to molecules*. New York: Scientific American Library.
- Stahl, C., Unkelbach, C., & Corneille, O. (2009). On the respective contributions of awareness of US valence and US identity in attitude formation through evaluative conditioning. *Journal of Personality and Social Psychology*, *97*, 404–420.
- Staats, A. W., & Staats, C. K. (1958). Attitudes established by classical conditioning. *Journal of Abnormal and Social Psychology*, *57*, 3740.
- Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*(3), 220–247.
- Strack, F., & Deutsch, R. (2004). Reflection and impulse as determinants of “conscious” and “unconscious” motivation. In J. P. Forgas, K. Williams, & S. Laham (Eds.), *Social motivation: Conscious and unconscious processes*. Cambridge, UK: Cambridge University Press.
- Tesser, A. (1978). Self-generated attitude change. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 11, pp. 289–338). New York: Academic Press.
- Tesser, A., & Martin, L. (1996). The psychology of evaluation. In E. T. Higgins, & A.W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 400–432). New York: Guilford Press.
- Thagard, P. (1991). Philosophical and computational models of explanation. *Philosophical Studies*, *64*(October), 87–104.
- Thelen, E., & Smith, L. B., (1994). *A dynamical systems approach to development of cognition and action*. Cambridge, MA: Bradford Books/MIT Press.
- Thurstone, L. L. (1928). Attitudes can be measured. *American Journal of Sociology*, *33*, 529–554.
- Thurstone, L. L. (1931). Measurement of social attitudes. *Journal of Abnormal and Social Psychology*, *26*, 249–269.
- Tooby, J., & Cosmides, L. (2005). Conceptual foundations of evolutionary psychology. In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (pp. 5–67). Hoboken, NJ: Wiley.
- Towles-Schwen, T., & Fazio, R. H. (2006). Automatically-activated racial attitudes as predictors of the success of interracial roommate relationships. *Journal of Experimental Social Psychology*, *42*, 698–705.
- Tulving, E. (1983). *Elements of episodic memory*. Oxford, UK: Oxford University Press.
- Tulving, E., & Craik, F. I. M. (2000). *Handbook of memory*. Oxford, UK: Oxford University Press.
- Vanman, E. J., Saltz, J. L., Nathan, L. R., & Warren, J. A. (2004). Racial discrimination by low-prejudiced whites: Facial movements as implicit measures of attitudes related to behavior. *Psychological Science*, *15*, 711–714.
- Van Overwalle, F., & Van Rooy, D. (2001). How one cause discounts or augments another: A connectionist account of causal competition. *Personality and Social Psychology Bulletin*, *27*, 1613–1626.
- Visser, P. S., Bizer, G. Y., & Krosnick, J. A. (2006). Exploring the latent structure of strength-related attitude attributes. In M. Zanna (Ed.), *Advances in experimental social psychology*. New York: Academic Press.
- Walther, E. (2002). Guilty by mere association: Evaluative conditioning and the spreading evaluation effect. *Journal of Personality and Social Psychology*, *82*, 919–924.
- Walther, E., & Nagengast, B. (2006). Evaluative conditioning and the awareness issue: Assessing contingency awareness with the four-picture recognition test. *Journal of Experimental Psychology: Animal Behavior Processes*, *32*, 454–459.
- Walther, E., Nagengast, B., & Trasselli, C. (2005). Evaluative conditioning in social psychology: Facts and speculations. *Cognition and Emotion*, *19*, 175–196.
- Wentura, D. (1999). Activation and inhibition of affective information: Evidence for negative priming in the evaluation task. *Cognition and 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 judgmental tendencies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 456–469.
- Wicker, A. W. (1969). Attitude versus actions: The relationship of verbal and overt behavioral responses to attitude objects. *Journal of Social Issues*, *25*(4), 41–78.
- Wilson, T. D., & Hodges, S. D. (1992). Attitudes as temporary constructions. In A. Tesser & L. Martin (Eds.), *The construction of social judgment* (pp. 37–65). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, *107*, 101–126.
- Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its 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.
- Wittenbrink, B., & Schwarz, N. (Eds.) (2007). *Implicit measures of attitudes*. New York: Guilford Press.
- Wojnowicz, M. T., Ferguson, M. J., Dale, R., & Spivey, M. J. (2009). The self-organization of explicit attitudes. *Psychological Science*, *20*, 1428–1435.
- Wyer, R. S. (1973). Category ratings as ‘subjective expected values’: Implications for attitude formation and change. *Psychological Review*, *80*(6), 445–467.
- Yeh, W., & Barsalou, L.W. (2006). The situated nature of concepts. *American Journal of Psychology*, *119*, 349–384.