In D. E. Carlston (Ed.). (2013). The Oxford handbook of social cognition (pp. 282-312). New York, NY: Oxford University Press.

CHAPTER 14

Dual Process Theories

Bertram Gawronski *and* Laura A. Creighton

Abstract

Dual process theories divide the realm of mental processes into two general categories depending on whether they operate automatically or in a controlled fashion. This chapter provides an overview of dual process theories in social psychology, focusing on their historical and conceptual developments. Identifying three general categories of dual process theories, the chapter distinguishes between domain-specific theories that focus on particular phenomena, generalized dual process theories that identify domain-independent principles underlying various kinds of phenomena, and formalized dual process theories that quantify the joint contributions of automatic and controlled processes to responses within a single task. The chapter also discusses critical arguments against each type of dual process theorizing, which are integrated in a general outlook on future directions.

Key Words: attitudes, attribution, automaticity, control, dual process theories, impression formation, persuasion, prejudice, stereotyping

Introduction

For the past three decades, a large body of research in social cognition has been shaped and guided by a class of theories that are generically described as *dual process theories* (Chaiken & Trope, 1999). The defining characteristic of these theories is that they divide the mental processes underlying social judgments and behavior into two general categories depending on whether they operate automatically or in a controlled fashion. Because the distinction between automatic and controlled processes has become a central component in virtually all areas of social psychology (see Chapter 12), it is difficult to imagine what contemporary social psychology would look like without the theoretical guidance of dual process theories.

Whereas early dual process theories focused primarily on domain-specific phenomena, such as persuasion (e.g., Chaiken, 1987; Petty & Cacioppo, 1986), attitude–behavior relations (e.g., Fazio, 1990; Wilson, Lindsey, & Schooler, 2000), prejudice and stereotyping (e.g., Devine, 1989), impression formation (e.g., Brewer, 1988; Fiske & Neuberg, 1990), and dispositional attribution (e.g., Gilbert, 1989; Trope, 1986), dual process theorizing in the past decade shifted toward integrative models that aim at identifying general principles assumed to be domain independent. These integrative models can be further divided into generalized dual process theories that describe mental processing by means of two domain-independent operating principles (e.g., Epstein, 1994; Kahneman, 2003; Lieberman, 2003; Smith & DeCoster, 2000; Strack & Deutsch, 2004) and formalized dual process theories that quantify the contributions of distinct mental processes to behavioral responses by means of mathematical modeling techniques (e.g., Payne, 2008; Sherman et al., 2008).

The aim of the current chapter is to provide a general overview of dual process theories in social cognition with the goal of identifying both historical and conceptual developments since their emergence in the mid-1980s. For this purpose, we first provide a brief discussion of the distinction between automatic and controlled processes, which serves as a basis for our overview of dual process theories. We then review the most influential phenomenon-specific theories that have set the foundation for the ubiquitous dual process paradigm within social psychology. Expanding on this review, the following two sections discuss the tenets of generalized and formalized dual process theories that have gained considerable impact within and beyond social psychology during the past decade. At the end of each section, we also discuss critical arguments that have been raised against phenomenon-specific, generalized, and formalized dual process theories, which are integrated in our outlook in the final part of this chapter.

Automaticity and Control

Dual process theories have their roots in the assumption that the universe of mental processes can be divided into two general classes: those that operate automatically and those that operate in a controlled fashion (Posner & Snyder, 1975; Shiffrin & Schneider, 1977). In social cognition, automatic processes are typically characterized in terms of four operating conditions: (1) they are elicited unintentionally; (2) they require little amounts of cognitive resources; (3) they cannot be stopped voluntarily; and (4) they occur outside of conscious awareness (Bargh, 1994). Conversely, controlled processes are characterized as those that (1) are initiated intentionally; (2) require considerable amounts of cognitive resources; (3) can be stopped voluntarily; and (4) operate within conscious awareness (for a more fine-grained analysis of these features and their interrelations, see Moors & De Houwer, 2006).

Initially, automatic and controlled processes were conceptualized in an all-or-none fashion, implying that a given process can be characterized either by the four features of automatic processing or by the four features of controlled processing. According to this dual mode view, the "four horsemen" of automaticity (i.e., intentionality, efficiency, controllability, awareness) constitute a fixed set of characteristics that are perfectly correlated (see Moors & De Houwer, 2006). However, challenging the usefulness of this all-or-none conceptualization, it soon turned out that there is virtually no process that meets all four operating criteria (Bargh, 1992). Instead, most processes studied within social psychology involve combinations of the proposed features, which classify them as "automatic" in one sense and "controlled" in another (Bargh, 1994). This insight has inspired a disjunctive conceptualization of automaticity, according to which a process can be characterized as automatic if it meets at least one of the four criteria of automaticity. That is, a process can be described as automatic if it is either (1) unintentional, (2) efficient, (3) uncontrollable, or (4) unconscious.

Even though the disjunctive treatment of automaticity is rather common in social psychology, it involves several problems. First, if the presence of a single feature is sufficient to call a process automatic or controlled, it is possible that a given process will have to be described as automatic and controlled at the same time. Needless to say, such a description can be rather confusing if it is not specified in which particular sense the process is described as automatic and in which sense it is described as controlled. Second, the 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 does not require a large amount of cognitive resources, whereas another Process B may be described as automatic because it is elicited unintentionally. Yet, the two processes may still differ in their operating conditions, if, for example, Process A, but not Process B, can be stopped voluntarily. Based on these considerations, several theorists have recommended that researchers should be more precise in their use of terminology by describing each feature of automaticity with its proper label, that is, unintentional, efficient, uncontrollable, or unconscious (Bargh, 1994; Moors & De Houwer, 2006). As we will outline in the following sections, different kinds of dual process theories emphasize different features of automaticity, which makes terminological precision particularly important to avoid conceptual confusion.

Phenomenon-Specific Dual Process Theories

As noted in the introductory section of this chapter, early dual process theories tended to be domain specific in that they focused on particular phenomena. In the current section, we first review the core assumptions of the most influential theories of this kind and then discuss criticism that has been raised against phenomenon-specific dual process theories.

Persuasion

Two of the most prominent dual process theories are the elaboration likelihood model (ELM; Petty & Cacioppo, 1986) and the heuristic systematic model (HSM; Chaiken, 1987) of persuasion. The central question of these models concerns the conditions under which different aspects of a persuasive message (e.g., strength of arguments, attractiveness of the source) influence the effectiveness of persuasive appeals.

ELABORATION LIKELIHOOD MODEL

The central notion of Petty and Cacioppo's (1986) ELM is that attitude change occurs along an elaboration continuum whereby persuasion is determined by how motivated and able an individual is to engage in effortful information processing (Figure 14.1). The basic assumption is that the higher an individual's cognitive elaboration, the more likely he or she is to process all object-relevant information. At the high end of the elaboration continuum, people

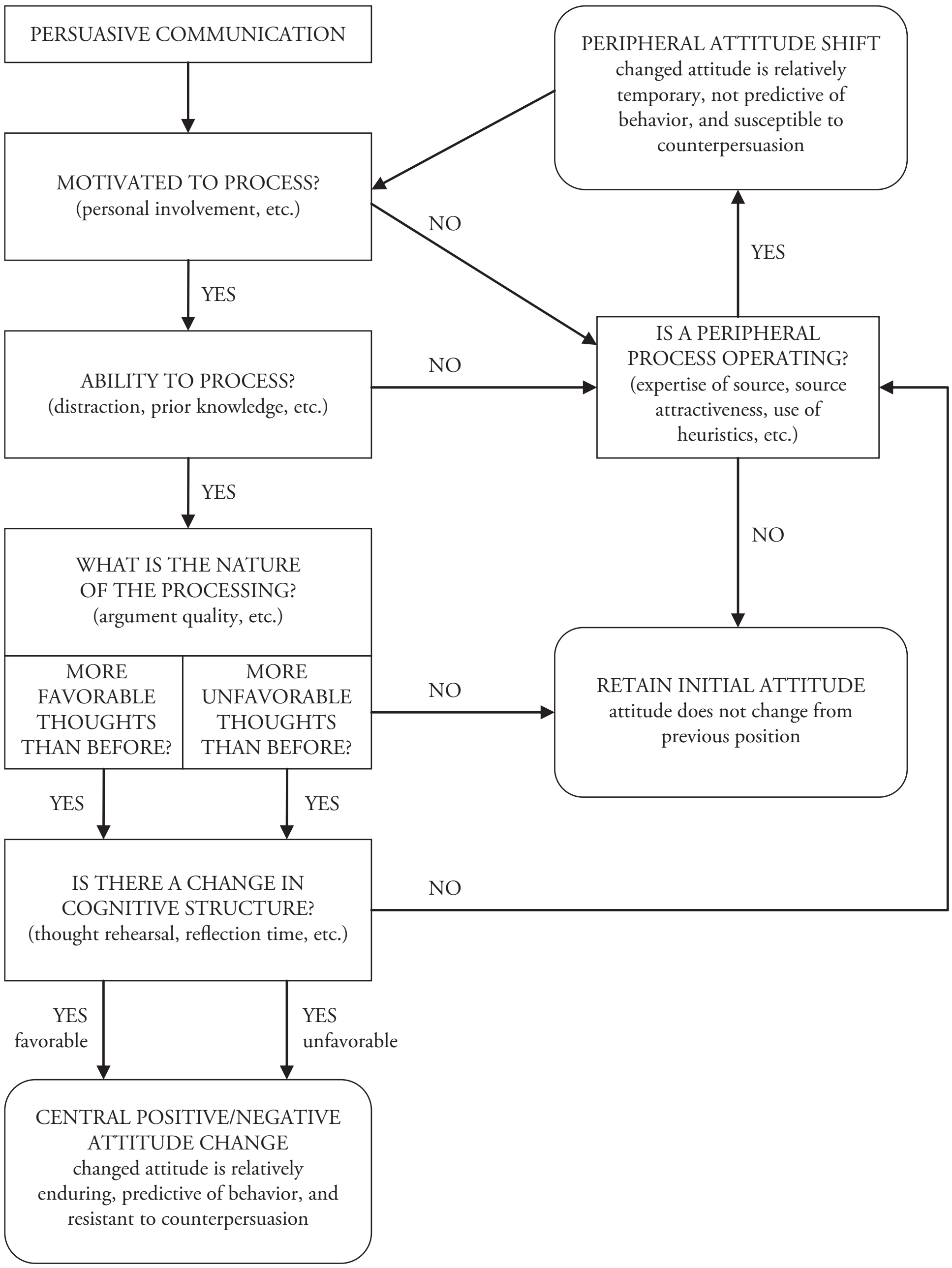

Figure 14.1 The elaboration likelihood model of persuasion. (Adapted from Petty & Cacioppo, 1986. Reprinted with permission.)

assess all of the available object-relevant information (e.g., strength of the presented arguments) and integrate this information with their stored knowledge in order to obtain a carefully considered (although not necessarily unbiased) evaluation (*central route*). Conversely, at the low end of the elaboration continuum, people engage in considerably less scrutiny of object-relevant information (*peripheral route*). When elaboration is low, attitude change can be effected from a cursory examination of the available information (e.g., by examining only a subset of the available information) or by the use of heuristics and other types of information processing shortcuts (e.g., *I agree with people I like*). Compared with attitudes that are changed through the central route, attitudes changed through the peripheral route tend to be relatively weak, susceptible to counterpersuasion, and less predictive of behavior (e.g., Petty, Cacioppo, & Schumann, 1983).

