Title

Two of what? A conceptual analysis of dual-process theories

Authors

Gawronski, B

Sherman, JW

Trope, Y

Publication Date

2023-12-10

Peer reviewed

In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.). (2014). Dual-process theories of the social mind (pp. 3-19). New York, NY: Guilford Press.

CHAPTER 1

Two of What?
A Conceptual Analysis of Dual-Process Theories

Bertram Gawronski, Jeffrey W. Sherman, and Yaacov Trope

The emergence of dual-process theories is probably one of the most significant theoretical developments in the history of social psychology. The overarching assumption of dual-process theorizing is that the mental processes underlying social phenomena can be divided into two distinct categories depending on whether they operate in an automatic or nonautomatic fashion.[1] Over the past decades, dual-process theories have made important contributions in virtually all areas of social psychology, and this volume showcases the most recent developments since Chaiken and Trope's (1999) seminal volume, *Dual-Process Theories in Social Psychology*.

Despite the ubiquity of dual-process theories in social psychology, their conceptual foundations have also been the subject of ongoing debates. In this chapter, we provide a metatheoretical analysis that aims at clarifying the explanatory function of dual-process theories, the conceptual nature of their underlying dualities, and structural features that characterize different types of dual-process theories. Expanding on this analysis, we discuss the criticism that dual-process theories are unfalsifiable and whether the realm of mental processes can indeed be divided into two distinct categories. Our conclusion is that dual-process theories have provided vital insights into the mental underpinnings of social phenomena. However, their explanatory and predictive value depends on (1) a clear distinction between operating principles and operating conditions, (2) conceptual rigor in the definition of the proposed dualities, (3) precise formulations of empirical hypotheses about covariations between processing dualities, and (4) clearly specified links between the hypothesized mental processes and the causal relations between stimuli and behavior they are supposed to explain. Expanding on this analysis, we conclude with a brief outlook on emerging themes and future directions in dual-process theorizing.

EXPLANATORY FUNCTION

A useful framework to clarify the explanatory function of dual-process theories is Marr's (1982) distinction among three levels of analysis in psychological research: the computational level, the algorithmic level, and the implementational level. According to Marr, research at the *computational level* is concerned with identifying relations between inputs (i.e., stimuli and their broader contexts) and outputs (i.e., judgments and behavior). The overarching goal of research at the computational level is to identify which types of inputs produce

which kinds of outputs under which contextual conditions. For example, a large body of research on behavioral priming can be described as computational, in that it focuses on the particular behaviors that are elicited by exposure to various kinds of prime stimuli (for a review, see Bargh, 2006). Research of this kind differs from research at the *algorithmic level*, which is concerned with the mechanisms that translate inputs into outputs. This level of analysis resonates with the goal of social-cognitive research, which aims at identifying the mental processes and representations underlying social behavior. For example, expanding on the identification of input–output relations in studies on behavioral priming, a considerable body of research investigated the mental mechanisms that mediate the effects of prime exposure on overt behavior, including motivational (e.g., Cesario, Plaks, & Higgins, 2006), misattribution (e.g., Loersch & Payne, 2011) and self-related (e.g., Wheeler, DeMarree, & Petty, 2007) processes. Finally, research at the *implementational level* is concerned with the physical systems that implement the mechanisms identified at the algorithmic level. In social psychology, this approach is prominently reflected in the emerging field of social neuroscience, which is concerned with the neural underpinnings of social judgments and social behavior (Cacioppo, Berntson, Sheridan, & McClintock, 2000; Ochsner & Lieberman, 2001). For example, expanding on mental process theories of prime-to-behavior effects (e.g., Cesario et al., 2006; Loersch & Payne, 2011; Wheeler et al., 2007), research at the implementational level may investigate the neural underpinnings of the mechanisms that mediate observed relations between certain kinds of primes and overt behavior.

Dual-process theories are located at Marr's (1982) algorithmic level of analysis, in the sense that they identify mental mechanisms that translate inputs into outputs. A central feature of dual-process theories is that they postulate two qualitatively distinct (sets of) mental processes that mediate between inputs and outputs. Some dual-process theories go beyond the algorithmic level by including assumptions about the neural substrates that implement the hypothesized processes (e.g., Lieberman, Gaunt, Gilbert, & Trope, 2002). Yet even these theories have their theoretical core at the algorithmic level, in that the identified substrates are directly linked to two (sets of) mechanisms that are claimed to translate inputs into outputs.

Although the proposed positioning of dual-process theories at Marr's (1982) algorithmic level may seem rather trivial, it helps to clarify the explanatory function of dual-process theories by specifying the empirical phenomena that dual-process theories aim to explain (*explanandum*) and the theoretical assumptions that are proposed to explain these phenomena (*explanans*). From an epistemological point of view, one could argue that research at the computational level aims to explain observed outputs by relating them to inputs that cause these outputs. Using the previous example of behavioral priming, exposure to a particular stimulus may serve as an explanation for an observed behavioral response to the extent that the stimulus can be said to cause the behavioral response. In other words, the observed behavior represents the phenomenon that needs to be explained, and exposure to the prime stimulus serves as the event that is supposed to explain the behavior (*causal explanation*). However, stating that exposure to the prime explains the behavioral response does not say anything about *how* the prime caused the observed behavior. This question is central in research at the algorithmic level, in which the causal relation between prime exposure and behavior represents a phenomenon that is in need of further explanation (De Houwer, 2011). Research at the algorithmic level provides an answer to this question by identifying the mental mechanisms that mediate the link between prime exposure and overt behavior (*mechanistic explanation*). In this sense, dual-process theories offer explanations of observed input–output relations by specifying the mental mechanisms that translate inputs into outputs. As we outline in the following sections, this conceptualization has important implications for dual-process theorizing in social psychology.

OPERATING PRINCIPLES VERSUS OPERATING CONDITIONS

The first important insight that can be gained from relating dual-process theories

to Marr's (1982) algorithmic level of analysis is that it resolves the common conflation of operating principles and operating conditions. Whereas the concept of *operating principles* refers to the mental mechanisms that translate inputs into outputs, the concept of *operating conditions* refers to the conditions under which a given process operates (Gawronski & Bodenhausen, 2009). A central characteristic of dual-process theories is that they are concerned with the question of whether the mental processes underlying social behavior operate in an automatic or nonautomatic fashion. This emphasis has sometimes led to the misunderstanding that labeling a process as automatic or nonautomatic is sufficient to characterize the intrinsic nature of that process. However, stating that a process operates in an automatic or nonautomatic fashion simply specifies *when* the process is assumed to operate; it does not specify *how* the process translates inputs into outputs. That is, characterizing a mental process as automatic versus nonautomatic specifies whether the process does or does not operate (1) when there is no conscious awareness, (2) when there is no goal to start the process, (3) when cognitive resources are reduced, and (4) when there is a goal to alter or stop the process (Bargh, 1994). Thus, although research investigating the unawareness, unintentionality, efficiency, and uncontrollability of a given process is essential to understand the boundary conditions of observed input–output relations (e.g., does a given input lead to a particular output when cognitive resources are reduced?), such research by itself does not address the question of *how* the human mind translates certain inputs into particular outputs.

