Social and Personality Psychology Compass 6/6 (2012): 438–454, 10.1111/j.1751-9004.2012.00438.x

Evaluation is a Dynamic Process: Moving Beyond Dual System Models

Jay J. Van Bavel[1]*, Yi Jenny Xiao[1], and William A. Cunningham[2]
[1] *New York University*
[2] *The Ohio State University*

Abstract
Over the past few decades, dual attitude/process/system models have emerged as the dominant framework for understanding a wide range of psychological phenomena. Most of these models characterize the unconscious and conscious mind as being built from discrete processes or systems: one that is reflexive, automatic, fast, affective, associative, and primitive, and a second that is deliberative, controlled, slow, cognitive, propositional, and more uniquely human. Although these models serve as a useful heuristic for characterizing the human mind, recent developments in social and cognitive neuroscience suggest that the human evaluative system, like most of cognition, is widely distributed and highly dynamic. Integrating these advances with current attitude theories, we review how the recently proposed Iterative Reprocessing Model can account for apparent dual systems as well as discrepancies between traditional dual system models and recent research revealing the dynamic nature of evaluation. Furthermore, we describe important implications this dynamical system approach has for various social psychological domains.

For nearly a century, psychologists have sought to understand the unconscious and conscious processes that allow people to evaluate their surroundings (Allport, 1935; Freud, 1930). Building on a model of the human mind rooted in classic Greek philosophy (Annas, 2001), many contemporary psychologists have characterized the mind as possessing discrete processes or systems: one that is evolutionarily older, reflexive, automatic, fast, affective, associative, and the other that is more uniquely human, deliberative, controlled, slow, cognitive, and propositional (see Figure 1). These dual process or system models have been highly influential throughout psychology for the past three decades (Chaiken & Trope, 1999). Indeed, a dual system model of the human mind permeates research in a wide range of psychological domains, such as attitudes and persuasion (Chaiken, 1980; Fazio, 1990; Gawronski & Bodenhausen, 2006; Petty & Cacioppo, 1986; Rydell & McConnell, 2006; Wilson, Samuel, & Schooler, 2000), stereotypes and prejudice (Crandall & Eshleman, 2003; Devine, 1989; Gaertner & Dovidio, 1986; Pettigrew & Meertens, 1995), person perception (Brewer, 1988; Fiske & Neuberg, 1990; Macrae & Bodenhausen, 2000), self-regulation (Baumeister & Heatherton, 1996; Freud, 1930; Hofmann, Friese, & Strack, 2009; Strack & Deutsch, 2004), moral cognition (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Haidt, 2001), learning and memory (Smith & DeCoster, 2000; Sun, 2002), and decision-making (Kahneman, 2003; Sloman, 1996).

Although dual system models provide generative frameworks for understanding a wide range of psychological phenomenon, recent developments in social and affective neuroscience suggest that the human evaluative system, like most of cognition, is widely distributed and highly dynamic (e.g., Ferguson & Wojnowicz, 2011; Freeman & Ambady,

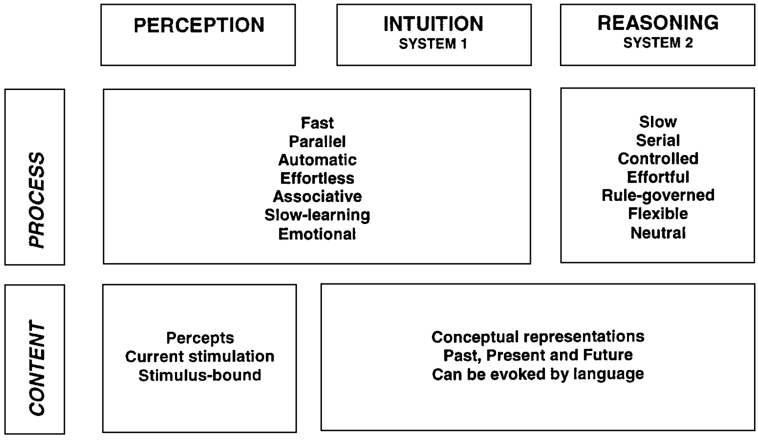

Figure 1 Illustrative example of the process and content of a dual system model (cited in Kahneman, 2003, p. 698).

2011; Scherer, 2009).[1] Integrating these advances with current attitude theories, we review how the recently proposed Iterative Reprocessing Model (Cunningham & Zelazo, 2007; Cunningham, Zelazo, Packer, & Van Bavel, 2007) can account for apparent dual systems as well as discrepancies between traditional dual system models and recent research revealing the dynamic nature of evaluation. The model also address why the nature of evaluative processing differs across people (e.g., Cunningham, Raye, & Johnson, 2005; Park, Van Bavel, Vasey, & Thayer, forthcoming). Although we focus primarily on dual models of attitudes and evaluation due to space constraints, we believe the premises of our dynamic model can be generalized to other domains where dual system models of typically invoked (Chaiken & Trope, 1999), including social cognition, self-regulation, prejudice and stereotyping, and moral cognition. Therefore, we very briefly discuss the implications of our model for these other domains in the final section of this paper and encourage interested readers to read our more extensive treatment of these issues in the domain of stereotypes and prejudice (Cunningham & Van Bavel, 2009a; Van Bavel & Cunningham, 2011) and emotion (Cunningham & Van Bavel, 2009b; Kirkland & Cunningham, 2011, forthcoming).

Attitudes and evaluation

Attitudes are one of the most central constructs in social psychology, yet there has been considerable debate regarding the most fundamental aspects of attitudes (Fazio, 2007; Schwarz & Bohner, 2001). Allport (1935) defined an attitude as "a mental and neural state of readiness, organized through experience, exerting a directive or dynamic influence upon the individual's response to all objects and situations with which it is related" (p. 810). Throughout the history of attitude research, theorists and researchers have attempted to provide a complete yet parsimonious definition of this construct. Well-known examples include the one-component perspective (Thurstone, 1928), the tripartite model (Affective, Behavior, Cognition; Katz & Stotland, 1959; Rosenberg & Hovland, 1960), and more recently, a host of dual attitudes (e.g., Greenwald & Banaji, 1995; Rydell & McConnell, 2006; Wilson et al., 2000) and dual process models (e.g., Chaiken, 1980; Fazio, 1990; Gawronski & Bodenhausen, 2006; Petty & Cacioppo, 1986).

It is widely assumed that attitudes are stored associations between objects and their evaluations, which can be accessed from memory very quickly with little conscious effort

(Fazio, 2001; Fazio, Sanbonmatsu, Powell, & Kardes, 1986; but see Schwarz, 2007). For example, people categorize positive and negative words more quickly when these words are preceded by a similarly valenced stimuli, suggesting that attitudes are automatically activated by the mere presence of the attitude object in the environment (Fazio et al., 1986). Moreover, people may have access to evaluative information about stimuli prior to their semantic content (Bargh, Litt, Pratto, & Spielman, 1989; but see Storbeck & Clore, 2007). Such research has led to the conclusion that the initial evaluative classification of stimuli as good or bad can be activated automatically and guide the perceiver's interpretation of his or her environment (Houston & Fazio, 1989; Smith, Fazio, & Cejka, 1996).