Addressing a common misconception about the ELM, Petty and Wegener (1999) pointed out that the influence of a particular persuasion variable will not necessarily increase or decrease as one moves along the elaboration continuum. For example, a common interpretation of the ELM is that high elaboration increases the impact of primary features of the persuasive message (e.g., argument quality), whereas low elaboration increases the impact of secondary cues (e.g., source attractiveness). This interpretation is qualified by the ELM's *multiple-roles hypothesis,* which specifies that a given variable can influence attitudes by different processes at different points along the elaboration continuum. To illustrate this assumption, consider a commercial ad in which a physically attractive source endorses either a beauty product (relevant source attractiveness) or the services of a roofing company (irrelevant source attractiveness; see Petty & Cacioppo, 1984). Under conditions of low elaboration, source attractiveness may have a positive effect on evaluations of both the beauty product and the roofing company by virtue of mere association. Under conditions of high elaboration, however, source attractiveness may influence evaluations through a different process with different outcomes for the two products. Whereas evaluations of the beauty product may be influenced by source attractiveness because of its perceived relevance for evaluating the product, the persuasive effect on evaluations of the roofing company may be attenuated because of the perceived irrelevance of source attractiveness. Overall, the multiple-roles hypothesis states that any persuasion variable can influence attitudes in four different ways: (1) by serving as an argument, (2) by serving as a cue, (3) by determining the extent of cognitive elaboration, and (4) by producing a bias in the processing of the available information (Petty & Wegener, 1999). These assumptions make the model very flexible in accounting for a variety of effects in the persuasion literature. At the same time, the multiple-roles hypothesis makes it difficult to predict the outcome for a given persuasion situation, if the conditions under which a particular variable will take on any of the proposed roles are not specified.

HEURISTIC SYSTEMATIC MODEL

Similar to Petty and Cacioppo's (1986) ELM, Chaiken's (1987) HSM describes two basic persuasion processes that may guide an individual's judgments of an attitude object (see also Chen & Chaiken, 1999; Chaiken, Liberman, & Eagly, 1989). *Systematic processing* involves comprehensive consideration of object-relevant information, which requires high levels of motivation and ability to engage in effortful processing. *Heuristic processing*, in contrast, relies on the activation, accessibility, and applicability of learned heuristics that require relatively few cognitive resources (e.g., *I agree with people I like*). According to the HSM, the likelihood that an individual engages in systematic processing is guided by the *sufficiency principle*, which states that the motivation to engage in systematic processing increases to the extent that an individual's desired level of confidence falls below his or her actual level of confidence. That is, individuals are more likely to engage in systematic processing when the difference between their desired and their actual levels of confidence is high. Conversely, people are more likely to engage in heuristic processing when the difference between their desired and actual levels of confidence is low. Importantly, systematic processing may not necessarily lead to unbiased judgments because systematic processing can be influenced by *defense motivation* and *impression management.* Defense motivation refers to the desire to defend preexisting attitudes, whereas impression management refers to the desire to hold attitudes that satisfy specific social goals.

Another central assumption of the HSM is that heuristic and systematic processing may co-occur and interact with each other to exert either independent or interdependent effects on evaluations. First, according to the model's *attenuation hypothesis*, systematic processing can completely override the effects of heuristic processing (e.g., Maheswaran & Chaiken, 1991). Such attenuation effects are likely to occur when systematic processing yields information that contradicts the validity of simple persuasion heuristics. Second, the information generated by heuristic and systematic processing may jointly influence evaluations in an additive manner (e.g., Maheswaran, Mackie, & Chaiken, 1992). According to the model's *additivity hypothesis,* such effects are likely to occur when the two processing modes do not yield conflicting reactions. Finally, when the message content is ambiguous, heuristic cues may bias the effects of systematic processing, as described by the model's *bias hypothesis*. For example, if the strength of a persuasive argument is ambiguous, the argument may be perceived as more convincing if it is presented by an expert than if it is presented by a layperson (e.g., Chaiken & Maheswaran, 1994).

In terms of its core assumptions, the HSM shows considerable resemblance to the ELM (Petty & Cacioppo, 1986). For example, both models maintain that attitude change can occur through either (1) systematic/central processing that requires some degree of motivation and capacity or (2) heuristic/peripheral processing that is assumed to require little motivation or capacity. However, the two models differ in their treatment of heuristic and peripheral processes (Chen & Chaiken, 1999). Whereas the ELM assumes an inverse relationship between central and peripheral processing along the elaboration continuum, the HSM assumes that systematic and heuristic processing may occur simultaneously with either independent or interactive effects. Hence, the ELM holds that there is a trade-off between peripheral and central processing, such that the importance of one processing mode decreases as the importance of the other processing mode increases. In contrast, under the HLM's conceptualization, individuals can engage in systematic and heuristic processing simultaneously.

Attitude–Behavior Relations

Expanding on the question of how attitudes are formed and changed, another class of dual process theories describes the mechanisms through which attitudes guide behavior. These models have been inspired by recurring debates about whether and to what extent attitudes influence behavior (e.g., Wicker, 1969). By shifting the focus from asking "*Do attitudes guide behavior?*" to the question, "*How do attitudes guide behavior?*" dual process theorizing provided important insights into the conditions under which attitudes do or do not influence behavior.

MODE MODEL

One such dual process theory is Fazio's (1990) *M*otivation and *O*pportunity as *DE*terminants (MODE) model, which specifies two distinct processes by which attitudes can guide behavior depending on the person's motivation and opportunity to engage in deliberate processing (for recent reviews, see Fazio, 2007; Olson & Fazio, 2009). A central component of the MODE model is the definition of *attitude* as the mental association between an object and a person's summary evaluation of that object (Fazio, 1995, 2007). To the extent that this association is sufficiently strong, the evaluation associated with the object may be activated automatically upon encountering that object (i.e., without intention to evaluate the object; see Fazio, Sanbonmatsu, Powell, & Kardes, 1986). Such automatically activated attitudes are assumed to influence an individual's spontaneous interpretation of the current situation, which in turn will guide the individual's behavior without him or her necessarily being aware of the attitude's influence (*spontaneous attitude–behavior process*). Alternatively, individuals may scrutinize specific attributes of the object and the current situation (*deliberate attitude–behavior process*). However, such deliberate analyses require that an individual has both the motivation and the opportunity (i.e., adequate time and cognitive resources) to engage in effortful information processing. Thus, to the extent that either the motivation or the opportunity to engage in effortful processing is low, automatically activated attitudes may guide behavior through their effect on the spontaneous construal of the current situation. However, if both the motivation and the opportunity to engage in effortful processing are high, the impact of automatically activated attitudes on behavior will depend on particular aspects of the current situation, including specific attributes of the attitude object or salient norms (Fazio, 1990; see also Ajzen & Fishbein, 1980).

An important application of the MODE model concerns the relation between explicit and implicit measures of attitudes (see Fazio & Olson, 2003, Olson & Fazio, 2009). According to the MODE model, implicit measures of attitudes—such as evaluative priming (Fazio, Jackson, Dunton, & Williams, 1995) or the implicit association test (Greenwald, McGhee, & Schwartz, 1998)—reduce participants' opportunity to engage in effortful processing. Consequently, participants' responses on these measures will directly reflect their automatically activated attitudes (but see Olson & Fazio, 2004). In contrast, verbally reported evaluations assessed by explicit measures are relatively easy to control. Thus, to the extent that participants lack either the motivation or the opportunity to engage in effortful processing, explicit measures should reveal the same automatically activated attitudes that are reflected in implicit measures. If, however, participants have both the motivation and the opportunity to engage in a deliberate analysis of the current situation, explicit measures may reflect whatever evaluation is implied by these inferences. Within the domain of racial prejudice, for example, the influence of an automatically activated prejudicial attitude on verbally reported evaluations may be mitigated if an individual is motivated to control his or her prejudiced reactions. Consequently, automatic prejudicial attitudes assessed by implicit measures should be directly reflected in explicit evaluative judgments when the motivation to control prejudiced reactions is low, but not when it is high (e.g., Dunton & Fazio, 1997). More generally, these assumptions imply that implicit attitude measures should be better predictors of spontaneous behaviors that are relatively difficult to control (e.g., nonverbal behavior), whereas explicit attitude measures should be better predictors of deliberate behaviors that are relatively easy to control (e.g., content of verbal responses).

DUAL ATTITUDE MODEL

Similar predictions are implied by Wilson and colleagues' (2000) dual attitude model, even though its underlying assumptions about attitude representation are quite different from the core assumptions of the MODE model. According to the dual attitude model, people may simultaneously hold two attitudes toward the same object, which are described as *implicit attitude* and *explicit attitude* (see also Greenwald & Banaji, 1995). Such dual representations are assumed to emerge when a previously acquired attitude is challenged by counterattitudinal information, and the newly formed, explicit attitude does not erase the previously acquired, implicit attitude from memory. To the extent that the old, implicit attitude is highly overlearned, it may be activated automatically. In contrast, retrieving the newly formed, explicit attitude from memory is assumed to require cognitive effort. As a result, earlier acquired, implicit attitudes should guide judgments and behavior when either the motivation or the capacity to engage in effortful processing is low. However, judgments and behavior should be influenced by more recently acquired, explicit attitudes when both the motivation and the capacity to engage in effortful processing are high.