In terms of Marr's (1982) framework, research on operating conditions is located at the computational level, in that it aims at identifying which types of inputs produce which kinds of outputs under which contextual conditions (e.g., does input *X* produce output *Y* when participants simultaneously perform a secondary task?). Although such research is essential to the concern with automaticity, the explanatory goal of dual-process theories goes beyond the computational level, in that they aim at specifying the mental mechanisms that translate inputs into outputs (e.g., what are the mental operations that translate input *X* into output *Y*?). From the perspective of Marr's algorithmic level, characterizations of a given process as automatic versus nonautomatic are not sufficient as a conceptual foundation of dual-process theories, because they fail to specify the nature of the processes that translate inputs into outputs. An illustrative example is the conscious–unconscious duality, which has been used in a manner suggesting that it refers to two qualitatively distinct mental processes (e.g., Baumeister, Masicampo, & Vohs, 2011; Dijksterhuis & Nordgren, 2006). Of course, it is possible that conscious and unconscious processing of input stimuli produces different behavioral outputs via two qualitatively distinct mechanisms. However, it is also possible that conscious and unconscious processes operate on the basis of the same mental structures involving the same mental operations (Huang & Bargh, in press. Simply stating that a mental process is conscious or unconscious does not specify *how* this process translates inputs into outputs, nor does it specify whether conscious and unconscious processing involve the same or different mental structures and operations. Thus, although dual-process theorizing is often equated with research on awareness, intentionality, efficiency, and controllability (i.e., operating conditions), its explanatory goal at the algorithmic level requires clear specifications of the mental mechanisms that translate inputs into outputs (i.e., operating principles).[2]

Another important caveat in this context is that different features of automatic processing do not necessarily co-occur (Bargh, 1994; Moors & De Houwer, 2006). Counter to early *dual-mode conceptualizations* assuming an all-or-none relation between different features of automaticity, the available evidence indicates that there is virtually no process that is characterized by all four features of automaticity. Instead, most processes studied within social psychology involve combinations of selected features, making them automatic in one sense and nonautomatic in another (Bargh, 1992). For example, a process may be unintentional and controllable, intentional and efficient, unintentional and resource-dependent, conscious and uncontrollable, unconscious and resource-dependent, controllable and

resource-independent, and so forth (e.g., Fujita, 2011; Gawronski & Bodenhausen, 2011; Hassin, Bargh, Engell, & McCulloch, 2009). This insight has inspired *disjunctive conceptualizations* of automaticity, according to which a process can be characterized as automatic if it meets at least one of the four criteria of automaticity. According to this view, a process can be described as automatic if it is (1) unconscious, (2) unintentional, (3) efficient, or (4) uncontrollable.

Although disjunctive treatments of automaticity are rather common in social psychology, they involve a number of problems (Gawronski & Creighton, 2013). First, if the presence of a single feature is sufficient to call a process automatic, it is possible that a given process has to be described as automatic and nonautomatic at the same time. Needless to say, such a description can cause considerable confusion if it does not specify in which particular sense the process is automatic and in which sense it is nonautomatic. Second, generic use of the term *automatic* to describe any of the four operating conditions can lead to confusion about conceptually distinct findings that are described with the same term. For example, a given Process A may be described as automatic because it is elicited unintentionally, whereas another Process B may be described as automatic because it does not require a large amount of cognitive resources. Yet despite their common description as automatic, the two processes may be fundamentally different, for example, if Process A is resource dependent and Process B is intentional.

Based on these considerations, several theorists recommended that researchers should be more precise in their use of terminology by describing each feature of automaticity with its proper label (i.e., unconscious, unintentional, efficient, or uncontrollable; Bargh, 1994; Moors & De Houwer, 2006). Importantly, such *decompositional conceptualizations* prohibit simple binary categorizations of mental processes as automatic or nonautomatic. Because the four features of automaticity do not necessarily co-occur, a decompositional classification schema involves 16 rather than two categories of potential operating conditions. On the basis of this conclusion, the assumption that the mental processes underlying social phenomena can be divided into two internally coherent categories depending on whether they operate in an automatic or nonautomatic fashion should be treated with caution. Of course, dual-process theories share the explanatory goal to identify the mental mechanisms that translate inputs into outputs and the assumption that input–output relations are mediated by two qualitatively distinct (sets of) processes that operate under different processing conditions. However, different theories emphasize different features of automaticity, and the nature of the proposed mechanisms is not necessarily equivalent.

TYPES OF DUAL-PROCESS THEORIES

Despite the concerns about generic treatments, the term *dual-process theory* is sometimes used in the singular to refer to an overarching theoretical idea that could be boiled down to a single theory (see Evans & Frankish, 2009). However, as the variety of contributions to this volume illustrates, there is not really a unifying "essence" that captures what might be regarded as the prototype of dual-process theories. Instead, there are important nuances that are easy to miss if different theories are treated as minor variations of the same prototype theory. Nevertheless, it seems possible to classify dual-process theories in terms of a few general characteristics, such as the phenomena they aim to explain (*explanandum*), the theoretical constructs they propose to explain these phenomena (*explanans*), and their mathematical formalization.

Explanandum

When dual-process theories started to emerge in the 1980s, their focus was mainly domain-specific, in that they aimed at explaining phenomena in particular areas of inquiry. Although some of these theories were based on general processing principles from cognitive psychology (e.g., Chaiken, 1987; Trope, 1986), their applications were specific to particular content domains within social psychology. Prominent examples include dual-process theories of persuasion (e.g., Chaiken, 1987; Petty & Cacioppo, 1986), attitude–behavior relations (e.g., Fazio, 1990), dispositional attri-

bution (e.g., Gilbert, 1989; Trope, 1986), prejudice and stereotyping (e.g., Devine, 1989), and impression formation (e.g., Brewer, 1988; Fiske & Neuberg, 1990). A shared feature of these theories is that they explain particular instances of input–output relations (*explanandum*) by postulating two (sets of) mechanisms by which the human mind translates inputs into outputs (*explanans*). For example, dual-process theories of persuasion aim at explaining the effects of different features of persuasive messages on attitudes. Their overarching goal is to understand the mental mechanisms by which different kinds of inputs (e.g., argument strength, source characteristics) are translated into outputs (i.e., attitude change). Similarly, dual-process theories of dispositional attribution aim at explaining the effects of different kinds of information about potential causes of a person's behavior on perceivers' trait impressions. Again, the overarching goal is to understand the mental mechanisms by which different kinds of inputs (e.g., behavioral information, situational information) are translated into outputs (i.e., dispositional attributions). In general, domain-specific dual-process theories differ in terms of the phenomena they aim to explain, in that they focus on input–output relations in different content areas (*explanandum*). Although the proposed explanations are specific to the phenomena of interest, their shared assumption is that different kinds of input–output relations are mediated by two (sets of) qualitatively distinct processes that operate under different conditions (*explanans*).