Dual attitudes and dual process models of attitudes

The recent development of a wide variety of implicit attitude measures (Petty, Fazio, & Briñol, 2009; Wittenbrink & Schwarz, 2007), including measures of human physiology (Cunningham, Packer, Kesek, & Van Bavel, 2009), has fueled an explosion of research on dual attitude/process/system models of attitudes and evaluations (see Table 1). Most of these models infer dual process architecture from observable dissociations between implicit and explicit measures of behavior (e.g., Dovidio, Kawakami, & Gaertner, 2002; McConnell & Leibold, 2001; Rydell & McConnell, 2006). Although many dual models generally share a common set of assumptions about the human mind, the specific features of each model differ. Therefore, we propose a rough taxonomy to characterize different classes of these models. ''Dual attitudes models'' tend to dichotomize the *representations* of attitudes into distinct automatic versus controlled constructs (Greenwald & Banaji, 1995; Rydell & McConnell, 2006; Wilson et al., 2000).[2] In contrast, ''dual process models'' tend to dichotomize the *processing* of attitudinal representations into automatic versus controlled processes.[3] There is considerable debate over whether these two types of processes are independent or communicate with one another (i.e., information from one system is available to the other system) (Fazio, 1990; Gawronski & Bodenhausen, 2006; Gilbert, Pelham, & Krull, 1988; Petty, Brinol, & DeMarree, 2007). In the latter case, interdependent dual process models have generally been proposed to operate in a corrective fashion, such that ''controlled'' processes can influence otherwise ''automatic'' responses (e.g., Wegener & Petty, 1997). Although dual *attitudes* models likely require dual *processes* to integrate different attitudinal representations into evaluations and behaviors, dual *process* models are less likely to require the assumption of dual *attitude representations* (e.g., Fazio, 1990). For the purpose of clarity, we use ''dual *system* models'' to capture models that assume dual attitudes and processes that do not interact (e.g., Rydell & McConnell, 2006; Wilson et al., 2000).

There are, of course, many ways to hook up a dual system (see Gilbert, 1999 for a discussion). A complete discussion of all possible dual models and interconnections between these systems is beyond the scope of this article. Therefore, we focus on several core premises that many models have in common. Likewise, we focus on the core premises from our own model – rather than an exhaustive discussion (e.g., Cunningham et al., 2007) – in order to communicate key similarities and differences between dual models and our proposed dynamic model.[4] Furthermore, we recognize that dual models and our proposed dynamic model do not exhaust all types of models of attitudes and evaluation – some extant models do include more than two processes (e.g., Beer, Knight, & D'Esposito, 2006; Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005) and many allow for interactive processes that operate in a *post hoc*, corrective fashion (e.g., Chen & Chaiken, 1999; Gawronski &

Table 1 A sample of dual models cited in attitude research arranged in chronological order

Model name	Core characteristics	Total citations	References
Heuristic versus systematic model	Heuristic versus systematic Effortless versus effortful	2202	Chaiken (1980)
Central and peripheral routes to persuasion	Central versus peripheral	4192	Petty and Cacioppo (1986)
Automatic and controlled components of stereotyping and prejudice	Automatic versus controlled Unintentional versus intentional Spontaneous versus deliberative	3363	Devine (1989)
MODE model	Unconscious versus conscious Spontaneous versus deliberative/reasoned	1424	Fazio (1990)
Implicit social cognition	Implicit versus explicit Automatic versus controlled	2660	Greenwald and Banaji (1995)
Two systems of reasoning	Associative versus Rule-based Deliberative/analytic versus automatic	1627	Sloman (1996)
Flexible Correction Model		302	Wegener and Petty (1997)
Dual-process models	Quick versus slow Effortless versus effortful Unconscious versus conscious Associative versus rule-based	610	Smith and DeCoster (2000)
Model of dual attitudes	Implicit versus explicit Automatic versus controlled	1209	Wilson et al. (2000)
Reflective and impulsive determinants of social behavior	Reflective versus impulsive	869	Strack and Deutsch (2004)
Predictive model of implicit and explicit attitudes	Implicit versus explicit Automatic versus controlled Unconscious versus conscious Effortless versus effortful	130	Perugini (2005)
Dual system of reasoning	Implicit versus explicit Fast versus slow Spontaneous versus deliberative Associative versus rule-based	131	Rydell and McConnell (2006)
APE model	Implicit versus explicit Automatic versus deliberative Associative versus propositional	566	Gawronski and Bodenhausen (2006)

Total citations obtained from Google Scholar on February 9, 2012 represent a proxy for scientific impact (i.e., actual number of scientific citations may be more or less than the number of citations reported here).

Bodenhausen, 2006; Strack & Deutsch, 2004; Wegener & Petty, 1997). However, few (to our knowledge) articulate how ''controlled'' processes might influence more ''automatic'' processes in an *a priori* fashion.

In the following sections, we describe an alternative, dynamic model of evaluation in which a constellation of widely distributed ''automatic'' and ''controlled'' processes interact in a dynamic fashion to process evaluations. This model characterizes the human brain as a parallel system that generates evaluations by integrating the results of computations performed by a widely distributed network of component processes (Frank, Cohen, & Sanfey, 2009; McClelland & Rumelhart, 1986; O'Reilly & Munakata, 2000; Rumelhart & McClelland, 1986). We place the terms ''automatic'' and ''controlled'' in quotations because we believe they do not reflect *bona fide* attitudes, processes or systems, but merely a useful heuristic to describe the number and nature of the particular interactive component processes currently involved in evaluation (Cunningham & Johnson, 2007).[5]

The iterative reprocessing model of evaluation

We argue that traditional conceptions of ''automatic'' and ''controlled'' processes as separable bona fide systems should be replaced by models that invoke the integration of multiple dynamic computational processes. In this paper, we describe one such approach – the Iterative Reprocessing (IR) Model (Cunningham & Zelazo, 2007; Cunningham et al., 2007). A fundamental assumption of the IR Model is that evaluative processes involve a series of iterative cycles: with every iteration, the current evaluation of a stimulus can be adjusted in light of additional contextual and motivational information to create an updated evaluation in line with finer stimulus detail, the context, and/or current goals. As such, the IR Model shares two important features with dual system/process models. First, stimuli evoke rapid perceptual and evaluative responses, and second, perceivers can become aware of these initial responses, and, with the right motivation and opportunity, modulate or elaborate upon them (for a more complete discussion of the similarities and differences, please see Cunningham & Zelazo, 2007; Cunningham et al., 2007). More pertinent to the current discussion, the IR Model differs from most dual process/system models in the following ways.

First, in contrast to dual *attitude* models (e.g., Rydell & McConnell, 2006; Wilson et al., 2000), the IR Model does not assume distinct (implicit versus explicit) attitudinal *representations* stored in memory. Rather, the IR model characterizes evaluation as an emergent property of multiple processes that unfold over time. As such, differences in evaluation are largely due to differences in information processing, rather than qualitatively different attitudinal representations stored in discrete memory systems. Distinguishing between attitudes as relatively stable representations, and evaluations as the current state of evaluative processing, requires an explanation of how attitudes are transformed into evaluations. The IR model resolves this issue by proposing a connectionist framework in which *attitudes* are represented as (relatively stable) unit weights, whereas *evaluations* reflect the current pattern of activation of the units (Cunningham et al., 2007).[6]