Another important implication of the dual attitude model is that implicit attitudes should be more difficult to change than explicit attitudes. Even though this prediction has been confirmed in several studies that used implicit measures to assess implicit attitudes and self-report measures to assess explicit attitudes (e.g., Gawronski & Strack, 2004; Gregg, Seibt, & Banaji, 2006), there is accumulating evidence that attitudes assessed with implicit measures can sometimes change rather quickly, with little counterattitudinal information (e.g., Gawronski & LeBel, 2008; Olson & Fazio, 2006). In addition, attitudes assessed with implicit measures have been shown to be highly context sensitive, such that the same object may elicit different evaluative responses as a function of the context in which it is encountered (for a review, see Gawronski & Sritharan, 2010). These findings have inspired the development of alternative models that were especially designed to explain different patterns of implicit and explicit attitude change (e.g., Gawronski & Bodenhausen, 2006a, 2011; Petty, Briñol, & DeMarree, 2007).

Prejudice and Stereotyping

One of the most striking findings in research on prejudice and stereotyping is that public opinion polls in North America showed a steady decline in negative evaluations of racial minority groups after World War II, whereas racial conflicts showed only a moderate reduction (e.g., Greeley & Sheatsley, 1971; Taylor, Sheatsley, & Greeley, 1978). This discrepancy inspired social psychologists to postulate more subtle forms of racial prejudice, such as modern (McConahay, 1986), aversive (Gaertner & Dovidio, 1986), or symbolic (Sears, 1988) racism. The general notion underlying these constructs is that racial prejudice has simply changed its face, rather than been abandoned. A very similar idea provided the inspiration for Devine's (1989) dissociation model, which was seminal in introducing the distinction between automatic and controlled processes to research on prejudice and stereotyping.

DISSOCIATION MODEL

A central aspect of Devine's (1989) dissociation model is the distinction between the *knowledge* of a social stereotype and the *belief* in the accuracy of that stereotype. According to Devine, both low-prejudice and high-prejudice individuals tend to be familiar with the contents of prevailing cultural stereotypes. However, the two groups differ with respect to their personal beliefs about the accuracy of these stereotypes. To the extent that stereotypic knowledge is acquired during early childhood and highly overlearned as a result of socialization processes, stereotypic knowledge is assumed to be activated automatically upon encountering members of stereotyped groups, and this occurs for both low-prejudice and high-prejudice individuals. In contrast, the rejection of stereotypic knowledge is assumed to be the result of egalitarian, nonprejudicial beliefs, which tend to be acquired later in the socialization process (e.g., Banse, Gawronski, Rebetez, Gutt, & Morton, 2010). Because these beliefs are less overlearned than earlier acquired stereotypic knowledge, replacing automatically activated stereotypes with egalitarian, nonprejudicial beliefs requires the operation of controlled processing. In other words, while the model assumes that automatic stereotype activation is equally strong and inescapable for both high-prejudice and low-prejudice individuals, the two groups differ at the level of controlled processing, such that low-prejudice but not high-prejudice individuals replace automatically activated stereotypes with egalitarian, nonprejudicial beliefs. In terms of the four characteristics of automaticity, these assumptions imply that the activation of social stereotypes occurs unintentionally, even though their impact on overt behavior can be controlled through effortful processes.

A notable aspect of Devine's (1989) dissociation model is that it implies a rather different view on the roles of personal beliefs and social contexts than the MODE model. Whereas the MODE model assumes that personal attitudes tend to be activated automatically and that the overt expression of these attitudes is sometimes suppressed when they conflict with social norms (Fazio et al., 1995), the dissociation model proposes that socially transmitted stereotypes are activated automatically and that the overt expression of these stereotypes are suppressed when they conflict with personal beliefs (Devine, 1989). In other words, whereas the MODE model locates an individual's authentic self at the level of automatic processes and extrinsic, social influences at the level of controlled processes, the dissociation model locates extrinsic, social influences at the level of automatic processes and the individual's authentic self at the level of controlled processes. Even though questions about what should be considered the "authentic self" are philosophical rather than empirical, these diverging views have important implications for the interpretation of automatic stereotypic biases (Gawronski, Peters, & LeBel, 2008). One example is the automatic tendency to misidentify harmless objects as weapons when they are held by a black person rather than a white person (for a review, see Payne, 2006). According to the MODE model, such unintentional errors reveal an individual's personal attitudes when the individual does not have the opportunity to adjust his or her automatic responses to egalitarian norms. In contrast, from the perspective of Devine's (1989) model, unintentional errors in weapon identification reveal the ubiquitous influence of cultural stereotypes that are independent of the individual's personal beliefs.

Impression Formation

Similar to Devine's (1989) dissociation model of prejudice and stereotyping, dual process theories of impression formation emphasize the role of social category information in early processing stages. However, whereas Devine's model focuses particularly on the unintentional activation versus controlled suppression of stereotypes, dual process theories of impression formation specify the conditions under which personal impressions of an individual are dominated by category-related or person-specific information.

CONTINUUM MODEL

Fiske and Neuberg's (1990) continuum model of impression formation proposes that the processes by which people form opinions of other individuals operate along a continuum that reflects the degree to which perceivers utilize category-related versus person-specific information (Figure 14.2). The basic assumption of the model is that category information enjoys general priority because the processing of such information does not require substantial amounts of cognitive resources. Specifically, perceivers are assumed to categorize individuals on the basis of salient category cues (e.g., gender, age, ethnicity), and this categorization is assumed to occur unintentionally upon encountering a target individual. Contingent on the relevance of the target for the perceiver's momentary goals, perceivers will direct their attention to individual attributes of the target, thereby moving toward the more thoughtful end of the processing continuum. If the target is judged to be irrelevant to the perceiver's momentary goals, the final impression of the target will be based exclusively on the initial categorization. If, however, the target is judged to be relevant to the perceiver's momentary goals, the perceiver will attempt to integrate person-specific attributes into a coherent impression of the target.

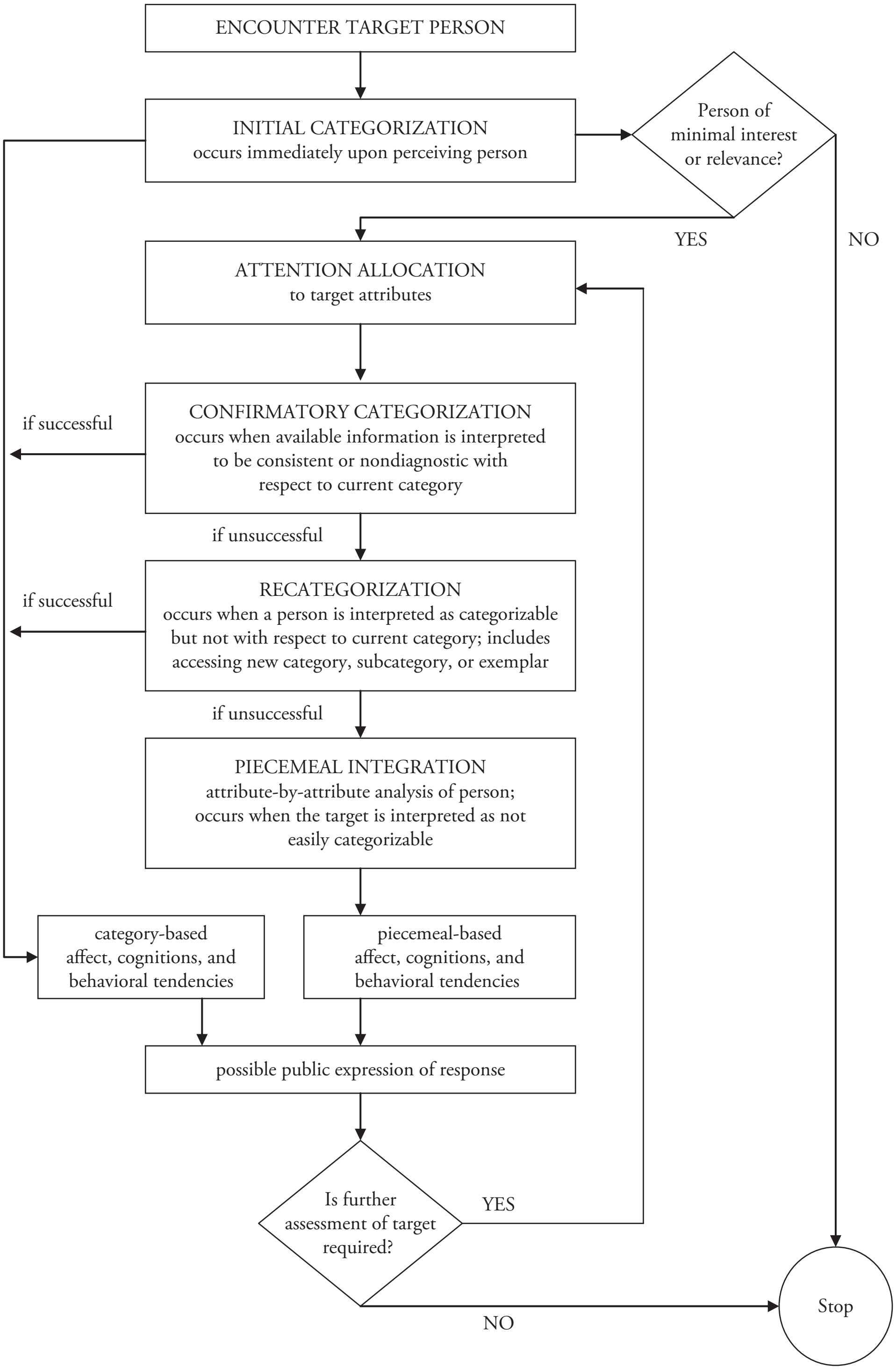

Figure 14.2 The continuum model of impression formation. (Adapted from Fiske & Neuberg, 1990. Reprinted with permission.)