With the beginning of the new millennium, the focus of dual-process theorizing shifted toward the development of integrative theories that aim at identifying general principles that are independent of particular content domains. A seminal contribution in this regard was Smith and DeCoster's (2000) conceptual integration of various domain-specific theories within a single dual-process framework. The central argument of their integrative account is that the multiple dualisms proposed by domain-specific theories reflect the operation of two basic processes that characterize any kind of human thought irrespective of its content: associative versus rule-based processes (cf. Sloman, 1996). This distinction has become the common denominator of various generalized dual-process theories, including theories that distinguish between reflective and impulsive processing (Strack & Deutsch, 2004), reflective and reflexive processing (Lieberman et al., 2002), and System 1 versus System 2 processing (Kahneman, 2003; Stanovich & West, 2000). Deviating from the initial focus on domain-specific phenomena, generalized dual-process theories are concerned with the basic architecture of information processing (Carruthers, 2009; Samuels, 2009). Thus, generalized dual-process theories differ from domain-specific dual-process theories in terms of the breadth of their *explanandum,* such that the latter aim at explaining input–output relations in particular content areas, whereas the former aim at providing a general account of the workings of the human mind.

Explanans

Beyond differences in terms of the phenomena that dual-process theories aim to explain, there are nontrivial differences in the mental constructs they propose to explain the phenomena of interest. Whereas some theories emphasize functionally distinct mental processes (*dual-process theories*), other theories attribute different behavioral outcomes to functionally distinct mental representations (*dual-representation theories*). Moreover, some theories include assumptions about both process and representation, assuming that different outcomes are the product of two functionally distinct processing systems (*dual-system theories*).

A defining feature of *dual-process theories* is that they explain different kinds of input–output relations on the basis of two (sets of) mental mechanisms. For example, Fiske and Neuberg's (1990) continuum model of impression formation distinguishes between categorization processes and piecemeal integration of individuating attributes. Similarly, Trope's (1986) two-stage model of dispositional attribution distinguishes between perceptual identification of trait-relevant cues and subsequent inference of dispositions. Many of these theories include empirical hypotheses about systematic covariations between operating principles and operating conditions, in that the proposed mechanisms are assumed

to operate under different conditions. For example, Fiske and Neuberg's (1990) continuum model assumes that the initial categorization of a target person requires little amounts of cognitive resources, whereas the piecemeal integration of individuating attributes is assumed to be resource-dependent. Similarly, Trope's (1986) two-stage model assumes that the perceptual identification of trait-relevant cues occurs unintentionally, whereas dispositional inference is assumed to be an intentional process (for a review, see Gawronski & Creighton, 2013). Thus, in addition to differences in the particular phenomena they aim to explain, dual-process theories can be characterized by their assumptions about the nature of the two processes they propose to explain these phenomena (i.e., operating principles) and the particular conditions under which each of the two processes is assumed to operate (i.e., operating conditions).

Deviating from the emphasis on processes, *dual-representation theories* attribute different behavioral outcomes to distinct mental representations. A useful example to illustrate the difference between dual-process and dual-representation theories is research comparing evaluative responses on traditional self-report measures to responses on performance-based tasks, such as the Implicit Association Test (Greenwald, McGhee, & Schwartz, 1998) or the evaluative priming task (Fazio, Jackson, Dunton, & Williams, 1995). A common finding in this area is that evaluative responses on the two kinds of measures show various dissociations, including different antecedents, different consequences, and discrepant evaluations of the same object (for reviews, see Friese, Hofmann, & Schmitt, 2008; Gawronski & Bodenhausen, 2006). Such dissociations have led some researchers to conclude that self-report measures reflect conscious, "explicit" attitudes, whereas performance-based tasks reflect unconscious, "implicit" attitudes (e.g., Greenwald & Banaji, 1995; Rydell & McConnell, 2006; see also Wilson, Lindsey, & Schooler, 2000). Yet counter to such dual-representation accounts, dual-process accounts assume that the two kinds of measures capture the same attitudinal representation, the primary difference being the reduced opportunity to control responses on performance-based measures (e.g., Fazio, 2007). Whereas dual-representation accounts are based on presumed parallels to the distinction between explicit and implicit memory, dual-process accounts emphasize the unintentional activation of attitudes and their reduced impact on overt responses when people have the motivation and opportunity to deliberate about specific attributes of the attitude object (for a more detailed discussion, see Payne & Gawronski, 2010).

A third category of theories explains different behavioral outcomes by the operation of two distinct processing systems. The shared assumption of such *dual-system theories* is that multiple psychological dualities are systematically correlated, thereby constituting two functionally distinct mental systems. Although dual-system theories differ in their assumptions about which dualities represent core features of the proposed systems, the hypothesized correlations between dichotomous characteristics are often depicted in lists of features that describe one of the two systems as *associative*, *automatic*, *slow-learning*, *experiential*, *affective*, *parallel*, and *holistic*, and the other one as *rule-based*, *nonautomatic*, *fast-learning*, *rational*, *cognitive*, *sequential*, and *analytic* (e.g., Epstein, 1994; Kahneman, 2003; Sloman, 1996; Smith & DeCoster, 2000). Although some dual-system theories limit their focus to particular content domains (e.g., Rydell & McConnell, 2006), the majority falls into the category of generalized theories that aim at providing a domain-independent account of the workings of the human mind (e.g., Smith & DeCoster, 2000; Strack & Deutsch, 2004).

Formalization

The different types of dual-process theories discussed so far share the feature that they rely only on the informal logic of verbally formulated propositions rather than mathematical formalization (for a notable exception, see Trope, 1986). Over the past decade, some researchers have started to develop formalized theories in which the mechanisms that are assumed to mediate input–output relations are specified in mathematical terms. The most prominent example is Jacoby's (1991) process dissociation (PD) model (for a review, see Payne & Bishara, 2009). Other, more complex theories rely on

multinomial modeling (e.g., Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005; Krieglmeyer & Sherman, 2012; Meissner & Rothermund, 2013; Nadarevic & Erdfelder, 2010; Payne, Hall, Cameron, & Bishara, 2010; Stahl & Degner, 2007) and applications of random-walk and diffusion modeling (e.g., Klauer, Voss, Schmitz, & Teige-Mocigemba, 2007). Regardless of the specifics of their mathematical underpinnings, formalized dual-process theories have at least three advantages over nonformalized theories.