In the context of evaluation, the weights can be conceptualized as having valence and intensity. Thus, what many dual process/system models term ''controlled'' processing may represent a change in the current activation pattern, and not necessarily in the unit weights (which can be activated easily, but generally change more slowly). This distinction allows for the apparent *stability* (i.e., the weights are relatively stable) and occasional *flexibility* (i.e., the patterns of activation are relatively variable) of implicit attitude measures without requiring distinct implicit versus explicit attitudinal representations. Imagine someone's attitude towards members of a racial minority. The weights may be conceptualized as having a negative valence with modest intensity. When this perceiver encounters a racial minority member, his/her evaluation (i.e., the current activation pattern) may be easily influenced by factors such as current context, motivation and goals. For example, even given relatively stable unit weights, the activation pattern regarding an other-race member may differ depending on whether the target belongs to the same team as the perceiver's (Van Bavel & Cunningham, 2009a). As such, evaluations may shift radically even if the underlying attitude toward the other race has not. In sum, evaluations are constructions consisting of relatively dynamic activation patterns that are sensitive to shifting contextual and motivational influences, and consist of a subset of input units with relatively stable weights, rather than being a veridical instantiation of our internal attitudes.[7]

Second, in contrast to most dual *process/system* models, the IR Model does not assume only two bona fide evaluative systems at work in the human brain (e.g., Gawronski & Bodenhausen, 2006; Smith & DeCoster, 2000; Strack & Deutsch, 2004). Instead, building on recent research on the functional anatomy of the human brain, the IR Model proposes that there are many highly interactive neural systems engaged in information processing (see Figure 1 for a simplified version of the model).[8] Importantly, information can propagate forward *or* backward through the system, meaning that evaluative processes are part of an iterative cycle that unfolds in a dynamic fashion over time and is mutually constrained by so-called bottom-up and top-down influences (termed bidirectional excitation), until the network eventually settles into a stable state. As shown in Figure 2, brain regions such as the amygdala, ventral striatum, and posterior orbitofrontal cortex are normally engaged in initial evaluative processing of stimuli. The thalamus, sensory cortices, and bodily states provide input into initial and subsequent iterations – even within a time period that is typically considered ‘‘automatic’’ (i.e., a few hundred milliseconds). As information about the stimulus is reprocessed, higher-order brain regions, such as the PFC, influence the evaluation, *and* also reseed initial evaluative processing

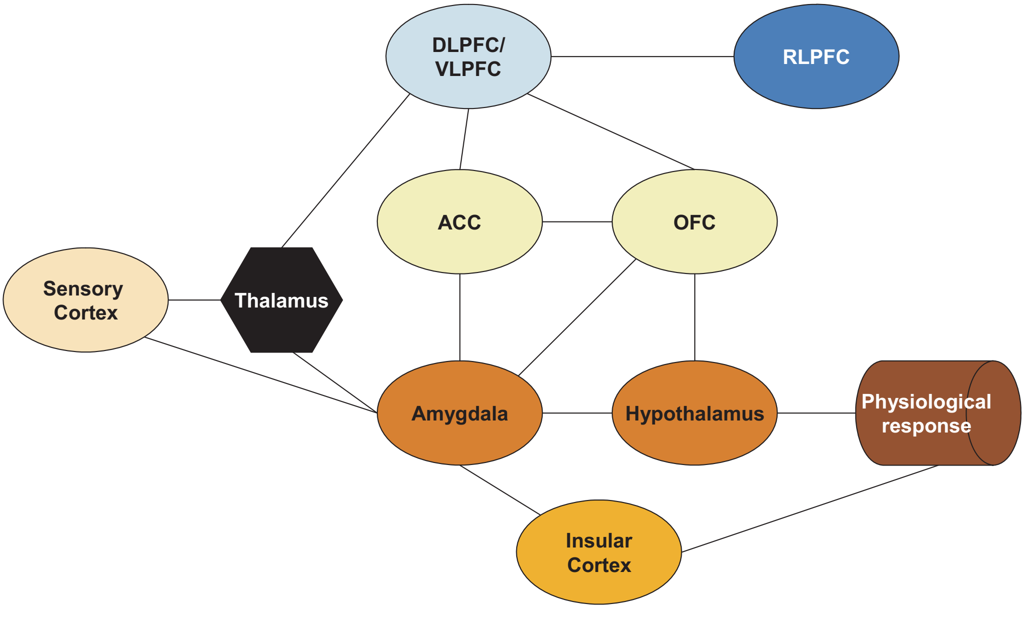

Figure 2 A simplified model of the brain regions underlying evaluation (reproduced with permission from Cunningham et al., 2007). Links between regions discussed in the text are denoted by solid lines (note that not all anatomical links or brain regions are represented). Information about a stimulus may be processed by the thalamus and projected to the amygdala, leading to an initial evaluation that is associated with a tendency to approach or avoid the stimulus. Additional iterations can also include processing by the insula, orbitofrontal cortex (OFC), and anterior cingulate cortex (ACC), as well as more detailed sensory processing. Visceral changes following evaluation are guided by the hypothalamus and other regions associated with autonomic control. Additional recruitment of the prefrontal cortex, especially regions of the ventrolateral prefrontal cortex (VLPFC), dorsolateral prefrontal cortex (DLPFC), and rostrolateral prefrontal cortex (RLPFC), may subserve goal-oriented reprocessing of stimuli and the regulation of evaluative processing by enhancing or suppressing features of the stimulus or situation. Information can propagate forward or backward through the system.

by influencing subsequent information processing (see Cunningham et al., 2007).[9] In this way, information is sent up and down the neuroaxis to generate and continually update an evaluation.

An important implication is that an evaluation not only can be updated based on additional contextual information and motivational factors, but these ostensibly top-down factors can shape ostensibly bottom-up responses. For example, the amygdala plays an important role in detecting an affectively significant stimulus, even when it is presented below conscious awareness (Cunningham et al., 2004; Morris, Ohman, & Dolan, 1998; Whalen et al., 1998), but is also sensitive to the influence of goals (Cunningham et al., 2005; Cunningham, Van Bavel, & Johnsen, 2008), expectations (Kim et al., 2004) and contextual factors (Forbes, Cox, Schmader, & Ryan, forthcoming; Van Bavel, Packer, & Cunningham, 2008). For example, White participants who were arbitrarily assigned to a mixed-race group showed greater amygdala activity towards in-group members, regardless of their race (Van Bavel et al., 2008). These results indicate that seemingly trivial aspects of the social context can shape amygdala responses in the presence of affectively significant cues, like race. In other words, ostensibly ''automatic'' components of the evaluative system, such as the amygdala, do not simply respond in a reflexive fashion to stimuli. Consistent with the IR Model, these component processes are sensitized to respond to motivationally relevant stimuli by higher-order systems.