Overall, the continuum model assumes that perceivers attempt to maintain the impression implied by their initial categorization while processing individual attributes of the target. To the extent that the additional information is interpreted to be consistent with the initially identified category, the final impression of the target will be based on the initial categorization. If, however, the additional information is inconsistent with the initial categorization, perceivers will attempt to recategorize the target in an attempt to find a more suitable category than the initial one. For example, if person-specific attributes of a target individual seem inconsistent with the impression implied by his or her category membership, perceivers may use subtypes to assign the target to a more appropriate category than the initial, general category (Richards & Hewstone, 2001). If this recategorization process successfully integrates the available information about the target, the final impression will be based on this newly applied category. However, if the attempt to recategorize the target fails, perceivers are assumed to move on to a process of piecemeal integration, in which they engage in an attribute-by-attribute assessment of individual characteristics of the target. Yet, according to Fiske, Lin, and Neuberg (1999), such piecemeal integration likely occurs quite rarely, given that perceivers tend to construct ad hoc theories to account for contradictory information in the initial stages of the impression formation continuum (e.g., Asch & Zuckier, 1984; Kunda, Miller, & Claire, 1990; Leyens & Yzerbyt, 1992).

DUAL PROCESS MODEL

Whereas Fiske and Neuberg's (1990) continuum model attributes a dominant role to category-based processing, Brewer's (1988) dual process model argues that impression formation may take either a top-down or a bottom-up route (see also Brewer & Harasty-Feinstein, 1999). Both routes are assumed to start with an automatic identification of salient features of the stimulus person (Figure 14.3). This processing step can be described as the mere recognition of feature configurations (e.g., male, dark skin color, business suit). To the extent that the target is irrelevant to the perceiver, the processing sequence is assumed to remain at this level without further processing of category-related or person-related implications of the identified features. If, however, the target is relevant to the perceiver, further processing of the identified features can take either a top-down or a bottom-up route depending on the relative involvement of the perceiver.

Bottom-up processing is assumed to occur under conditions of high involvement, in which perceivers are assumed to adopt an interpersonal orientation. In this person-based processing mode,

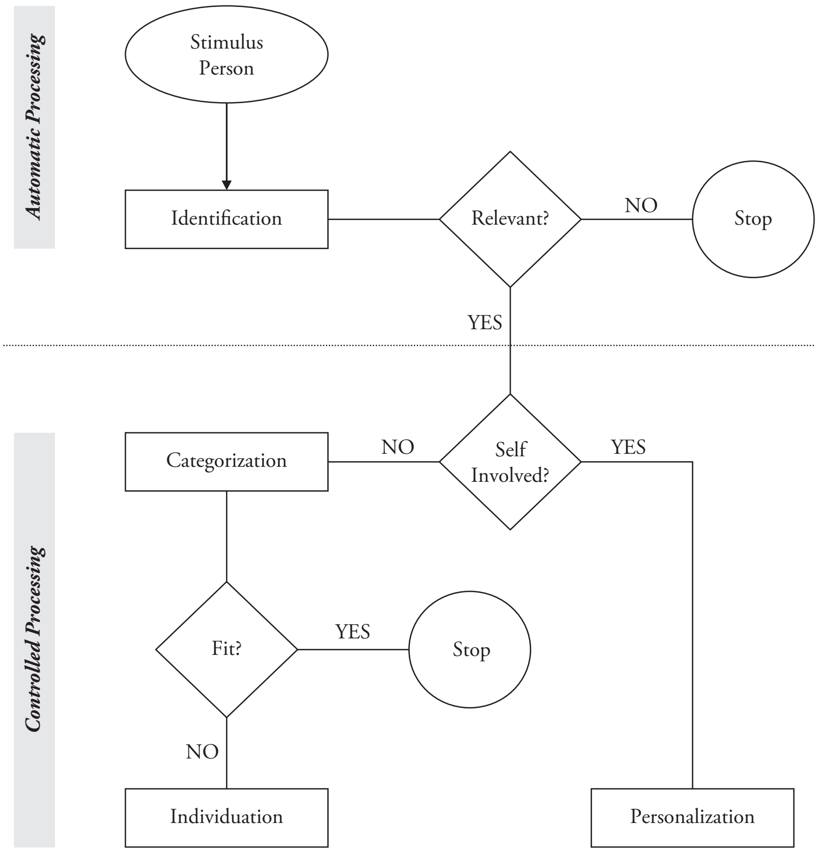

Figure 14.3 The dual-process model of impression formation. (Adapted from Brewer, 1988. Reprinted with permission.)

perceivers are assumed to draw inferences directly from the identified features, which are integrated into a coherent impression of the target (*personalization*). Depending on the motivation and ability to engage in effortful processing, this person-based impression may be more or less complex. In other words, personalization is not an effortful process per se. Rather, the degree of cognitive elaboration is assumed to influence the complexity of the final impression, such that low elaboration will lead to relatively simple person-based impressions, whereas high elaboration will lead to relatively complex person-based impressions.

Top-down processing is assumed to occur under conditions of low involvement, in which perceivers are assumed to adopt an intergroup orientation. In this category-based processing mode, the target is initially categorized on the basis of salient features (e.g., black businessman). This categorization process, in turn, may activate stereotypic contents associated with the applied category, which serve as a filter for the integration of other target-specific information. Whereas target-specific information that is related to the stereotypic content of the category will be integrated into a coherent impression, target-specific information that is unrelated to the category stereotype will be ignored. To the extent that the category-related target information is consistent with the category stereotype, the process is assumed to stop at this point, leading to a target impression in line with the category stereotype (*stereotyping*). If, however, the category-related target information is inconsistent with the category stereotype, the inconsistency has to be resolved in order to achieve a coherent impression of the target. The result of the latter process is an individuated impression of the target, which is based on a systematic integration of target-specific information (*individuation*). However, this integration is still regarded as a category-based process, given that the initial categorization of the target serves as a filter for the processing of category-related versus category-unrelated target information. Thus, like person-based processing, category-based processing can be more or less effortful, such that stereotyping is the likely outcome of low elaboration, whereas individuation usually requires high elaboration.

Dispositional Attribution

Another important question in the context of impression formation is how perceivers make sense of other people's behavior. To describe the processes that underlie inferences from observed behavior, social psychologists in the 1960s proposed various theories of causal (e.g., Kelley, 1967) and dispositional (e.g., Jones & Davis, 1965) attribution (see Chapter 6). However, deviating from the predictions of these models, empirical research soon demonstrated that perceivers tend to give more weight to dispositional compared with situational factors (e.g., Jones & Harris, 1967; Ross, Amabile, & Steinmetz, 1977). This tendency to overestimate the role of dispositional compared with situational factors has become known as the *fundamental attribution error* (Ross, 1977). A particular instantiation of the fundamental attribution error is the *correspondence bias* (Gilbert & Malone, 1995; Jones, 1990), which is defined as the tendency to draw correspondent dispositional inferences from observed behavior even if the behavior is constrained by situational factors (for a discussion of conceptual differences between the fundamental attribution error and the correspondence bias, see Gawronski, 2004). In the 1970s, the discrepancy between theoretically derived predictions and empirical results led to the odd situation that the models that had originally been designed to describe and explain perceivers' inferences acquired a normative status, such that empirically observed deviations were depicted as judgmental biases or errors (see Kruglanski & Ajzen, 1983) instead of counterevidence against the proposed models. This situation did not change until the emergence of dual process theories in the mid-1980s. These theories turned attention back to describing the processes that underlie perceivers' inferences, with a particular focus on explaining when and why the correspondence bias occurs.

THREE-STAGE MODEL

One such dual process theory is Gilbert's (1989) three-stage model of dispositional inference. According to this model, dispositional inferences involve three sequential processes that are claimed to require different amounts of cognitive resources: (1) *behavioral categorization* (i.e., what is the actor doing?), (2) *dispositional characterization* (i.e., what disposition does the behavior imply?), and (3) *situational correction* (i.e., what situational determinants might have caused the behavior?). Whereas behavioral categorization and dispositional characterization are assumed to occur unintentionally without requiring large amounts of cognitive resources, situational correction is assumed to be an intentional, relatively effortful process. Applied to the correspondence bias, these assumptions imply that the tendency to draw correspondent dispositional inferences from situationally constrained behavior should be lower when perceivers have both the motivation and the cognitive capacity to engage in the effortful process of situational correction. However, the tendency to commit the correspondence bias should be enhanced if either the motivation or the cognitive capacity to engage in effortful processing is low. These predictions have been confirmed in several studies that investigated effects of processing motivation (e.g., D'Agostino & Ficher-Kiefer, 1992; Vonk, 1999; Webster, 1993) and cognitive capacity (e.g., Gilbert, Pelham, & Krull, 1988) on dispositional inferences from situationally constrained behavior.

An important extension of Gilbert's (1989) three-stage model was proposed by Krull (1993), who merged Gilbert's (1989) assumptions about the effortfulness of situational correction with previous research on judgmental anchoring in dispositional inference (Quattrone, 1982). Deviating from Gilbert's (1989) assumption that social inferences follow a fixed sequence, Krull (1993) argued that the particular sequence of processes depends on the inferential goal of the perceiver. According to Krull, perceivers interested in inferring an actor's disposition will (1) categorize the behavior, (2) characterize a corresponding disposition, and then (3) correct these characterizations for situational constraints. If, however, perceivers are interested in the causal impact of situational factors, they will (1) categorize behavior, (2) characterize the situation, and then (3) correct these characterizations for dispositional factors. In other words, the contents of both the characterization and the correction stage are assumed to depend on the inferential goal of the perceiver. Because correction processes are assumed to require more capacity compared with characterization processes, motivation and cognitive capacity to engage in effortful processing should have different effects on social inferences as a function of perceivers' inferential goals. Specifically, reduced cognitive elaboration should increase the tendency to commit the correspondence bias when perceivers have the goal of inferring dispositional characteristics of the actor. In contrast, reduced cognitive elaboration should have the opposite effect when perceivers are interested in characteristics of the situation (e.g., Krull, 1993; Krull & Dill, 1996; Krull & Erickson, 1995).