First, formalized theories allow researchers to quantify the contribution of multiple distinct processes to a given behavioral outcome, which is not possible on the basis of verbally formulated theories. The significance of this characteristic is reflected in the principle of *equifinality,* which refers to cases in which combinations of different processes produce the same behavioral outcome. For example, in research on self-regulation, two people may show the same behavioral response when (1) the initial impulse and inhibitory control are weak or (2) the initial impulse and inhibitory control are strong (Sherman et al., 2008). Formalized theories are able to capture such complex interplays by providing quantitative estimates for each of the proposed processes.

Second, formalized theories have the advantage that their logical coherence can be tested by means of their mathematical constraints (Klauer, in press). For many verbally formulated theories it can be very difficult to identify logical inconsistencies between their core assumptions, particularly when they involve a large number of theoretical claims. To illustrate this problem, consider a theory that includes a set of N propositions. Logically, any subset of $N - 1$ of these propositions could be internally consistent even if the entire set N is inconsistent. This possibility implies that an exhaustive consistency assessment of N propositions requires scrutiny of 2^N cases. The resulting capacity problem is illustrated by the fact that, even if each case could be examined in a millionth of a second, an exhaustive consistency check for a theory including 100 propositions would take longer than the universe has existed (see Johnson-Laird, 2012). Formalized theories avoid this problem, because their underlying mathematical constraints help to identify inconsistent assumptions at various stages of the research process, including the initial specification of a theory, the generation of predicted data patterns, the estimation of model parameters from observed data, and the assessment of fit between predicted and observed data (Klauer, in press).

Third, formalized theories have the advantage that they directly link the proposed processes to relations between inputs and outputs. In a strict sense, psychological measures do not assess mental processes or mental representations, but their behavioral outcomes (De Houwer, 2011). Although it is rather common to treat behavioral outcomes as "proxies" for mental constructs, such treatments involve the logical fallacy of *affirming the consequent*, also known as *reverse inference* (Gawronski & Bodenhausen, in press-a). One possibility to avoid this problem is to distinguish clearly between the behavioral outputs captured by psychological measures and the mental constructs that are proposed to explain input–output relations. The validity of dual-process theories (like any other mental process theory) can then be tested by deriving predictions about input–output relations and the particular conditions under which they should emerge (Gawronski & Bodenhausen, in press-b). Yet the derivation of such predictions requires "bridging" assumptions that link the proposed mental constructs to inputs and outputs (McGrath, 1981). Domain-specific dual-process theories typically include such assumptions, but these theories have been criticized for referring to particular contents (e.g., argument strength vs. source characteristics) in linking phenomenon-relevant inputs to the proposed processes (e.g., central/systematic vs. peripheral/heuristic processing). This strategy seems problematic to the extent that the processing of different information involves the same mental operations irrespective of its content (see Kruglanski, Erb, Pierro, Mannetti, & Chun, 2006; Sherman, 2006). Generalized dual-process theories avoid this problem by postulating two (sets of) content-independent processes. However, many of these theories focus primarily on the internal architecture of the human mind, without specifying links between the proposed processes and the relevant inputs. Formalized dual-process theories avoid both problems by (1) proposing

general processes that are independent of particular contents and (2) directly linking the proposed processes to observed input–output relations in their mathematical formulations.

An important aspect of formalized dual-process theories is the relation between their processing parameters and the concept of automaticity. For example, the processing parameters of Jacoby's (1991) PD model have sometimes been interpreted as capturing automatic and nonautomatic processes by virtue of its underlying mathematical structure. This interpretation is reflected in the common depiction of the two parameter estimates as *automatic* (using the acronym A) and *controlled* (using the acronym C). However, such depictions conflate the difference between operating principles and operating conditions. Although formalized theories provide clear specifications of *how* inputs are translated into outputs, their mathematical underpinnings do not have any implications for *when* the proposed processes operate (i.e., when there is no conscious awareness; when cognitive resources are reduced; when there is no goal to start the process; when there is a goal to alter or stop the process). Questions about the conditions under which a given process operates cannot be addressed by mathematical formalizations but have to be answered on the basis of empirical data. For example, whether the particular processes captured by the two PD parameters depend on the availability of cognitive resources has to be investigated by testing effects of cognitive load or time pressure on the two parameters, and there is nothing in the underlying PD formulas that would guarantee one or the other outcome. The same is true for all other features of automaticity, including unawareness, unintentionality, and uncontrollability (for a more detailed discussion, see Gawronski & Creighton, 2013).

COVARIATION OF DUALITIES

Many dual-process theories hypothesize systematic covariations between two or more psychological dualities (Samuels, 2009). At the most basic level, these covariations involve empirical relations between functionally distinct processes and the conditions under which these processes operate. For example, a common assumption of domain-specific dual-process theories is that the processes underlying the effects of certain kinds of inputs are intentional and resource-dependent, whereas those underlying the effects of other kinds of inputs are unintentional and resource-independent. Examples include processing constraints on the effects of central versus peripheral cues in dual-process theories of persuasion (e.g., Petty & Cacioppo, 1986), behavioral versus situational information in dual-process theories of dispositional attribution (e.g., Gilbert, 1989), and category versus individuating information in dual-process theories of impression formation (e.g., Fiske & Neuberg, 1990). Similarly, dual-representation theories often attribute a particular feature of automatic processing (e.g., unconscious) to one representation and the opposite feature (e.g., conscious) to the other representation (e.g., Greenwald & Banaji, 1995). Yet the idea of covariation between dualities is most explicit in dual-system theories, which propose systematic relations between multiple distinct dualities (e.g., Epstein, 1994; Kahneman, 2003; Lieberman et al., 2002; Sloman, 1996; Smith & DeCoster, 2000; Stanovich & West, 2000; Strack & Deutsch, 2004).

Although assumptions about covariations between psychological dualities are very common, their epistemic value depends on two conceptual issues (see Moors, Chapter 2, this volume). First, it is important that the two categories underlying a given duality are characterized by a clear demarcation, so that they constitute nonoverlapping categories. For example, the distinction between categorical and individuating information in dual-process theories of impression formation has been criticized for being ambiguous as to whether a given target characteristic should be regarded as categorical information (e.g., man vs. woman; adolescent vs. elderly) or as an individuating attribute (e.g., male vs. female; young vs. old) (see Kunda & Thagard, 1996). Second, it is important to distinguish between conceptual definitions and empirical hypotheses when proposing systematic relations between two or more dualities. For example, to the extent that associative processes are *defined* as unconscious and propositional processes

are *defined* as conscious (e.g., Mitchell, De Houwer, & Lovibond, 2009), the relation between the two dualities would be purely semantic rather than empirical (cf. Smedslund, 2000; Wallach & Wallach, 1994). In this case, it would make no sense to test empirically whether propositional processes require conscious awareness, because that is how propositional processes are defined in the first place. Any process that is conscious would be propositional by definition. Yet, if the associative–propositional dualism is defined without reference to the unconscious–conscious dualism (e.g., Gawronski & Bodenhausen, 2011), assumptions about systematic covariations between the two dualities could be subject to empirical tests to the extent that (1) the associative–propositional distinction specifies particular input–output relations that can be expected on the basis of the two processes (operating principles), and (2) the unconscious–conscious distinction specifies the conditions under which these input–output relations should emerge (operating conditions). Thus, a basic requirement for conceptually sound dual-process theorizing is that the categories of conceptually distinct dualities are defined in a manner that avoids semantic overlap between these dualities (Moors, Chapter 2, this volume). Moreover, although conceptual definitions and empirical hypotheses are often conflated in claims about covariations between multiple dualities, it is important that the categories constituting a given duality are defined in a manner that allows unambiguous demarcations between these categories.