Third, an important implication of the IR Model is that goals and contexts may not only influence the current evaluation, but that prior states of the evaluative system (time – 1) set the stage for a rapid appraisal of the same or different subsequent stimuli (at

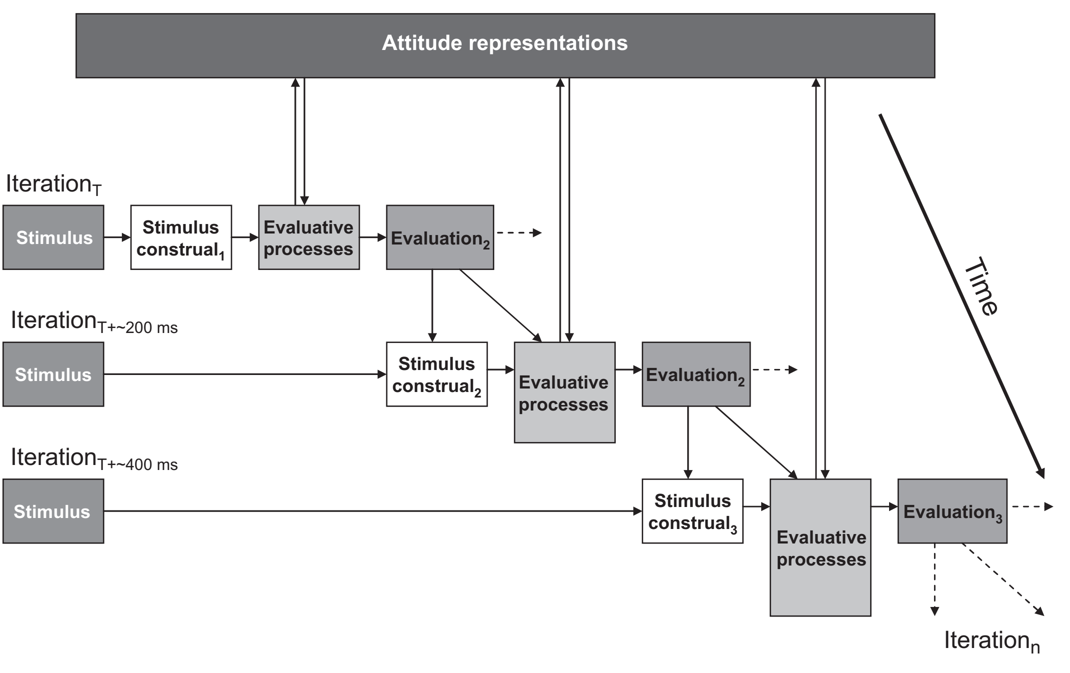

Figure 3 The perceptual and evaluative cycle (reproduced with permission from Cunningham et al., 2007). During each iteration, evaluative processes retrieve attitudinal representations to generate an evaluation relying on a particular construal(s) of the stimulus. This evaluation may influence the next iteration of evaluative processing, direct behavior, or both. In general, the complexity of evaluative processing (and the resulting evaluation) increases with additional iterations.

time = 0) (Smith et al., 1996; also see Figure 3). We speculate here for the first time that this process may be unconscious in some circumstances and may therefore be described as a *pre-appraisal* process. Thus, the preceding context and motivational state of an organism influence ongoing evaluative processes, and preceding evaluative processes can influence the perceived stimuli, context and motivational state of an organism (e.g., Ferguson & Bargh, 2004; Van Bavel & Cunningham, 2009a). Although most dual models do not argue that prior states are irrelevant, these models are nevertheless surprisingly silent on this issue. As a consequence, relatively little work has been conducted on the influence of prior states on ostensibly automatic evaluations of stimuli (e.g., Ferguson & Bargh, 2004). By making this aspect of evaluation explicit in our dynamic model of evaluation, we hope to motivate research that integrates work on emotion, belief, and goals with the work on attitudes and evaluation to provide a more complete understanding of the human evaluative system.

In a recent test of this idea, we used electroencephalography (EEG) to examine the influence of motivational state on the earliest aspects of racial bias (Cunningham, Van Bavel, Arbuckle, Packer, & Waggoner, 2012). We manipulated approach and avoidance motivational states by having participants push or pull a joystick for a block of trials in which they were randomly presented with a series of White and Black faces. Consistent with previous work on own-race bias, we observed relatively greater responses to White than Black faces in a very early positive going event-related potential occurring approximately 100 ms after stimulus onset (P100). Critically, this racial bias was reduced in the approach condition, suggesting that very early aspects of social perception may be modulated by motivational states (see also Amodio, 2010). Thus, in contrast to most dual process/system models, the evaluation unfolds over time in a dynamic fashion and motivation shapes rapid responses prior to exerting corrective control. As a consequence, chronic and contextual differences in motivation ultimately shape downstream perceptions (e.g., Caruso, Mead, & Balcetis, 2009; Xiao & Van Bavel, forthcoming), and behaviors (e.g., Van Bavel & Cunningham, 2012; Van Bavel, Swencionis, O'Connor, & Cunningham, 2012).

Fourth, the IR model specifically addresses the graded and multi-dimensional nature of automaticity. According to the IR model, the evaluative process arises through multiple iterations – so-called cognitive and affective processes work in concert rather than independently. Although the first iteration(s) in the evaluative process may appear similar to the ''automatic'' processes described by most traditional dual system models (e.g., Devine, 1989; Gawronski & Bodenhausen, 2006; Greenwald & Banaji, 1995), the IR model conceptualizes evaluations as emerging in a graded fashion, such that *relatively* more ''automatic'' evaluations reflect fewer iterations (Cunningham et al., 2007). Thus, many interactive dual process models are encapsulated within a dynamic systems approach. However, attention, perception, and evaluation unfold in a dynamic fashion, as different processes and concerns are integrated in an evaluation (see Brosch & Van Bavel, 2012 for an example in the domain of emotional attention). It is difficult to imagine how this constellation of multiple interactive components can be neatly carved into two discrete systems. The IR model assumes that brain systems are organized hierarchically – rather than independently – such that what are often considered ''automatic'' processes influence and are influenced by what are often considered ''controlled'' processes in a dynamic, iterative fashion.[10]

To examine the interactions between ''automatic'' and ''controlled'' components of social evaluation, we examined the brain activity of White participants while they viewed Black and White faces during event-related functional magnetic resonance imaging (Cunningham et al., 2004). Consistent with previous research, when the faces were presented for 30 ms, amygdala activity was greater for Black than for White faces (see Cunningham & Van Bavel, 2009a; Van Bavel & Cunningham, 2009b for a review). In addition, participants with the largest difference in amygdala activation between Black and White faces had the largest racial bias on an indirect behavioral attitude measure (i.e., the Implicit Association Test). However, when the faces were presented for 525 ms, this pattern of racial bias in amygdala activity was significantly reduced and regions of frontal cortex associated with controlled processing (i.e., lateral prefrontal cortex, anterior cingulate) showed greater activation for Black than for White faces. Moreover, activity in the dorsolateral prefrontal cortex and anterior cingulate was associated with a reduction in Black-White differences in amygdala activity from the 30-ms to the 525-ms condition. The results of this study are consistent with the notion that that ''controlled'' processes may modulate ''automatic'' components of evaluation (see also Forbes et al., forthcoming). As we noted above, this research is consistent with several, but not all, dual process models.

Although one may consider the relatively coarse perceptual and evaluative information that guides an initial evaluation ''automatic,'' once additional iterations of processing occur spanning multiple processes, the distinction between ''automatic'' and ''controlled'' becomes less clear. Moreover, according to our model, so-called top-down processes can shape initial responses – perhaps ''unconsciously'' – and bottom-up processes can alter ''conscious'' experience and behavior. In other words, the dynamic and interactive nature of this system challenges the very definition of automaticity employed in several models of attitudes and evaluations. We believe this feature of our dynamic model of evaluation is a departure from the vast majority of dual models – even ones that allow for a *post hoc* correction of an automatic response. Moreover, it calls into question the meaning of terms like ''bottom-up'' and ''top-down'' that invoke temporal order (e.g., fast versus slow), cognitive processes (e.g., unconscious versus conscious) and neural mechanisms (e.g., the occipital versus frontal cortices).[11]

Implications and future research

Dual models have permeated a number of core research domains in psychology. In the following section, we very briefly describe how our dynamic approach to evaluation might have important implications for dominant theoretical perspectives in stereotyping and prejudice (Devine, 1989; Dovidio, Evans, & Tyler, 1986; Fazio, Jackson, Dunton, & Williams, 1995; Gaertner & Dovidio, 1986), social cognition and person perception (Brewer, 1988; Fiske & Neuberg, 1990), self-regulation (Baumeister & Heatherton, 1996; Carver, 2005; Hofmann et al., 2009; Strack & Deutsch, 2004), and moral cognition (Greene et al., 2001; Haidt, 2001). However, we believe that these insights might be fruitfully applied to virtually any other domain in psychology that has been influenced by dual notions of the human mind (Chaiken & Trope, 1999).