TWO-STAGE MODEL

Another influential model that aims to describe the processes underlying dispositional attributions is Trope's (1986) two-stage model. According to this model, trait judgments are the product of two sequential processes, which are described as *identification* and *inference* (Figure 14.4). At the identification stage, perceivers categorize momentarily available cues in trait-relevant terms. These cues may be related to the actor's behavior (*behavioral cues*), the situational context of the behavior (*situational cues*), or the actor's personal characteristics or group membership (*prior cues*). For example, a person's behavior might be categorized as friendly or hostile, the situational context as facilitating friendly or hostile reactions, and the actor as belonging to a stereotypically friendly or hostile group. To the extent that the relevant cues within each of the three dimensions are unambiguous, they fully constrain

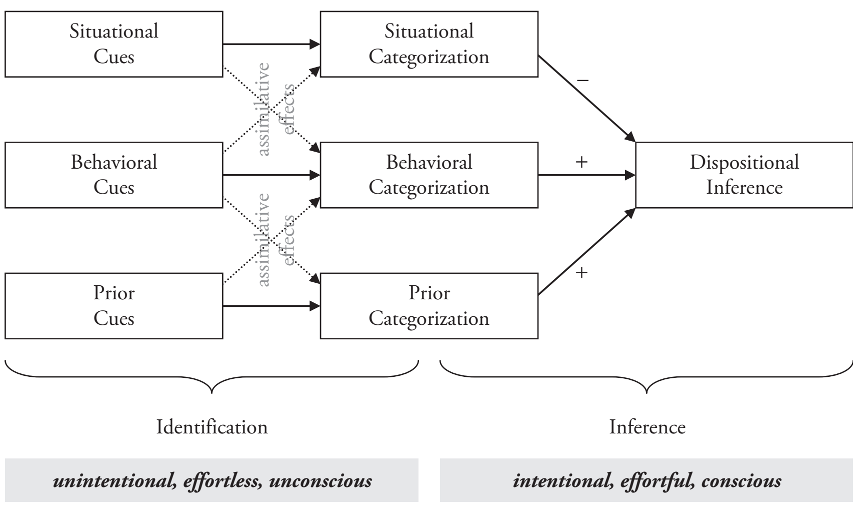

Figure 14.4 The two-stage model of dispositional inference. (Adapted from Trope, 1986. Reprinted with permission.)

their corresponding categorizations. That is, behavioral cues fully determine the categorization of the behavior, situational cues fully determine the categorization of the situation, and prior cues fully determine the categorization of the actor. If, however, a particular cue is ambiguous, its categorization may be biased by contextual cues in an assimilative manner. For example, if an actor's behavior is ambiguously hostile, it may be perceived as more hostile when the situational context is known to facilitate hostile rather than friendly behavior (e.g., Snyder & Frankel, 1976; Trope, Cohen, & Alfieri, 1991; Trope, Cohen, & Maoz, 1988) or when the actor belongs to a stereotypically hostile rather than a stereotypically friendly group (e.g., Duncan, 1976; Gawronski, Geschke, & Banse, 2003; Hugenberg & Bodenhausen, 2003; Kunda & Sherman-Williams, 1993; Sagar & Schofield, 1980). Similar effects may occur for the categorization of ambiguous situational cues (e.g., Trope & Cohen, 1989) and the categorization of ambiguous prior cues (e.g., Hugenberg & Bodenhausen, 2004), both of which may be biased by unambiguous behavioral cues. Because perceivers tend to consider their subjective categorizations "as perceptual givens rather than as context-derived" (Trope & Gaunt, 1999, p. 170), deliberate correction of such biased perceptions is rather unlikely, even when the validity of the biasing contextual information is discredited afterward (e.g., Trope & Alfieri, 1997). In combination with studies showing biasing effects of contextual cues under conditions of depleted cognitive resources (e.g., Trope & Alfieri, 1997; Trope & Gaunt, 2000), these findings suggest that the processes involved at the identification stage operate unintentionally, efficiently, and outside of conscious awareness (for a review, see Trope & Gaunt, 1999).

Once behavioral, situational, and prior cues have been categorized, the outputs of the identification stage serve as inputs for more or less deliberate dispositional inferences (see Figure 14.4). At this stage, perceivers' categorizations of the behavior, the situation, and the actor are integrated into a unified judgment of the actor's disposition. However, in contrast to the generally assimilative nature of contextual influences at the identification stage, dispositional judgments at the inferential stage are influenced by the three kinds of information in different ways. Whereas both behavioral and prior information influence trait judgments in a positive direction, situational information has a subtractive effect. For example, behavior that is categorized as hostile will facilitate correspondent inferences of dispositional hostility. Similarly, categorization of the actor as a member of a stereotypically hostile group will also promote inferences of dispositional hostility. Situational cues identified to provoke hostile reactions, however, should have an inhibitory effect on inferences of dispositional hostility. Such information should discount the informational value of the other two dimensions, and hence reduce correspondent inferences of dispositional hostility (see Kelley, 1972). Importantly, whereas contextual influences at the identification stage are assumed to operate efficiently, unintentionally, and outside of conscious awareness, the integration of the different kinds of information at the inference stage is assumed to be a conscious, intentional process that requires varying amounts of cognitive resources depending on the salience of the three kinds of information (Trope & Gaunt, 2000).

Like Gilbert's (1989) three-stage model, Trope's (1986) two-stage model implies that the tendency to draw correspondent dispositional inferences from situationally constrained behavior should be more pronounced when cognitive elaboration is low than when it is high (but see Trope & Gaunt, 2000, for a qualification of this prediction). However, Trope's model also implies that perceivers may sometimes draw strong correspondent inferences from situationally constrained behavior even when perceivers consider the importance of situational factors at the inference stage. To illustrate this case, consider a study by Snyder and Frankel (1976) in which participants were presented with a videotaped interview of a target behaving somewhat anxiously. The sound of the interview was muted, so that participants focused only on the target's behavior. Half of the participants were told that that the target person was interviewed about an anxiety-provoking topic (e.g., sexual fantasies), whereas the remaining half were told that the topic was not anxiety provoking (e.g., favorite books). Results showed that participants in both conditions inferred equal levels of dispositional anxiety from the observed behavior. At first glance, this result may suggest that participants in the anxious topic condition ignored the anxiety-provoking topic as a situational factor for the observed behavior when they made their attributions of dispositional anxiety. However, Snyder and Frankel's (1976) data also showed that participants perceived the target's behavior as more anxious when they believed that she was interviewed about sexual fantasies than when they believed that she was interviewed about favorite books. Given that higher levels of perceived *behavioral* anxiety should promote corresponding attributions of *dispositional* anxiety, the biasing effect of situational information on the identification of the target's behavior may have concealed the subtractive influence of situational information at the inference stage. In other words, participants perceived the target's behavior as more anxious when the interview topic was anxiety provoking, but then they considered the anxiety-provoking nature of the interview topic as a situational explanation for this behavior when they drew correspondent dispositional inferences. These considerations imply that the presumed interview topic may influence dispositional attributions in Snyder and Frankel's (1976) paradigm in two different ways. First, the interview topic may have an *indirect positive effect* on attributions of dispositional anxiety that is meditated by the biased identification of the target's behavior. Second, the interview topic may have a *direct negative effect* if it is considered a situational explanation for the observed behavior (e.g., Trope & Alfieri, 1997). Because the proposed indirect effect of situational information is opposite to the proposed direct effect, the two effects may sometimes cancel each other out on measures of trait attribution unless the proposed indirect effect is statistically controlled (e.g., Gawronski et al., 2002). Importantly, this pattern of simultaneous, yet antagonistic, direct and indirect effects implies that perceivers may draw strong correspondent inferences from situationally constrained behavior, even if they consider situational factors to have a strong impact on the target's behavior. In other words, they commit the correspondence bias, even though they do not commit the fundamental attribution error (Gawronski, 2004).

Criticism of Phenomenon-Specific Dual Process Theories

Overall, the reviewed phenomenon-specific dual process theories have gained strong empirical support, and many of them have been seminal in shaping the field of social psychology over the past decades. Nevertheless, these theories have also been the target of criticism. One of the most important arguments raised against these models is that they equate distinct contents with distinct processes (e.g., Kruglanski, Pierro, Mannetti, Erb, & Chun, 2007). For example, Kruglanski and Thompson (1999) argued that dual process theories of persuasion tend to conflate different types of evidence (e.g., message arguments, characteristics of the source) with different types of processes (e.g., peripheral vs. central; heuristic vs. systematic), even though the processes by which the two kinds of information are integrated into attitudinal judgments may be the same (e.g., Erb, Pierro, Mannetti, Spiegel, & Kruglanski, 2007). According to this criticism, the two kinds of information may simply differ in their levels of complexity, thereby requiring different amounts of processing resources. Thus, if the complexity of the two kinds of information is controlled or independently manipulated, message arguments may influence attitudes even under conditions of low elaboration (e.g., when their complexity is low), or source characteristics may influence attitudes only when cognitive elaboration is high (e.g., when their complexity is high).

Similar arguments have been raised against dual process theories of dispositional attribution (e.g., Chun, Spiegel, & Kruglanski, 2002) and dual process theories of impression formation (e.g., Chun & Kruglanski, 2006). For example, challenging Trope's (1986) assumption that the biasing influence of contextual cues in the categorization of ambiguous behavior is an efficient process, Chun and colleagues (2002) have shown that the effects of assimilative behavior identification are attenuated by cognitive load when the categorization of ambiguous behavior is relatively difficult. Along the same lines, Chun and Kruglanski (2006) showed that individuating information has a stronger impact compared with category information under conditions of resource depletion if the complexity of the individuating information is lower compared with the complexity of the category information. This finding challenges a central assumption of Fiske and Neuberg's (1990) continuum model, which implies that category information enjoys general priority. As an alternative to phenomenon-specific dual process theories, Kruglanski and colleagues (2007) proposed a unimodel of human judgment, which proposes that different kinds of evidence (e.g., message arguments vs. source characteristics; category vs. individuating information) are integrated in a single inferential process whose outcome depends on several judgmental parameters, including information relevance, processing difficulty, processing motivation, processing capacity, and processing order.

A related concern about phenomenon-specific dual process theories is that they all seem to address very similar issues. Yet, it remains unclear how the different models are related to each other and what general principles of information processing underlie the studied phenomena. This concern echoes criticism by Kruglanski and colleagues'

(2007) unimodel, according to which different kinds of information are integrated in a single epistemic process that is influenced by several processing parameters. An alternative response is implied by generalized dual process theories, which identify two general processing principles that are assumed to provide the foundation for the reviewed domain-specific phenomena.

Generalized Dual Process Theories

Generalized dual process theories aim at identifying domain-independent principles of social information processing that underlie various kinds of phenomena. Most of these theories fall into the category of dual system theories, in that they propose two processing systems that operate on the basis of qualitatively distinct principles. In the following sections, we first review the core assumptions of the most prominent examples, and then discuss criticism that has been raised against generalized dual process theories.