ARE DUAL-PROCESS THEORIES UNFALSIFIABLE?

The quest to distinguish clearly between conceptual definitions and empirical hypotheses is particularly important in the context of criticism that dual-process theories are unfalsifiable (e.g., Keren & Schul, 2009). If there is no demarcation between the conceptual definitions of the proposed dualities and empirical hypotheses about covariations between dualities, there is a considerable risk that the value of dual-process theories will be reduced to post hoc categorizations of empirical effects. An illustrative example is the common equation of System 1 processing with resource-independence and System 2 processing with resource-dependence (e.g., Dhar & Gorlin, 2013). The problem is that such an equation does not offer anything beyond post hoc classifications of observed effects (Gawronski, 2013).To the extent that a given effect is resource-independent it will be categorized as being due to System 1, but it will be attributed to System 2 if it is resource dependent. Moreover, if an effect that was initially attributed to System 1 turns out to depend on cognitive resources, this effect would simply be recategorized as the product of System 2, and vice versa. Without a clear specification of the operating principles of System 1 and System 2 processing, the theory does not impose any constraints on the interpretation of a given result. Thus, criticism of dual-process theories as being unfalsifiable can be avoided by (1) conceptually precise definitions of the proposed dualities and (2) clear formulations of empirical hypotheses about their covariations.

To the extent that the two requirements are met, dual-process theories are in fact highly falsifiable, counter to the common criticism that they are unfalsifiable. According to Popper (1934), the falsifiability of a given theory increases with the number of events that are prohibited by the theory (see Gawronski & Bodenhausen, in press). Thus, by assuming systematic overlap between two conceptually distinct dualities, dual-process theories can be said to prohibit two out of four possible events. For example, dual-process theories of dispositional attribution claiming that the impact of situational information is mediated by a resource-dependent process, whereas the impact of behavioral information is mediated by a resource-independent process, would be disconfirmed by any finding showing that either (1) the impact of behavioral information is resource-dependent or (2) the impact of situational information is resource-independent (e.g., Krull, 1993; Trope & Gaunt, 2000). Importantly, the number of events prohibited by dual-process theories—and thus their falsifiability—increases with the number of proposed covariations between dualities. For example, a dual-systems theory assuming systematic overlap between four conceptually distinct dualities would prohibit 14 out of 16 possible combinations. Such a

theory would be highly falsifiable, because it would be disconfirmed by any one of the 14 cases that are prohibited by the theory. To be sure, certain assumptions about systematic overlap between multiple dualities may be rejected as empirically implausible or inconsistent with the available evidence (Keren & Schul, 2009; Kruglanski & Gigerenzer, 2011). For example, the assumption that the associative–propositional, affective–cognitive, and unconscious–conscious dualities have systematic overlap may be rejected on the grounds that there are cognitive associative processes, unconscious cognitive processes, conscious affective processes, and so forth. However, because such rejections are based on empirical arguments, they invalidate rather than support the common contention that dual-process theories are unfalsifiable.

Despite our rejection of generalized claims of unfalsifiability, we believe that there is another, less obvious feature that can make dual-process theories unfalsifiable. Many domain-specific dual-process theories were formulated in a manner such that the relevant inputs were specified at the level of external stimuli (e.g., source characteristics, situational information, category membership). The same was true for the proposed outputs, which were clearly specified at the level of overt responses (e.g., evaluative judgments, trait judgments). These specifications made it relatively easy to determine whether a theoretically derived prediction about input–output relations is consistent or inconsistent with a given finding. With the emergence of generalized dual-process theories, the focus shifted toward properties of the mind's internal architecture. In these theories, many of the core assumptions are specified at the mental level, without reference to external stimuli and overt responses. For example, in Kahneman's (2003) dual-system theory of judgment and choice, System 1 processing is claimed to involve a process of attribute substitution, in which a target attribute is substituted by a heuristic attribute that comes more readily to mind. Moreover, System 2 processing is assumed to monitor and, if necessary, modify or override the outputs of System 1. However, the theory itself does not specify which stimulus attributes count as target attributes and which ones should be regarded as heuristic attributes.[3] Thus, enhanced effects of any given attribute may be explained by either (1) dominance of System 1 processing involving the replacement of a target attribute by another heuristic attribute or (2) dominance of System 2 processing involving the replacement of the output of System 1 processing. In the absence of "bridging" assumptions that link the hypothesized mental constructs to external inputs and overt outputs, it can be rather difficult to determine whether a given theory is consistent or inconsistent with a particular finding (McGrath, 1981). This difficulty can lead to the impression that dual-process theories are unfalsifiable, although their assumptions about covariations between two or more dualities clearly prohibit a large number of possible events. In other words, although dual-process theories are falsifiable *in principle* if they are evaluated in terms of Popper's (1934) structural criterion, some of them are *practically* unfalsifiable, because they are formulated in a manner that makes it difficult to determine which relations between external stimuli and overt behavioral responses would be inconsistent with these theories. This feature allows researchers to apply some dual-process theories to virtually every empirical outcome in a post hoc fashion (e.g., post hoc categorizations of a given effect as being driven by System 1 vs. System 2). However, it is rather difficult to derive a priori predictions that could pose a strong challenge to these theories.

HOW MANY PROCESSES ARE THERE?

A common question about dual-process theories is whether the realm of mental processes can indeed be meaningfully divided into two distinct categories. Whereas some theorists have argued that dual-process theories clearly demonstrated their explanatory and predictive value (Deutsch & Strack, 2006), others have argued for the superiority of single-process (e.g., Kruglanski et al., 2006) or multiple process alternatives (e.g., Sherman, 2006). Yet when discussing the question of how many processes there "really" are, it is important to note that existence claims—including claims about the existence of one, two, or multiple processes—are ontological in nature. In the

philosophy of science, ontological claims fall into the realm of metaphysics, which means that they cannot be tested empirically (e.g., Popper, 1934; Quine, 1960). From this perspective, it is not possible to test whether there are one, two, or multiple processes. However, researchers can make decisions about the usefulness of ontological claims by empirically testing assumptions *about* the proposed processes. To the extent that dual-process theories help us to understand past observations of input–output relations (*explanation*) and correctly forecast future observations of input–output relations (*prediction*), there is reason to believe that their assumptions are correct, including their ontological assumptions about the existence of certain mental processes. However, if the predictions of a given theory are continuously disconfirmed, it seems likely that researchers will at some point reject its underlying ontological claims (cf. Peters & Gawronski, 2011). Note, however, that in such cases it is not the existence-claim itself that is confirmed or disconfirmed, but the assumptions that are made about the proposed entities. In this sense, the number of mental processes is not an objective fact that can be studied independently of the assumptions that are made about them. Instead, the number of mental processes that we assume to exist inherently depends on our theories, in that we accept the ontology of those theories that help us to understand past observations (*explanation*) and correctly forecast future observations (*prediction*).