Research on implicit cognition has suggested that evaluations, intuitions, stereotypes and prejudice can operate automatically and efficiently without the perceiver's conscious control and despite best intentions to the contrary (Greenwald & Banaji, 1995). In the past 15 years, a number of implicit measures have been developed (see Petty et al., 2009; Wittenbrink & Schwarz, 2007) to examine the operating principles of these ''automatic'' versus ''controlled'' components of the human mind. Initial evidence of the relatively modest association between implicit and explicit measures (see Banaji, 2004 for a discussion) and the dissociation between implicit and explicit forms of behavior (e.g., Jellison, McConnell, & Gabriel, 2004) were taken as support for dual models. Although dual models dominated this field of research for the past two decades, recent empirical findings have provided evidence that the strong form of dual models struggles to explain.

For instance, a core feature of automatic processes is their uncontrollability (Bargh, 1994). However, research has demonstrated the malleability of ostensibly automatic process. Influences on these processes span a wide range, including different sources of motivation – affiliative motives, self-image motives, motivation to be non-prejudiced, motivation channeled by social identities, stereotype suppression strategies – implementation intentions and affective priming, and a wide range of contextual factors (Blair, 2002). For example, mere categorization with an arbitrary group is sufficient to attenuate very rapid racial biases towards in-group members (Van Bavel & Cunningham, 2009a). Likewise, egalitarian goals (Moskowitz, Gollwitzer, Wasel, & Schaal, 1999) and deliberate implementation intentions (Gollwitzer & Schaal, 1998) can reduce biases on implicit measures of intergroup bias (Mendoza, Gollwitzer, & Amodio, 2010; Stewart & Payne, 2008). Researchers have demonstrated and are continuing to uncover the malleability of automatic processes, and their susceptibility to influence from various top-down processes and contextual factors. According to the IR model, situational factors may interact with other higher-order influences, such as motivation, to modulate automatic processes. For example, by incorporating contextual backgrounds (e.g., threatening versus non-threatening) into an evaluative priming procedure, individuals high in non-prejudice motivation were able to inhibit their racial bias on an implicit measure when contextual cues were associated with prejudice (e.g., threatening cues) (Maddux, Barden, Brewer, & Petty, 2005). The IR Model, due to its dynamic nature, can account for these apparent "top-down" influences by invoking a mechanism by which motivation can alter the rapid construal of stimuli (see Van Bavel & Cunningham, 2011 for a more detailed discussion).

Likewise, rather than portraying self-regulation as an effort to overcome hedonic impulses in favor of more deliberate evaluations (e.g., Hofmann et al., 2009), we argue that self-regulation should be considered a dynamic process (e.g., Carver & Scheier, 1998). For instance, a series of studies by Ferguson and Bargh (2004) found that active goal pursuit led to more positive automatic evaluation of goal-relevant objects relative to goal-irrelevant objects, which also predicted downstream behavioral intentions towards these goal-relevant stimuli (see also Hofmann, Deutsch, Banaji, & Lancaster, 2010). Research on the malleability of evaluative processes speaks to the highly interactive nature of motivational states and evaluative processes. As such, the IR approach suggests that successful long-term self-regulation will likely require chronic re-construals of tempting situations and stimuli (Magen, 2007) and more reflexive action control (Bargh, Gollwitzer, Lee-Chai, Barndollar, & Trötschel, 2001).

Similarly, our approach makes a number of different predictions than the extant dual process models of moral cognition (Greene et al., 2001; Haidt, 2001). Foremost, we predict that people can guide early moral intuitions through reasoned processes. For example, people may be able to flexibly construe a situation or stimulus in moral or alternative terms depending on their goals and beliefs, which will direct attention, modulate perception and guide consequent emotional intuitions (Van Bavel, Packer, Johnsen, & Cunningham, 2012). As such, reason need not be a slave to the passions (but see Haidt, 2001). Indeed, "reasoning" processes may able to guide whether or not people even detect a morally relevant stimulus in the first place through *pre-appraisal* processes. Deliberating about one's values and beliefs may not only justify or correct for an initial emotional intuition, but may sensitize one to certain actions or events before they occur, which would then have down-steam consequences for emotional responding.

Discussion

> "Psychology also attempts to conceptualize what it is doing…by the construction of oppositions – usually binary ones…far from providing the rungs of a ladder by which psychology gradually climbs to clarity, this form of conceptual structure leads rather to an ever increasing pile of issues, which we weary of or become diverted from, but never really settle." –Newall (1973, p. 287)

Dual models have dominated the study of human psychology for the past century, and descriptions of the human mind for much longer (Descartes, 1641/1984). Western intellectual tradition is littered with metaphors designed to capture the internal battle between automatic and controlled processes. Characterizations of the battle between systems range from men wrestling with wild steeds and raging bulls to stubborn elephants. Likewise, due to their intuitive appeal and heuristic value, dual system models have dominated our conception of attitudes and evaluation for several decades. The popularity of these dual models has shaped the way we conceive both the attitude construct and the evaluative process. Specifically, attitudes and evaluations have been treated as having two distinct components or processes – one implicit component operating automatically, and the other explicit component operating under conscious control. In contrast to dual models, we argue that traditional conceptions of "automatic" and "controlled" processes may be best characterized as a heuristic that reflects a host of dynamic, underlying computational processes.

In contrast, the fundamental assumption underlying our model is that brain systems are organized hierarchically, such that so-called "automatic" processes influence *and* are influenced by "controlled" processes. Therefore, controlled processes do not merely override automatic ones – these processes work in a dynamic, interactive fashion to generate evaluations. The IR model proposes there are many highly interactive neural systems engaged in information processing, as opposed to only two qualitatively distinct processes at work in the human brain, which is consistent with recent developments in affective and cognitive neuroscience. These influences come online in a continuous and interactive fashion.

Moreover, the IR model does not make the assumption that there are distinct (implicit versus explicit) attitudinal representations in memory, but proposes that different evaluations are emergent properties of differences in information processing. Following connectionist models of memory, attitudes can be conceptualized as the different connection weights, whereas evaluations are the current activation pattern of the units. As such, we make a distinction between attitudes and evaluations. We believe this approach has important implications for a wide variety of topic areas within psychology.

According to the Iterative Reprocessing (IR) Model, evaluative processes are part of an iterative cycle: with every iteration, the current evaluation of a stimulus can be adjusted in light of additional contextual and motivational information in order to create an updated evaluation in line with finer stimulus detail, the context, and/or current goals. Although bona fide dual models have heuristic value, they cannot bear the weight of recent evidence on the continuous nature of automaticity and susceptibility of automatic evaluation to the influence of various processes and contextual factors. On the other hand, our approach not only reconciles these discrepancies, but also respects recent developments in affective and cognitive neuroscience on the highly dynamic nature of the human evaluative system.