Cognitive-Experiential Self-Theory

The groundwork for generalized dual process theories was put in place by Epstein's (1994) cognitive-experiential self-theory (CEST), which is based on his foundational work on the nature of the self-concept as a global theory of personality (Epstein, 1973). In general terms, CEST proposes two interacting systems that are characterized by different processing principles. The first is described as the *experiential system*; the second is described as the *rational system* (for an overview, see Epstein & Pacini, 1999).

The experiential system is assumed to operate in an automatic, effortless manner on the basis of associative connections that are closely linked to affective principles of pleasure and pain (i.e., what feels good or bad). Encoding of reality in the experiential system is claimed to occur in concrete images, metaphors, and narratives, involving holistic responses that are oriented toward immediate action. As such, responses driven by the experiential system are characterized by broad, schematic generalizations that tend to be incoherent, crudely integrated, and context specific. Changes in the experiential system are assumed to occur slowly, requiring repetitive or relatively intense experiences. These processing principles are assumed to be rooted in brain structures that developed early in evolution and that have not been replaced by more recently evolved structures that build the foundation for the second, rational system.

The rational system is characterized by intentional, effortful processing that is based on logical relations between elements (i.e., what is rational). Encoding of reality in the rational system is claimed to occur in abstract symbols, words, and numbers, involving analytic responses that are oriented toward delayed action. Thus, responses driven by the rational system are characterized by differentiated, highly integrated representations that tend to be abstract, logically coherent, and context independent. Changes in the rational system are assumed to occur more quickly compared with the experiential system, with changes depending on argument strength and availability of new evidence.

According to Epstein (1994), the two systems operate in parallel, such that each system can independently produce its own response tendency. In cases in which these response tendencies are incongruent, people tend to experience a "conflict between the head and the heart" (p. 710), such that the experiential system may produce an intuitive, affective response tendency that conflicts with a rational, logical response tendency produced by the rational system. At the same time, the two systems may interact with each other, such that preconscious processes in the experiential system may continuously influence conscious processing in the rational system. However, the proposed interaction between the two systems is assumed to be asymmetrical because influences from the experiential system usually remain outside of conscious awareness. As a result, these influences often remain uncontrolled by the rational system because there is no awareness that there is anything to control to begin with. Still, there can be individual and situational variations in the relative dominance of the two systems. For example, Epstein, Pacini, Denes-Raj, and Heier (1996) developed the Rational-Experiential Inventory (REI), which includes two individual difference measures that are specifically designed to identify stable individual differences in the dominance of intuitive-experiential and analytical-rational thinking styles. Other moderating factors include situational circumstances and emotional arousal. Whereas circumstances that require a formal analysis of the current situation are assumed to give priority to the rational system, emotional arousal is assumed to shift the balance toward the experiential system.

Although the role of CEST in empirical research has mostly taken the form of a conceptual framework for the interpretation of results rather than a source for the deduction of testable predictions, it clearly deserves the credit of setting the groundwork for generalized dual process theorizing. In fact, many of its core assumptions can be found in its successors, some of which have been more successful in generating novel predictions through the specification of earlier claims and the inclusion of new propositions. During the past few years, these theories have gained more impact compared with CEST. Yet, as the first theory of this kind, CEST still enjoys the status of being the most frequently cited generalized dual process theory.

Associative Versus Ruled-Based Processing

Another milestone in the development of generalized dual process theories is Smith and DeCoster's (2000) conceptual integration of various domain-specific dual process theories. Drawing on Sloman's (1996) distinction between associative and rule-based processing and McClelland and colleagues' work on fast-learning versus slow-learning memory systems (McClelland, McNaughton, & O'Reilly, 1995), Smith and DeCoster (2000) argued that the phenomena identified and studied by domain-specific dual process theories reflect the operation of two distinct memory systems that are guided by different processing principles: a slow-learning system that is characterized by associative processing and a fast-learning system that is characterized by rule-based processing. Associative processing is further specified as being structured by similarity and contiguity, drawing on simple associations between objects and events that are learned slowly over many experiences. Associative processing is assumed to occur automatically in a parallel fashion without awareness of the involved processing steps, even though their output may be accessible to conscious awareness. Rule-based processing, in contrast, is characterized as being structured by language and logic, drawing on symbolically represented rules that can be learned quickly with very few experiences. Attributing a dominant role to associative processing, rule-based processing is further assumed to occur optionally in a sequential fashion if both the motivation and the capacity to engage in effortful processing are present. Its processing steps are often accessible to conscious awareness, such that the applied rules of inference can be verbalized. Similar to Epstein's (1994) CEST, Smith and DeCoster (2000) propose an asymmetrical interaction between the two memory systems, such that rule-based processing may draw on inputs from both memory systems, whereas associative processing is exclusively based on the slow-learning system.

Using the distinction between associative and rule-based processing in the two memory systems, Smith and DeCoster (2000) integrated the phenomena identified and studied by various domain-specific dual process theories within a single generalized framework. For example, peripheral/heuristic processing in dual process theories of persuasion (Chaiken, 1987; Petty & Cacioppo, 1986) is characterized as the use of well-learned associations of salient cues (e.g., source attractiveness) with positive or negative evaluations. Central/systematic processing, in contrast, is described as the effortful search for relevant information that is evaluated using rule-based processes on the basis of logical principles. Along the same lines, automatic attitude activation in Fazio's (1990) MODE model is described as automatic access to evaluations that are associated with an attitude object through repeated pairings. Deliberate analysis of an object's attributes, in contrast, is characterized as the search for and appraisal of relevant information on the basis of logical rules of inference. Correspondent dispositional inferences in Gilbert's (1989) three-stage model are described as the use of traits that are semantically associated with a person's observed behavior, whereas inferences about alternative causes (e.g., situational factors) are assumed to involve rule-based processes that are guided by principles of logical inference. Similar considerations apply to Devine's (1989) dissociation model of prejudice and stereotyping, such that automatic stereotype activation is assumed to be the result of highly overlearned associations between social groups and stereotypic information, whereas suppression of these automatically activated stereotypes involves effortful access to personal beliefs in order to override the impact of stereotypic information. Finally, automatic categorization in dual process theories of impression formation (Brewer, 1988, Fiske & Neuberg, 1990) is described as the use of information that is associated with a person's salient category (e.g., gender, race, age), whereas individuation involves the processing and appraisal of multiple individual characteristics to form a personal impression.

System 1 and System 2 Processing

Working toward a theoretical integration of earlier research on heuristics and biases (for reviews, see Gilovich, Griffin, & Kahneman, 2002; Kahneman, Slovic, & Tversky, 1982), Kahneman (2003) presented a generalized dual process theory that shares many features with Smith and DeCoster's (2000) and Epstein's (1994) models. To this end, Kahneman (2003) distinguished between two systems, generically described as *System 1* and *System 2* (cf. Stanovich & West, 2000), that are assumed to underlie intuition versus reasoning. Sharing characteristics of basic perceptual processes, intuitive processing in System 1 is described as fast, parallel, automatic, effortless, associative, slow learning, and emotional. In contrast, reasoning processes in System 2 are described as slow, serial, controlled, effortful, rule governed, fast learning, and emotionally neutral. At the same time, information processing in the two systems is assumed to differ from basic perceptual processes, in that both intuition and reasoning can be evoked by verbal information, involving conceptual representations of the past, the present, and the future. These features differ from basic perceptual processes, which involve stimulus-bound percepts that are driven by current stimulation. Thus, whereas the outputs of System 1 may be described as intuitive *impressions*, the outputs of System 2 are *judgments* that can be based on impressions or on deliberate reasoning. In that sense, an important function of System 2 is to monitor the activities and inputs of System 1. If no intuitive response is generated by System 1, judgments and behavior are exclusively computed by System 2. If, however, System 1 provides an intuitive response as input for System 2, System 2 may either (1) endorse this response, (2) adjust the response for other features that are recognized to be relevant, (3) correct the response for a recognized bias, or (4) block the response from overt expression if it is identified to violate a valid rule of inference.

Whereas the intuitive responses generated by System 1 are determined by the accessibility of mental contents (Higgins, 1996), processing in System 2 is guided by the application of logical rules of inference. In the case of heuristic judgments, highly accessible contents in System 1 will often pass the monitoring activities of System 2 through a process of attribute substitution. In general terms, attribute substitution occurs when an "individual assesses a specified target attribute of a judgment object by substituting a related heuristic attribute that comes more readily to mind" (Kahneman, 2003, p. 707). This process can be illustrated with the well-known Linda problem that has been used to demonstrate the conjunction fallacy (Tversky & Kahneman, 1983, p. 297): "*Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations*." According to Tversky and Kahneman (1983), people commit the conjunction fallacy when they judge Linda as more likely to be (1) a bank teller and active in the feminist movement than (2) a bank teller. Because the conjunction of two distinct events can never be more likely than one of the two events by itself, such a judgment violates basic principles of statistical probability. Drawing on Kahneman's (2003) theoretical conceptualization, this judgmental tendency can be explained by accessibility-driven attribute substitution, in that individuals substitute a relevant judgmental attribute (i.e., statistical probability) with an irrelevant, yet highly accessible attribute (i.e., feature similarity). Other examples of accessibility-driven attribute substitution include framing effects, in which people tend to show a preference for positively over negatively framed objects, events, and decisions (e.g., a sausage that is described as 80% lean vs. 20% fat), even if the two descriptions are semantically equivalent.

Although Kahneman (2003) adopted the System 1/System 2 distinction from Stanovich and West's (2000) influential work on rationality in human reasoning, his reanalysis of key findings of the heuristics and biases research program contributed significantly to the current dominance of this distinction in the literature on judgment and decision making. This impact went even beyond the traditional boundaries of psychology, in that philosophers have started to use it as a basis for speculations about the general architecture of the human mind (e.g., Carruthers, 2009; Samuels, 2009). Yet, similar to Epstein's (1994) CEST, the theory has been criticized for providing no more than a conceptual framework that can be applied to empirical data in a post hoc fashion without generating novel predictions that could be empirically confirmed or disconfirmed (Keren & Schul, 2009). We will return to this concern in our overarching discussion of criticism against generalized dual process theories.