An important issue in this context is the quest for *parsimony,* which stipulates that researchers should favor theories that involve fewer assumptions to explain a particular empirical finding (Gawronski & Bodenhausen, in press). Single-process theorists sometimes appeal to the quest for parsimony, arguing that dual-process theories are less parsimonious than single-process theories, because they postulate two qualitatively distinct processes rather than a single one. However, an often-overlooked aspect of parsimony is that it refers to the *total* number of theoretical assumptions that are required to explain a given finding rather than the number of assumptions of what might be considered the core of a given theory. To explain a particular finding, single-process theories have to rely on a host of additional assumptions over and above the hypothesis that information processing is guided by a single process (e.g., modulation of this process by five conceptually distinct parameters; see Kruglanski et al., 2006). Thus, when evaluating theories on the basis of their parsimony, it does not suffice to count the number of processes they propose. What matters for the criterion of parsimony is the total number of assumptions that is required to explain a given finding.

A final issue concerns the most appropriate level of abstraction in theorizing about mental processes. In some sense, every mental process can be described by the shared feature that they follow if–then rules. Some researchers have interpreted this possibility as an argument for the superiority of single-process theories that attribute all kinds of input–output relations to the same process of rule-based inference (e.g., Kruglanski et al., 2006). However, descriptions of mental processes at such a high level of abstraction do not go far beyond claiming that all mental processes follow some kind of regularity instead of being random (Gawronski & Creighton, 2013). Moreover, to the extent that such if–then conditionals directly refer to inputs as the antecedent and outputs as the consequent, their theoretical claims would be located at Marr's (1982) computational level of analysis. However, computational analysis fails to specify the mental processes and representations that translate inputs into outputs at the algorithmic level, which is a central concern of dual-process theorizing. In this sense, the debate between advocates of single-process and dual-process theories seems at least partly spurious, in that the conflicting claims refer to different levels of analysis (cf. De Houwer & Moors, in press; see also Sherman, 2006).

Similar considerations apply to the debate between dual-process and multiple-process theories. In some sense, every process can be further analyzed for subcomponents that constitute this process. For example, although the distinction between associative and propositional processes may serve as the conceptual foundation of a dual-process theory of evaluation (Gawronski & Bodenhausen, 2006), the functional principles of the two processes may be further divided depending on whether they operate during the formation or the expression of evaluative

representations. Moreover, one could argue that the process of propositional validation involves a number of subprocesses, including the default affirmation of validity, the monitoring of consistency, and the resolution of inconsistency (Gawronski & Bodenhausen, 2011). In this sense, every dual-process theory may be regarded as an oversimplification, in that its proposed dualities can always be divided into multiple subordinate processes. In fact, for many theories that have been categorized as *dual-process theories,* the term is actually a misnomer given that the majority of these theories propose more than two processes that are involved in the translation of inputs into outputs.

We argue that the optimal level of theoretical abstraction cannot be determined a priori but has to be decided on the basis of the research question (Sherman, 2006). For many social psychological questions, the conceptual distinctions proposed by dual-process theories have clearly demonstrated their value in explaining and predicting the phenomena of interest. However, for other questions, more fine-grained theories may be needed to account fully for the available evidence. Yet regardless of the chosen level of abstraction, it is important that the proposed process dimensions be clearly defined and their conceptual definitions be distinguished from empirical hypotheses about covariations between different dimensions.

EMERGING THEMES AND FUTURE DIRECTIONS

An interesting line of research that is starting to emerge from dual-process theorizing concerns the mental processes underlying operating conditions. This statement may seem a little puzzling given our strong emphasis on the distinction between operating principles and operating conditions. Yet it is entirely possible to stipulate a clear distinction between operating principles and operating conditions and, at the same time, ask questions about the processes underlying operating conditions.

Traditionally, dual-process theories aimed at identifying the mental processes underlying particular kinds of input–output relations. These processes are conceptually defined by their operating principles, and assumptions about their operating conditions reflect empirical hypotheses about whether they operate (1) when there is no conscious awareness, (2) when there is no goal to start the process, (3) when cognitive resources are reduced, and (4) when there is a goal to alter or stop the process. Operating conditions are typically investigated by means of appropriate manipulations, including supraliminal versus subliminal presentation times (awareness), the presence versus absence of instructions to perform a particular mental operation (intentionality), the relative difficulty of a simultaneously performed secondary task (efficiency), and instructions not to perform a particular mental operation (controllability). Any of these manipulations can be conceptualized as involving contextual inputs (or input characteristics) that may moderate the relation between a primary input and its output (De Houwer & Moors, 2012). In line with this conceptualization, Marr (1982) located operating conditions at the computational rather than the algorithmic level, which is consistent with our emphasis on the distinction between operating principles and operating conditions.

Nevertheless, it is certainly possible to investigate the mental processes by which the inputs of operating conditions influence the input–output relations produced by another process. In this case, the moderating effect of a given input on other input–output relations represents the phenomenon that needs to be explained (*explanandum*), and researchers may develop mental process theories that explain *how* this input moderates input–output relations (*explanans*). For example, researchers may investigate the mental processes by which the affordances of secondary tasks influence the operation of another mental process. Similarly, researchers may be interested in the mental processes by which instructions to inhibit a particular process influence the operation of that process. Such theories differ from traditional dual-process theories, in that they focus on how secondary processes influence the operation of the primary processes postulated by traditional dual-process theories. Applied to the four features of automaticity, potential themes of such theories may include (1) the processes by which subliminal versus supraliminal stimulus presentations moderate the

execution of a primary process, (2) the processes by which the presence versus absence of instructions to perform a particular mental operation moderates the execution the primary process, (3) the processes by which secondary task performance moderates the execution of a primary process, and (4) the processes by which instructions not to perform a particular mental operation moderate the execution of the primary process.

What may already be clear from this conceptualization is that research on the mental processes underlying operating conditions requires a sufficient understanding of the input–output relations that are produced by the relevant primary processes. Although such an assumption may seem questionable to critics of dual-process theorizing, cognitive science has a long history of research on similar questions, including the nature of consciousness (Baars, 2002), working memory (Baddeley, 2010), and executive control (Braver, 2012). Of course, the paradigms in this research tend to be much less complex than the ones typically used by social psychologists. Nevertheless, we believe that research under the guidance of dual-process theories has accumulated a sufficiently large body of knowledge to move to the next level of inquiry by studying the interplay between the proposed primary processes and the secondary processes that determine the conditions of their operation. To the extent that the theories inspired by this research adhere to the distinction between operating principles and operating conditions for both primary and secondary processes, they may also provide valuable insights into when and why certain features of automaticity do or do not co-occur (Moors & De Houwer, 2006). Examples of emerging themes in this regard include recent claims about the resource-independence of inhibitory control (e.g., Fujita, 2011) and the resource-dependence of unconscious processing (e.g., Hassin et al., 2009).