In short, we argue that evaluation is shaped by the nature of the immediately active representations and the current dynamics of the system (including a combination of internal and external factors that constrain information processing). We believe that the proposed dynamic model of evaluation has implications not only for our understanding of the concept of automaticity (including its four components),[12] but should also lead to novel constructs and processes (e.g., pre-appraisal") (because pre-appraisal is a process not a construct) and predictions for research on topics as disparate as stereotypes and prejudice, self-regulation, and morality. We also propose future research questions that could specifically address predictions consistent with the IR model, but not necessarily with traditional dual models. More generally, we argue that dynamic models of human psychology offer a more powerful framework for explaining the extant literature on attitudes and evaluation, more accurately reflect the underlying cognitive and neural mechanisms that guide behavior, and offer a number of provocative directions for future research.

Acknowledgment

The authors would like to thank Jim Uleman, Jamil Zaki, Michael Robinson, three anonymous reviewers, Sophie Wharton, and several members of the New York University *Social Perception and Evaluation Lab* for their thoughtful comments on this manuscript.

Short Biographies

Jay Van Bavel is broadly interested in the dynamic nature of perception and evaluation. Specifically, his work examines how our social identities and personal values shape our perceptions and evaluations of the social and physical world. He approaches these issues from a social neuroscience perspective, blending theory and methods from social psychology and cognitive neuroscience. He received the SPSSI Dissertation Award and was a finalist for the SESP Dissertation Award in 2009. He currently is an Assistant Professor at the New York University Department of Psychology. He completed his BA in psychology at the University of Alberta, a PhD in psychology at the University of Toronto and a postdoctoral fellowship at The Ohio State University.

Jenny Xiao is broadly interested in studying social categorization and intergroup relations, by blending theories and techniques from social cognitive psychology and neuroscience. Specifically, her work explores how high-level social psychological constructs such as social identity alter low-level cognitive and perceptual processes. For example, she has found that social group membership and intergroup threat work in concert to shape perceptions and representations of physical distance, which in turn lead to detrimental intergroup consequences. Jenny has received graduate student research awards and fellowships from SPSSI and SPSP. Jenny Xiao received a dual BA in psychology and biology from Bard College and is currently a PhD student In the Social Psychology Program at New York University.

William Cunningham's research examines the dynamic nature of cognition and emotion. Specifically, his work examines core questions in affective science (the nature and representation of attitudes and emotions) from a cognitive neuroscience perspective. He received the SAGE Young Award in 2009 from the Foundation for Personality and Social Psychology, and the Janet T. Spence Award for Transformative Early Career Contributions in 2011. He currently is on faculty at The Ohio State University, and is associate professor of psychology and the associate director of the Center for Cognitive and Behavioral Brain Imaging. Before OSU, he was on faculty at the University of Toronto. He holds a BA and MA in psychology from the College of William & Mary, and a PhD in psychology from Yale University.

Endnotes

* Correspondence address: Department of Psychology, New York University, 6 Washington Place, New York, NY 10003, USA. Email: jay.vanbavel@nyu.edu

[1] We note that there is not a clear consensus on this issue. Several models grounded in social and affective neuroscience characterize the brain in terms of two discrete systems (e.g., Lieberman, 2003; Smith & DeCoster, 2000).
[2] The terms and definitions we have attributed for ''dual attitude'', ''dual process'' and ''dual system'' models are not necessarily consensual. Due to the proliferation of dual models, we have used the terminology to create a rough taxonomy of popular models.
[3] It is worth noting that dual *system* models may require dual *processes* to integrate attitudinal representations into evaluation and behavior.
[4] Although it is beyond the scope of the current article, in future work we intend to specify how our approach differences from other models by focusing on the core principles underlying different models (e.g., interactive versus non-interactive, two processes versus multiple processes, etc.) rather than summing across a large class of theories that often differ on these principles.
[5] Our intuitions about the presence of dual attitudes/processes/systems may reflect our meta-cognition concerning the number or nature of component processes engaged in a particular task (see Johnson & Hirst, 1993 for a similar argument).
[6] Eagly and Chaiken (2007) make a similar definitional distinction between ''inner tendencies'' and ''evaluations''. However, to our knowledge, their definition is not based on the same underlying computational principles.
[7] Although we argue that evaluations (i.e., patterns of activation) are constructed (Schwarz, 2007), we also believe that relatively stable attitudes (i.e., weights) inform evaluations (Fazio, 2007).
[8] See work on parallel distributed process models of social cognition for a similar point (e.g., Kunda, 1999; Kunda & Thagard, 1996; Smith, 1996).
[9] Not only is neural connectivity largely bi-directional, recent research shows that the sub-cortical input to the cortex is actually fairly weak, whereas the putative ''feedback'' connections are immense (Douglas & Martin, 2004). These findings have lead to new theoretical models of brain function arguing that the backward connections actual play the primary role in generating predictions whereas the forward connections are more likely to play the role of providing what is traditionally considered feedback (i.e., prediction error signals; Friston, 2005). Indeed, the modulation afforded by cortical systems (memory, attention, motivation) is likely to affect large masses of cells than the sensory signals themselves (Logothetis, 2008). In other words, ''controlled'' processes, mediated by signals from the frontal and parietal networks, can incorporate expectations, goals, bodily states and contextual information into representations that are deemed most relevant in a current context (see Miller & Cohen, 2001), which can then lead to different ''automatic'' evaluative responses.
[10] Thus, our model is not only different from classic dual process/system models, but also models that propose a single mental process (e.g., Kruglanski & Thompson, 1999).
[11] Although these levels of explanation are often in agreement with our model, we argue that it is important to treat these levels of explanation independently to fully understand the dynamics of the evaluative system. Indeed, the semantics often make it difficult for us to communicate how this model is different from other models because terms like ''control'' or ''top-down'' usually evoke a set of assumptions that may not be captured by a specific phenomenon.
[12] Indeed, we argue that awareness – a core dimension of automaticity that is subjectively experienced as dichotomous – may be more accurately characterized as graded rather than all-or-none (Kouider, de Gardelle, Sackur, & Dupoux, 2010).

References

Allport, G. (1935). Attitudes. In C. Murchison (Ed.), *A Handbook of Social Psychology* (pp. 798–844). Worcester, MA: Clark University Press.

Amodio, D. M. (2010). Coordinated roles of motivation and perception in the regulation of intergroup responses: Frontal cortical asymmetry effects on the P2 event-related potential and behavior. *Journal of Cognitive Neuroscience*, **22**, 2609–2617.

Annas, J. (2001). *Voices of Ancient Philosophy: An Introductory Reader*. New York, NY, USA: Oxford University Press.

Banaji, M. R. (2004). The opposite of a great truth is also true: Homage of Koan #7. In J. T. Jost & M. R. Banaji (Eds.), *Perspectivism in Social Psychology: The yin and Yang of Scientific Progress* (pp. 127–140). Washington, DC: American Psychological Association.

Bargh, J. A. (1994). The Four Horsemen of automaticity: Awareness, efficiency, intention, and control in social cognition. In J. R. S. Wyer & T. K. Srull (Eds.), *Handbook of Social Cognition* (2nd edn, pp. 1–40). Hillsdale, NJ: Erlbaum.

Bargh, J. A., Gollwitzer, P. M., Lee-Chai, A., Barndollar, K., & Trötschel, R. (2001). The automated will: Nonconscious activation and pursuit of behavioral goals. *Journal of Personality & Social Psychology*, **81**, 1014–1027.