Reflection–Reflexion Model

The reflection–reflexion model proposed by Lieberman (2003; see also Lieberman, Gaunt, Gilbert, & Trope, 2002) combines the basic idea of the dual system approach with recent advances in social cognitive neuroscience (see Chapter 34). Deviating from conceptualizations that describe automatic processes as more efficient variants of insufficiently practiced controlled processes (e.g., Bargh, 1997), the reflection–reflexion model argues that automatic and controlled processes use qualitatively distinct representations that have their basis in distinct neural structures. According to Lieberman (2003), the automatic–controlled distinction is insufficient, if not misleading, because it simply describes the *operating conditions* of a given process (i.e., when does the process operate?) without specifying the underlying *computational properties* (i.e., what is the process doing?). Because the reflection–reflexion model assumes distinct representational underpinnings, it also allows for complex interactions between automatic and controlled processes. Such interactions are difficult to reconcile with the view that automatic and controlled processes draw on the same underlying mental representations, the only difference being that automatic processes operate more efficiently as a result of practice.

The first system, called the *X-system* with reference to the term *reflexive*, is proposed to involve the amygdala, basal ganglia, and lateral temporal cortex. Reflexive processes in the X-system link affect and meaning to currently represented stimuli by means of simple stimulus-stimulus associations (semantic meaning) or stimulus-outcome associations (affective meaning). These associations build the foundation for a person's implicit theories and generalized expectations about the world, which are assumed to develop slowly over time and over extended periods of learning. The neurons in the X-system are highly interdependent, in that they are mutually influenced by the neurons they are influencing. As a result, the activation of associations in the X-system operates in a parallel fashion on the basis of similarity and pattern matching, such that observed relations of the format "if *p*, then *q*" will create a reflexive tendency to draw the logically incorrect inference "if *q*, then *p*."

The second system, called the *C-system* with reference to the term *reflective*, is proposed to involve the anterior cingulate cortex, prefrontal cortex, and medial temporal lobe. The operation of the C-system is assumed to be conditional on the failure of the X-system to achieve a momentary goal, such that reflective processes are initiated only if (1) the implicit theories or generalized expectancies in the X-system are violated, or (2) there are no implicit theories or generalized expectancies in the X-system that are applicable to guide behavior in a novel situation. As such, the primary function of the C-system is to handle context-specific "exceptions to the rule" for which the generalizations in the X-system are not prepared. Reflective inferences in the C-system are assumed to operate in a sequential manner on the basis of causal and logical relations, which allows the C-system to block logically incorrect inferences of *p* in the presence of *q* from observed relations of the format "if *p*, then *q*."

Arguably, the most significant contribution of the reflection–reflexion model is the identification of distinct neural underpinnings of automatic and controlled processes. Thus, it does not seem surprising that the theory has enjoyed its strongest influence in the area of social cognitive neuroscience (for a review, see Lieberman, 2007). Yet, in traditional social psychology, its impact has mostly taken the form of citations in sample lists of generalized dual process theories, with few examples of studies that have tested novel predictions derived from the theory.

Reflective–Impulsive Model

One the most influential dual system theories to date is Strack and Deutsch's (2004) reflective–impulsive model (RIM) of social behavior. The RIM argues that social behavior is guided by two simultaneously operating systems of information processing, which are described as the *reflective system* (RS) and the *impulsive system* (IS), respectively. Even though both systems are assumed to operate in parallel, the IS enjoys priority over the RS because the RS operates only under conditions of sufficient cognitive capacity, whereas information processing in the IS is assumed to be resource independent. Similar to other dual system theories, the RIM states that the IS operates on the basis of simple associative links between elements that are formed and activated according to the principles of similarity and contiguity. Information processing in the RS, in contrast, is assumed to involve propositionally represented relations between elements, which are tagged with truth values (i.e., true vs. false). These operating characteristics make the RS capable of various operations that cannot be performed by the IS, the most important of which are the processing of negations and representations of the future. Thus, even though accessible associations in the IS provide the basis for propositional representations in the RS, their functionally distinct operating principles can have different behavioral implications if processing in the RS involves the negation of activated associations in the IS (see Deutsch, Gawronski, & Strack, 2006; Gilbert, 1991) or the delay of gratification (see Metcalfe & Mischel, 1999).

Another central assumption of the RIM concerns the translation of mental representations into behavior. The RIM assumes that the RS and the IS influence behavior through a common pathway that includes the activation of behavioral schemata of varying abstractness (Norman & Shallice, 1986). These behavioral schemata can be activated directly through the spread of activation from momentarily accessible associations in the IS, which may elicit an impulsive tendency to either approach or avoid a given object. Alternatively, behavioral schemata may be activated indirectly through behavioral intentions generated in the RS, which are guided by (1) the subjective hedonic quality of future states that may result from a given behavior (i.e., value) and (2) the subjective probability with which the behavior may produce the focal outcome (i.e., expectancy).

Going beyond other dual process models that focus primarily on cognitive and affective processes, the RIM attributes an important role to motivational processes, which may operate in the IS in at least two different ways. First, the RIM assumes that motivational orientations to approach or avoid an object may be elicited by mere processing of positive or negative information, mere perception of approach or avoidance movements, the experience of positive or negative affect, or the execution of approach or avoidance motor actions. Conversely, motivational orientations of approach or avoidance are assumed to facilitate the processing of information, the experience of affective states, and the execution of behavior that are compatible with the current motivational state (see Neumann, Förster, & Strack, 2003). Second, the RIM integrates basic principles of homoeostatic dysregulation, such that deprivation of basic needs is assumed to activate behavioral schemata that are linked to successful satisfaction of those needs through a history of past experiences.

To date, the RIM enjoys the status of being the most influential dual system theory in the generation of empirical research. This research includes a wide range of topics within and beyond social psychology, such as the roles of impulse and control in self-regulation (e.g., Hofmann, Rauch, & Gawronski, 2007), limits in the processing of negations (e.g., Deutsch et al., 2006), the relation between personality traits and behavior (e.g., Back, Schmukle, & Egloff, 2009), and emotional effects of food deprivation (e.g., Hoefling et al., 2009).

Associative-Propositional Evaluation Model

Even though Gawronski and Bodenhausen's (2006a, 2006b, 2007, 2011) associative–propositional evaluation (APE) model was originally designed to resolve various inconsistencies in the literature on implicit and explicit attitude change, its emphasis on general principles of information processing make it more similar to generalized dual process theories than phenomenon-specific dual process theories of persuasion. The theoretical core of the APE model is the distinction between associative and propositional processes. Associative processes are defined as the *activation* of associations in memory, which is driven by principles of similarity and contiguity. Propositional processes are defined as the *validation* of the information implied by activated associations, which is assumed to be guided by principles of logical consistency. To the extent that the propositional information implied by activated associations is consistent, it will be used for judgments and behavioral decisions. If, however, the information implied by activated associations is inconsistent, aversive feelings of dissonance will induce a tendency to resolve the dissonance-provoking inconsistency before a judgment or behavioral decision is made (see Festinger, 1957). In general terms, inconsistency may be resolved either by rejecting one of the propositions that are involved in inconsistent belief systems or by searching for an additional proposition that resolves the inconsistency (Gawronski & Strack, 2004). To the extent that the inconsistency is resolved by rejecting one of the involved propositions, activated associations and endorsed propositional beliefs will show a dissociation because mere negation of a proposition (e.g., "it is not true that old people are bad drivers") does not necessarily deactivate the mental associations that underlie that proposition (i.e., the association between the concepts *old people* and *bad drivers*). The prototypical example of such cases are dissociations between explicit and implicit measures, given that explicit self-report measures assess the outcome of propositional validation processes (e.g., survey questions asking participants to report their agreement or disagreement with a particular statement), whereas implicit measures (e.g., implicit association test, sequential priming tasks) provide a proxy for the activation of associations in memory. Even though the original formulation of the APE model was primarily concerned with the role of associative and propositional processes in evaluation (Gawronski & Bodenhausen, 2006a, 2006b), its basic principles are equally applicable to nonevaluative, semantic information (Gawronski, LeBel, & Peters, 2007; for an example, see Peters & Gawronski, 2011).

By emphasizing the ubiquitous roles of associative and propositional processes in social information processing, the APE model has a strong resemblance to other generalized dual process theories. However, the APE model also includes a number of assumptions that distinguish it from these models. First, whereas most generalized dual process theories propose the existence of two systems with distinct operating characteristics (e.g., Epstein, 1994; Kahneman, 2003; Lieberman, 2003; Smith & DeCoster, 2000; Strack & Deutsch, 2004), the APE model does not make any assumptions about inherent links between processes and systems (Gawronski & Bodenhausen, 2011). In that sense, the APE model can be described as a *dual process theory* in the original sense of the term, whereas most other generalized dual process theories represent examples of *dual system theories*. Second, the APE model does not make any claims implying that one process is fast learning whereas the other is slow learning. To the contrary, the APE model argues that either the outcome of associative processes or the outcome of propositional processes can be more or less robust against external influences depending on (1) the nature of the external influence, and (2) mutual interactions between the two processes (Gawronski & Bodenhausen, 2006a, 2007, 2011). Third, the APE model makes specific assumptions about how associative and propositional processes interact with each other. These assumptions imply a number of novel predictions about the conditions under which a given factor should produce (1) changes in explicit but not implicit evaluations (e.g., Gawronski & Strack, 2004; Gregg, Seibt, & Banaji, 2006); (2) changes in implicit but not explicit evaluations (e.g., Gawronski & LeBel, 2008; Grumm, Nestler, & von Collani, 2009); (3) corresponding changes in explicit and implicit evaluations, with changes in implicit evaluations being mediated by changes in explicit evaluations (e.g., Peters & Gawronski, 2011; Whitfield & Jordan, 2009); and (4) corresponding changes in explicit and implicit evaluations, with changes in explicit evaluations being mediated by changes in implicit evaluations (e.g., Gawronski & LeBel, 2008; Whitfield & Jordan, 2009). Finally, the APE model draws a sharp line between operating principles (i.e., associative vs. propositional) and operating conditions (i.e., automatic vs. controlled) of mental processes, in that operating principles represent definitions of what a given process is doing, whereas operating conditions represent empirical assumptions about the conditions under which the process is operating. Importantly, the model assumes that there is no perfect overlap between operating principles and operating conditions, such that both associative and propositional processes can have features of automatic or controlled processes (Gawronski & Bodenhausen, 2007, 2009, 2011). For example, even though the activation of associations may occur unintentionally, associations may also be activated intentionally. In a similar vein, complex propositional inferences may require cognitive effort, but propositional inferences may occur quickly and effortlessly if the complexity of these inferences is low (see Gawronski & Bodenhausen, 2007, 2011).