CONCLUSION

Emerging in the 1980s, the first generation of dual-process theories has yielded a multiplicity of dualities in theorizing about social information processing. Cumulatively, this work has demonstrated the ability of dual-process theories to be generative and applicable to a broad spectrum of social psychological phenomena. Since the publication of Chaiken and Trope's (1999) seminal volume, *Dual-Process Theories in Social Psychology*, these theories have continued to be highly generative and have further expanded the range of application. At the same time, attempts at integration have shifted the focus to formulating unitary distinctions between different mental processes that cut across multiple domains. Our hope is that future advances will help dual-process theories to further enhance the basic science desiderata of conceptual coherence, precision, generality, and testability without sacrificing their applicability to the problems humans face as members of dyads, groups, and societies.

In this spirit, our main goal in this chapter has been to analyze the explanatory function of dual-process theories, the conceptual nature of their underlying dualities, and the structural features that characterize different types of dual-process theories. Drawing on Marr's (1982) distinction among computational, algorithmic, and implementational levels of analysis, we have argued that a central explanatory function of dual-process theories is to specify the mental mechanisms that translate inputs into outputs. From this perspective, descriptions of a given process as automatic versus nonautomatic simply specify *when* the process is assumed to operate (operating conditions). However, they do not specify *how* the proposed process translates inputs into outputs (operating principles). The latter question requires clear specifications of the underlying mental operations, which are essential for mechanistic explanations at Marr's algorithmic level of analysis. To the extent that dual-process theories include precise and nonoverlapping definitions of the proposed dualities, their hypotheses about covariations between processing dualities are highly falsifiable, in that they prohibit a considerable number of possible events. Yet the explanatory and predictive value of dual-process theories essentially depends on (1) a clear distinction between operating principles and operating conditions, (2) conceptual rigor in the definition of the proposed dualities, (3) precise formulations of empirical hypotheses about covariations between processing dualities, and (4) clearly specified links between

the hypothesized mental processes and the causal relations between stimuli and behavior they are supposed to explain.

NOTES

1. Automatic processes are often contrasted with controlled processes. Yet the term *control* has been used to refer to either (1) a particular feature of nonautomatic processing (i.e., controllability) or (2) an umbrella concept subsuming multiple different features of nonautomatic procesing (i.e., awareness, intentionality, resource dependence, controllability). To avoid conceptual confusion, we use the term *nonautomatic* as the semantic antonym of the term *automatic* instead of the more common term *controlled* (see Moors & De Houwer, 2006).
2. Some dual-process theories derive empirical assumptions about operating conditions from their conceptual definitions of operating principles (e.g., inferences about resource dependence from the conceptual distinction between sequential and parallel processing). Although such inferences are theoreticaly valuable to clarify the conceptual basis of the derived hypotheses about operating conditions, it is important to note that operating principles and operating conditions are nevertheless conceptually distinct in these theories, in that the former describe the nature of the mental processes that translate inputs into outputs, whereas the latter refer to the particular conditions under which these processes are assumed to operate.
3. Note that the specification of heuristic attributes as coming more readily to mind does not provide a clear conceptual demarcation, because the accessibility of any given attribute can vary as a function of its salience.

REFERENCES

Baars, B. J. (2002). The conscious access hypothesis: Origins and recent evidence. *Trends in Cognitive Sciences, 6,* 47–52.

Baddeley, A. (2010). Working memory. *Current Biology, 20,* 136–140.

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

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

Bargh, J. A. (2006). What have we been priming all these years?: On the development, mechanisms, and ecology of nonconscious social behavior. *European Journal of Social Psychology, 36,* 147–168.

Baumeister, R. F., Masicampo, E. J., & Vohs, K. D. (2011). Do conscious thoughts cause behavior? *Annual Review of Psychology, 62,* 331–361.

Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences, 16,* 106–113.

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: Erlbaum.

Cacioppo, J. T., Berntson, G. G., Sheridan, J. F., & McClintock, M. K. (2000). Multilevel integrative analyses of human behavior: Social neuroscience and the complementing nature of social and biological approaches. *Psychological Bulletin, 126,* 829–843.

Carruthers, P. (2009). An architecture for dual reasoning. In J. St. B. T. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 109–127). New York: Oxford University Press.

Cesario, J., Plaks, J. E., & Higgins, E. T. (2006). Automatic social behavior as motivated preparation to interact. *Journal of Personality and Social Psychology, 90,* 893–910.

Chaiken, S. (1987). The heuristic model of persuasion. In M. P. Zanna, J. M. Olson, & C. P. Herman (Eds.), *Social influence: The Ontario Symposium* (Vol. 5, pp. 3–39). Hillsdale, NJ: Erlbaum.

Chaiken, S., & Trope, Y. (Eds.). (1999). *Dual-process theories in social psychology.* New York: Guilford Press.

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

De Houwer, J. (2011). Why the cognitive approach in psychology would profit from a functional approach and vice versa. *Perspectives on Psychological Science, 6,* 202–209.

De Houwer, J., & Moors, A. (2012). How to

define and examine implicit processes? In R. Proctor & J. Capaldi (Eds.), *Implicit and explicit processes in the psychology of science* (pp. 183–198). New York: Oxford University Press.

De Houwer, J., & Moors, A. (in press). Levels of analysis. In B. Gawronski & G. V. Bodenhausen (Eds.), *Theory and explanation in social psychology*. New York: Guilford Press.

Deutsch, R., & Strack, F. (2006). Duality models in social psychology: From dual processes to interacting systems. *Psychological Inquiry, 17,* 166–172.

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

Dhar, R., & Gorlin, M. (2013). A dual-system framework to understand preference construction processes in choice. *Journal of Consumer Psychology, 23,* 528–542.

Dijksterhuis, A., & Nordgren, L. F. (2006). A theory of unconscious thought. *Perspectives on Psychological Science, 1,* 95–109.

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist, 49,* 709–724.

Evans, J. St. B. T, & Frankish, K. (Eds.). (2009). *In two minds: Dual processes and beyond*. New York: Oxford University Press.

Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. *Advances in Experimental Social Psychology, 23,* 75–109.

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.

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. *Advances in Experimental Social Psychology, 23,* 1–74.

Friese, M., Hofmann, W., & Schmitt, M. (2008). When and why do implicit measures predict behaviour? Empirical evidence for the moderating role of opportunity, motivation, and process reliance. *European Review of Social Psychology, 19,* 285–338.