Bargh, J. A., Litt, J., Pratto, F., & Spielman, L. A. (1989). On the preconscious evaluation of social stimuli. In A. F. Bennett & K. M. McConkey (Eds.), *Cognition in Individual and Social Contexts* (Vol. 3, pp. 357–370). Amsterdam: Elsevier-North Holland.

Baumeister, R. F., & Heatherton, T. F. (1996). Self-regulation failure: An overview. *Psychological inquiry*, **7**, 1–15.

Beer, J. S., Knight, R. T., & D'Esposito, M. (2006). Controlling the Integration of Emotion and Cognition: The Role of Frontal Cortex in Distinguishing Helpful from Hurtful Emotional Information. *Psychological Science*, **17**, 448–453.

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

Brewer, M. B. (1988). A dual process model of impression formation. In T. K. Srull & R. S. Wyer (Eds.), *Advances in Social Cognition* (Vol. 1, pp. 1–36). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Brosch, T., & Van Bavel, J. J. (2012). *The Flexibility of Emotional Attention: Accessible Social Identities Guide Rapid Attentional Orienting*. New York: Department of Psychology, New York University. Unpublished manuscript.

Caruso, E., Mead, N., & Balcetis, E. (2009). Political partisanship influences perception of biracial candidates' skin tone. *Proceedings of the National Academy of Sciences*, **106**, 20168–20173.

Carver, C. S. (2005). Impulse and constraint: Perspectives from personality psychology, convergence with theory in other areas, and potential for integration. *Personality and Social Psychology Review*, **9**, 312–333.

Carver, C. S., & Scheier, M. F. (1998). *On the Self-Regulation of Behavior*. New York: Cambridge University Press.

Chaiken, S. (1980). Heuristic versus systematic processing and the use of source versus message cues in persuasion. *Journal of Personality & Social Psychology*, **39**, 752–766.

Chaiken, S., & Trope, Y. (1999). *Dual-Process Theories in Social Psychology*. New York: Guilford Press.

Chen, S., & Chaiken, S. (1999). The heuristic-systematic model in its broader context. In S. Chaiken & Y. Trope, (Eds.), *Dual-Process Theories in Social Psychology*. (pp. 73–96). New York: Guilford Press.

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

Crandall, C. S., & Eshleman, A. (2003). A justification-suppression model of the expression and experience of prejudice. *Psychological Bulletin*, **129**, 414–446.

Cunningham, W. A., & Johnson, M. K. (2007). Attitudes and evaluation: Toward a component process framework. In E. Harmon-Jones & P. Winkielman (Eds.), *Social Neuroscience: Integrating Biological and Psychological Explanations of Social Behavior* (pp. 227–245). New York: The Guilford Press.

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., Packer, D. J., Kesek, A., & Van Bavel, J. J. (2009). Implicit measures of attitudes: A physiological approach. In R. E. Petty, R. H. Fazio & P. Brinol (Eds.), *Attitudes: Insights From the new Implicit Measures* (pp. 485–512). New York: Psychology Press.

Cunningham, W. A., Raye, C. L., & Johnson, M. K. (2005). Neural correlates of evaluation associated with promotion and prevention regulatory focus. *Cognitive, Affective & Behavioral Neuroscience*, **5**, 202–211.

Cunningham, W. A., & Van Bavel, J. J. (2009a). A neural analysis of intergroup perception and evaluation. In G. G. Berntson & J. T. Cacioppo (Eds.), *Handbook of Neuroscience for the Behavioral Sciences* (pp. 975–984). Hoboken, NJ: Wiley.

Cunningham, W. A., & Van Bavel, J. J. (2009b). Varieties of emotional experience: Differences in object or computation? *Emotion Review*, **1**, 56–57.

Cunningham, W. A., Van Bavel, J. J., Arbuckle, N. L., Packer, D. J., & Waggoner, A. S. (2012). *Automatic Social Perception is Flexible: Approach and Avoidance Motivational States Shape P100 Responses to Other-Race Faces*. Columbus: Department of Psychology, Ohio State University. Unpublished manuscript.

Cunningham, W. A., Van Bavel, J. J., & Johnsen, I. R. (2008). Affective flexibility: Evaluative processing goals shape amygdala activity. *Psychological Science*, **19**, 152–160.

Cunningham, W. A., & Zelazo, P. D. (2007). Attitudes and evaluations: A social cognitive neuroscience perspective. *Trends in Cognitive Sciences*, **11**, 97–104.

Cunningham, W. A., Zelazo, P. D., Packer, D. J., & Van Bavel, J. J. (2007). The iterative reprocessing model: A multi-level framework for attitudes and evaluation. *Social Cognition*, **25**, 736–760.

Descartes, R. (1641/1984). Meditations on First Philosophy. In J. Cottingham, R. Stoothoff & D. Murdoch Trans. *The Philosophical Writings of René Descartes* (Vol. 2, pp. 1–62). Cambridge: Cambridge University Press.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, **56**, 5–18.

Douglas, R. J., & Martin, K. A. C. (2004). Neuronal circuits of the neocortex. *Annual Review of Neuroscience*, **27**, 419–451.

Dovidio, J. F., Evans, N., & Tyler, R. B. (1986). Racial stereotypes: The contents of their cognitive representations. *Journal of Experimental Social Psychology*, **22**, 22–37.

Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology*, **82**, 62–68.

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

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

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

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

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality & Social Psychology*, **50**, 229–238.

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

Ferguson, M. J., & Wojnowicz, M. T. (2011). The when and how of evaluative readiness: A social cognitive neuroscience perspective. *Social & Personality Psychology Compass*, **5**, 1018–1038.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 23, pp. 1–74). New York: Academic Press.

Forbes, C. E., Cox, C. L., Schmader, T., & Ryan, L. (forthcoming). Negative stereotype activation alters interaction between neural correlates of arousal, inhibition and cognitive control. *Social Cognitive and Affective Neuroscience*.

Frank, M. J., Cohen, M. X., & Sanfey, A. G. (2009). Multiple systems in decision making: A neurocomputational perspective. *Current Directions in Psychological Science*, **18**, 73–77.

Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review*, **118**, 247–279.

Freud, S. (1930). *Civilization and its Discontents*. London: Hogarth.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B*, **360**, 815–836.

Gaertner, S. L., & Dovidio, J. F. (1986). The aversive form of racism. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, Discrimination, and Racism* (pp. 61–89). Orlando, FL: Academic Press.

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.

Gilbert, D. T. (1999). What the mind's not. In S. Chaiken & Y. Trope (Eds.), *Dual Process Theories in Social Psychology* (pp. 3–11). New York: Guilford.

Gilbert, D. T., Pelham, B. W., & Krull, D. S. (1988). On cognitive busyness: When person perceivers meet persons perceived. *Journal of Personality and Social Psychology*, **55**, 685–694.

Gollwitzer, P. M., & Schaal, B. (1998). Metacognition in action: The importance of implementation intentions. *Personality and Social Psychology Review*, **2**, 124–136.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, **293**, 2105–2108.

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

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, **108**, 814–834.

Hofmann, W., Deutsch, R., Banaji, M. R., & Lancaster, K. (2010). Cooling the heat of temptation: Mental self-control and the automatic evaluation of tempting stimuli. *European Journal of Social Psychology*, **40**, 17–25.

Hofmann, W., Friese, M., & Strack, F. (2009). Impulse and self-control from a dual-systems perspective. *Perspectives on Psychological Science*, **4**, 162–176.