Fujita, K. (2011). On conceptualizing self-control as more than the effortful inhibition of impulses. *Personality and Social Psychology Review, 15*, 352–366.

Gawronski, B. (2013). What should we expect from a dual-process theory of preference construction in choice? *Journal of Consumer Psychology, 23,* 556–560.

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

Gawronski, B., & Bodenhausen, G. V. (2009). Operating principles versus operating conditions in the distinction between associative and propositional processes. *Behavioral and Brain Sciences, 32,* 207–208.

Gawronski, B., & Bodenhausen, G. V. (2011). The associative–propositional evaluation model: Theory, evidence, and open questions. *Advances in Experimental Social Psychology, 44,* 59–127.

Gawronski, B., & Bodenhausen, G. V. (in press-a). Theory evaluation. In B. Gawronski & G. V. Bodenhausen (Eds.), *Theory and explanation in social psychology*. New York: Guilford Press.

Gawronski, B., & Bodenhausen, G. V. (in press-b). Social-cognitive theories. In B. Gawronski & G. V. Bodenhausen (Eds.), *Theory and explanation in social psychology*. New York: Guilford Press.

Gawronski, B., & Creighton, L. A. (2013). Dual-process theories. In D. E. Carlston (Ed.), *The Oxford handbook of social cognition* (pp. 282–312). New York: Oxford University Press.

Gilbert, D. T. (1989). Thinking lightly about others: Automatic components of the social inference process. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 189–211). New York: Guilford Press.

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

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

Hassin, R. R., Bargh, J. A., Engell, A. D., & McCulluch, K. C. (2009). Implicit working memory. *Consciousness and Cognition*, *18,* 665–678.

Huang, J. Y., & Bargh, J. A. (in press). The Selfish Goal: Autonomously operating motivational structures as the proximate cause of human judgment and behavior. *Behavioral and Brain Sciences*.

Jacoby, L. L. (1991). A process-dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*, 513–541.

Johnson-Laird, P. N. (2012). Mental models and consistency. In B. Gawronski & F. Strack (Eds.), *Cognitive consistency: A fundamental principle in social cognition* (pp. 225–244). New York: Guilford Press.

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

Keren, G., & Schul, Y. (2009). Two is not always better than one: A critical evaluation of two-systems theories. *Perspectives on Psychological Science, 4,* 533–550.

Klauer, K. C. (in press). Mathematical modeling. In B. Gawronski & G. V. Bodenhausen (Eds.), *Theory and explanation in social psychology*. New York: Guilford Press.

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

Krieglmeyer, R., & Sherman, J. W. (2012). Disentangling stereotype activation and stereotype application in the Stereotype Misperception Task. *Journal of Personality and Social Psychology*, *103*, 205-224.

Kruglanski, A. W., Erb, H.-P., Pierro, A., Mannetti, L., & Chun, W. Y. (2006). On parametric continuities in the world of binary either ors. *Psychological Inquiry, 17,* 153–165.

Kruglanski, A. W., & Gigerenzer, G. (2011). Intuitive and deliberative judgments are based on common principles. *Psychological Review, 118,* 97–109.

Krull, D. S. (1993). Does the grist change the mill?: The effect of the perceiver's inferential goal on the process of social inference. *Personality and Social Psychology Bulletin, 19,* 340–348.

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., Gaunt, R., Gilbert D. T., & Trope, Y. (2002). Reflexion and reflection: A social cognitive neuroscience approach to attributional inference. *Advances in Experimental Social Psychology, 34,* 199–249.

Loersch, C. & Payne, B. K. (2011). The situated inference model: An integrative account of the effects of primes on perception, behavior, and motivation. *Perspectives on Psychological Science, 6,* 234–252.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York: Freeman.

McGrath, J. E. (1981). Dilemmatics: The study of research choices and dilemmas. *American Behavioral Scientist, 25,* 179–210.

Meissner, F., & Rothermund, K. (2013). Estimating the contributions of associations and recoding in the Implicit Association Test: The ReAL model for the IAT. *Journal of Personality and Social Psychology, 104*, 45–69.

Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences*, *32*, 183–198.

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

Nadarevic, L., & Erdfelder, E. (2011). Cognitive processes in implicit attitude tasks: An experimental validation of the Trip Model. *European Journal of Social Psychology, 41*, 254–268.

Ochsner, K. N., & Lieberman, M. D. (2001). The emergence of social-cognitive neuroscience. *American Psychologist, 56*, 717–734.

Payne, B. K., & Bishara, A. J. (2009). An integrative review of process dissociation and related models in social cognition. *European Review of Social Psychology, 20,* 272–314.

Payne, B. K., & Gawronski, B. (2010). A history of implicit social cognition: Where is it coming from? Where is it now? Where is it going? In B. Gawronski, & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 1–15). New York: Guilford Press.

Payne, B. K., Hall, D. L., Cameron, C. D., & Bishara, A. J. (2010). A process model of affect misattribution. *Personality and Social Psychology Bulletin, 36*, 1397–1408.

Peters, K. R., & Gawronski, B. (2011). Are we puppets on a string?: Comparing the effects of contingency and validity on implicit and explicit evaluations. *Personality and Social Psychology Bulletin, 37,* 557–569.

Petty, R. E., & Cacioppo, J. T. (1986). The Elaboration Likelihood Model of persuasion.

Advances in Experimental Social Psychology, 19, 123–205.

Popper, K. R. (1934). *Logic of scientific discovery.* New York: Basic Books.

Quine, W. V. O. (1960). *Word and object.* 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.

Samuels, R. (2009). The magical number two, plus or minus: Dual-process theory as a theory of cognitive kinds. In J. St. B. T. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 129–146). New York: Oxford University Press.

Sherman, J. W. (2006). On building a better process model: It's not only how many, but which ones and by which means. *Psychological Inquiry, 17,* 173–184.

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

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

Smedslund, G. (2000). A pragmatic basis for judging models and theories in health psychology: The axiomatic method. *Journal of Health Psychology, 5,* 133–149.

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.

Stahl, C., & Degner, J. (2007). Assessing automatic activation of valence: A multinomial model of EAST performance. *Experimental Psychology, 54,* 99–112.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences, 23,* 645–726.

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

Trope, Y. (1986). Identification and inferential processes in dispositional attribution. *Psychological Review, 93,* 239–257.

Trope, Y., & Gaunt, R. (2000). Processing alternative explanations of behavior: Correction or integration? *Journal of Personality and Social Psychology, 79,* 344–354.

Wallach, L., & Wallach, M. A. (1994). Gergen versus the mainstream: Are hypotheses in social psychology subject to empirical test? *Journal of Personality and Social Psychology, 67,* 233–242.

Wheeler, S. C., DeMarree, K. G., & Petty, R. E. (2007). Understanding the role of the self in prime-to-behavior effects: The active-self account. *Personality and Social Psychology Review, 11,* 234–261.

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