Houston, D. A., & Fazio, R. H. (1989). Biased processing as a function of attitude accessibility: Making objective judgments subjectively. *Social Cognition*, **7**, 51–66.

Jellison, W. A., McConnell, A. R., & Gabriel, S. (2004). Implicit and explicit measures of sexual orientation attitudes: Ingroup preferences and related behaviors and beliefs among gay and straight men. *Personality and Social Psychology Bulletin*, **30**, 629–642.

Johnson, M. K., & Hirst, W. (1993). MEM: Memory subsystems as processes. In A. F. Collins, S. E. Gathercole, M. A. Conway & P. E. Morris (Eds.), *Theories of Memory* (pp. 241–286). East Sussex, England: Erlbaum.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, **58**, 697–720.

Katz, D., & Stotland, E. (1959). A preliminary statement to a theory of attitude structure and change. *Psychology: A study of a science*, **3**, 423–475.

Kim, H., Somerville, L. H., Johnstone, T., Polis, S., Alexander, A. L., Shin, L. M. et al. (2004). Contextual modulation of fMRI responsivity to surprised faces. *Journal of Cognitive Neuroscience*, **16**, 1730–1745.

Kirkland, T., & Cunningham, W. A. (2011). Neural basis of affect and emotion. *WIRES: Cognitive Science*, **2**, 656–665.

Kirkland, T., & Cunningham, W. A. (forthcoming). Mapping emotions through time: How affective trajectories inform the language of emotion. *Emotion*.

Kouider, S., de Gardelle, V., Sackur, J., & Dupoux, E. (2010). How rich is consciousness? The partial awareness hypothesis. *Trends in Cognitive Sciences*, **14**, 301–307.

Kruglanski, A. W., & Thompson, E. P. (1999). Persuasion by a single route: A view from the unimodel. *Psychological Inquiry*, **10**, 83–109.

Kunda, Z. (1999). Parallel processing of stereotypes and behaviors. In S. Chaiken & Y. Trope (Eds.), *Dual-Process Theories in Social Psychology* (Vol. xiii, 657, pp. 314–322). New York: The Guilford Press.

Kunda, Z., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review*, **103**, 284–308.

Lieberman, M. D. (2003). Reflective and reflexive judgment processes: A social cognitive neuroscience approach. In J. P. Forgas, K. D. Williams & W. von Hippel (Eds.), *Social Judgments: Explicit and Implicit Processes* (pp. 44–67). New York: Cambridge University Press.

Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature*, **453**, 869–878.

Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual Review of Psychology*, **51**, 93–120.

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.

Magen, E. G. J. J. (2007). Harnessing the need for immediate gratification: Cognition reconstrual modulates the reward value of temptations. *Emotion*, **7**, 415–428.

McClelland, J. L., & Rumelhart, D. E. (Eds.). (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. (Vol. II). Cambridge, MA: MIT 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.

Mendoza, S. A., Gollwitzer, P. M., & Amodio, D. M. (2010). Reducing the expression of implicit stereotypes: Reflexive control through implementation intentions. *Personality and Social Psychology Bulletin*, **36**, 512–523.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, **24**, 167–202.

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 & Social Psychology*, **77**, 167–184.

O'Reilly, R. C., & Munakata, Y. (2000). *Computational Explorations in Cognitive Neuroscience*. Cambridge, MA: The MIT Press.

Park, G. H., Van Bavel, J. J., Vasey, M. W., & Thayer, J. F. (forthcoming). Cardiac vagal tone predicts inhibited attention to fearful faces. *Emotion*.

Perugini, M. (2005). Predictive models of implicit and explicit attitudes. *British Journal of Social Psychology*, **44**, 29–45.

Pettigrew, T. F., & Meertens, R. W. (1995). Subtle and blatant prejudice in western Europe. *European Journal of Social Psychology*, **25**, 57–75.

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

Petty, R. E., & Cacioppo, J. T. (1986). The Elaboration Likelihood Model of persuasion. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 19, pp. 123–205). New York: Academic Press.

Petty, R. E., Fazio, R. H., & Briñol, P. (Eds. 2009). *Attitudes: Insights From the new Implicit Measures*. New York, NY: Psychology Press.

Rosenberg, M. J., & Hovland, C. I. (1960). Cognitive, affective, and behavioral components of attitudes. Attitude organization and change: An analysis of consistency among attitude components. In M. J. Rosenberg, C. I. Hovland, W. J. McGuire, R. P. Abelson & J. W. Brehm (Eds.), *Attitude Organization and Change. An Analysis of Consistency Among Attitude Components* (pp. 1–14). New Haven, CT: Yale University Press.

Rumelhart, D. E., & McClelland, J. L. (Eds.). (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (Vol. I). Cambridge, MA: MIT Press.

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

Scherer, K. R. (2009). The dynamic architecture of emotion: Evidence for the component process model. *Cognition and Emotion*, **23**, 1307–1351.

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* (pp. 436–457). Balden, MA: Blackwell Publishers.

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

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

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

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

Stewart, B. D., & Payne, B. K. (2008). Bringing automatic stereotyping under control: Implementation intentions as efficient means of thought control. *Personality and Social Psychology Bulletin*, **34**, 1332–1345.

Storbeck, J., & Clore, G. L. (2007). On the interdependence of cognition and emotion. *Cognition & Emotion*, **21**, 1212–1237.

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

Sun, R. (2002). *Duality of the Mind*. Mahwah, NJ: Lawrence Erlbaum Associates.

Thurstone, L. L. (1928). Attitudes can be measured. *American Journal of Sociology*, **33**, 529–555.

Van Bavel, J. J., & Cunningham, W. A. (2009a). Self-categorization with a novel mixed-race group moderates automatic social and racial biases. *Personality and Social Psychology Bulletin*, **35**, 321–335.

Van Bavel, J. J., & Cunningham, W. A. (2009b). A social cognitive neuroscience approach to intergroup perception and evaluation. In W. P. Banks (Ed.), *Encyclopedia of Consciousness* (pp. 379–388). New York: Academic Press.

Van Bavel, J. J., & Cunningham, W. A. (2011). A social neuroscience approach to self and social categorisation: A new look at an old issue. *European Review of Social Psychology*, **21**, 237–284.

Van Bavel, J. J., & Cunningham, W. A. (2012). *The Social Dynamics of Person Memory: The Significance and Social Role of One's Social Identity Shape Attention and Memory*. Department of Psychology, New York University. Unpublished manuscript.

Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2008). The neural substrates of in-group bias: A functional magnetic resonance imaging investigation. *Psychological Science*, **19**, 1131–1139.

Van Bavel, J. J., Packer, D. J., Johnsen, I. R., & Cunningham, W. A. (2012). *Moral Evaluations: Fast, Extreme, and Universally Prescriptive*. Department of Psychology, New York University. Unpublished manuscript.

Van Bavel, J. J., Swencionis, J. K., O'Connor, R. C., & Cunningham, W. A. (2012). Motivated social memory: Belonging needs moderate the own-group bias in face recognition. *Journal of Experimental Social Psychology*, **48**, 707–713.

Wegener, D. T., & Petty, R. E. (1997). The flexible correction model: The role of naive theories of bias in bias correction. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 29, pp. 141–208). San Diego: Academic Press.

Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M., & Jenike, M. A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience*, **18**, 411–418.

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

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

Xiao, Y. J., & Van Bavel, J. J. (forthcoming). See your friends close, and your enemies closer: Social identity and identity threat shape the representation of physical distance. *Personality and Social Psychology Bulletin*.