Criticism of Generalized Dual Process Theories

During the past decade, generalized dual process theories have been extremely influential in guiding interpretations of empirical data far beyond the boundaries of social psychology. In fact, their steadily increasing impact seems to be associated with a decreasing influence of their phenomenon-specific precursors, some of which have already acquired the status of "historical milestones" that have ceased to inspire novel research. However, despite their overwhelming influence, generalized dual process theories have also been the target of criticism.

One argument that has been raised against the distinction between associative and rule-based processes is that information processing can be parsimoniously described in terms of general *if–then* rules (e.g., Kruglanski & Dechesne, 2006; Kruglanski, Erb, Pierro, Mannetti, & Chun, 2006). Because this description can be applied to both associative and rule-based processes, the proposed distinctions between qualitatively distinct processes is claimed to be misleading. In an abstract sense, all inference processes could be described as rule-based in terms of general *if–then* rules. According to Kruglanski and colleagues, such rule-based inferences may be influenced by various judgmental parameters, such as subjective relevance, task demands, cognitive capacity, accessibility, and different kinds of motivations. Thus, data patterns that have been interpreted as reflecting the operation of qualitatively distinct processes may in fact reflect the operation of a single rule-based process that is modulated by the proposed processing parameters.

In response to this criticism, however, one could object that the proposed interpretation in terms of general *if–then* rules seems too abstract to specify how exactly rule-based processes are operating. In fact, claiming that all psychological processes follow general *if–then* rules does not go far beyond stating that all psychological processes follow identifiable regularities instead of being completely random. Because all of the reviewed models agree that both associative and rule-based processes follow identifiable regularities and that neither of them is random, one could argue that the single-process criticism misconstrues the basic distinctions of generalized dual process theories. According to these theories, associative processes are guided by similarity and contiguity, whereas rule-based processes are guided by syllogistic relations and logical consistency. Moreover, by introducing several parameters that presumably modulate a single rule-based process, the proposed alternative implicitly assumes that the operation of these parameters does not require further specification—an assumption that is rejected by generalized dual process theories. For example, stating that rule-based inferences are influenced by subjective relevance does not clarify how relevance is determined in the first place. Similarly, stating that accessibility influences what information is considered in rule-based inferences does not clarify why some information tends to be more accessible than other information. From a dual process perspective, one could argue that the process of determining the subjective relevance of accessible information is functionally equivalent to the process of propositional validation, whereas accessibility is determined by the similarity matching principles of associative processes.

Another, more fundamental criticism is that generalized dual process theories have a tendency to create lists of process characteristics that may not necessarily overlap (Keren & Schul, 2009). This criticism has been raised in particular against dual system theories, which describe the proposed systems by means of several features that may not covary. For example, describing one system as affective, automatic, holistic, and associative and the other one as cognitive, controlled, analytic, and logical (e.g., Epstein & Pacini, 1999) raises questions about whether cognitive processes cannot be automatic, holistic, or associative. This criticism resembles earlier concerns about the conflation of different features of automaticity, stating that a given process rarely meets all of the four criteria (i.e., unintentional, efficient, uncontrollable, unconscious; see Bargh, 1994).

Related to this argument, generalized dual process theories have been criticized for lacking conceptual precision, which makes it difficult to identify empirical evidence that could disprove them. In fact, it has been argued that these theories provide no more than nominal lists of definitions instead of empirically testable predictions. For example, Keren and Schul (2009) raised concerns that dual system theories in particular can accommodate virtually every empirical finding in a post hoc fashion. At the same time, they do not provide precise predictions that could disconfirm them, nor do they include specific claims that could empirically distinguish them from other dual system theories. Even though these concerns seem applicable to several of the reviewed models, it is important to note that they also apply to the single process models that have been proposed as superior alternatives. Of course, this does not invalidate the criticism of imprecision and lack of testable predictions. To the contrary, the fact that theories with a high degree of *generality* often lack the level of *specificity* that is required for the derivation of testable predictions (Quine & Ullian, 1978) has important implications for theory construction in social psychology. We will return to this issue in the final section of this chapter.

Formalized Dual Process Theories

Simultaneous to the emergence of generalized dual process theories, social cognition researchers realized that many behaviors that had been presumed to reflect automatic processes are not process pure, but instead conflate the joint contributions of automatic and controlled processes. In addition, many researchers became concerned about method-related confounds between processes and tasks, for example, when automatic processes are assessed with one type of task (e.g., implicit measures) and controlled processes with another (e.g., explicit measures). These concerns have led researchers to adopt mathematical modeling procedures from cognitive psychology to quantify the joint contributions of automatic and controlled processes to behavioral responses within a single task.

Control-Dominating Process Dissociation Model

The most prominent of these modeling techniques is Jacoby's (1991) process dissociation (PD) model. The basic idea of PD models is that automatic and controlled processes sometimes work in concert to produce a behavioral response, while at other times automatic and controlled processes work in opposition to each other (for reviews, see Payne, 2008; Payne & Bishara, 2009). For example, many implicit measures—such as the implicit association test (Greenwald et al., 1998) or evaluative priming (Fazio et al., 1995)—involve one class of trials that is described as *compatible* and another class of trials that is described as *incompatible* (see Gawronski, Deutsch, LeBel, & Peters, 2008). The basic idea is that both automatic and controlled processes will lead to the correct response on compatible trials (thereby facilitating fast and accurate responses). On incompatible trials, however, only controlled processes will lead to the correct response, whereas automatic processes will lead to the incorrect response (thereby inhibiting fast and accurate responses). For example, in an implicit association test designed to measure automatic racial bias (Greenwald et al., 1998), controlled identification of a black or white face will produce an accurate key response in both the compatible and the incompatible block. In contrast, automatic racial bias will produce an accurate key response in the compatible block, but an incorrect response in the incompatible block.

These influences can be depicted graphically as processing trees that describe how automatic and controlled processes may interactively determine correct versus incorrect responses on a given task. In Jacoby's (1991) PD model, controlled processes are assumed to dominate, in that automatic processes influence the behavioral outcome only if controlled processes fail (Figure 14.5a). If controlled processes succeed (depicted as *C* in Figure 14.5a), participants will show the correct response on both compatible and incompatible trials. If,

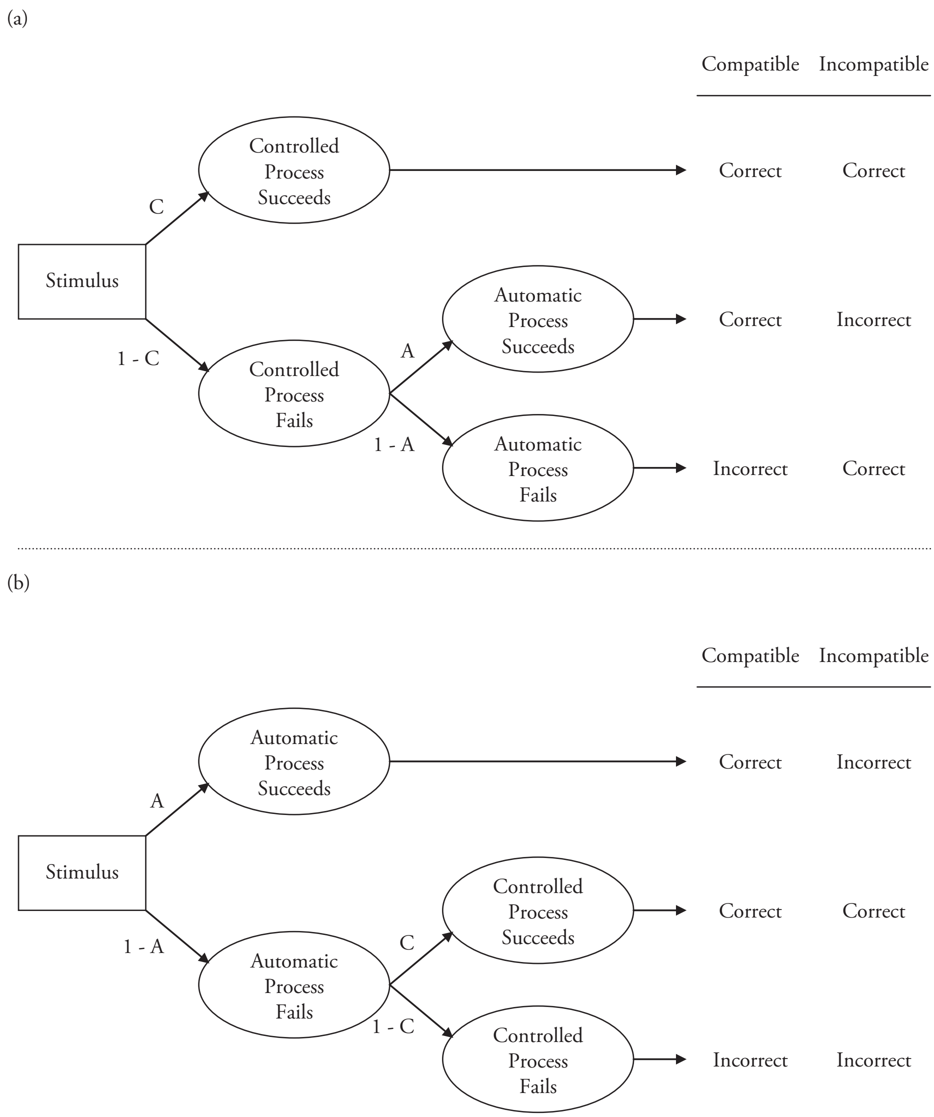

Figure 14.5 Two variants of process dissociation models of automatic and controlled processing. Panel (a) depicts a model in which controlled processes are assumed to dominate; panel (b) depicts a model in which automatic processes are assumed to dominate. (Adapted from Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005. Reprinted with permission.)

however, controlled processes fail (depicted as *1 – C* in Figure 14.5a), the behavioral outcome depends on the operation of automatic processes. If automatic processes drive the response (depicted as *A* in Figure 14.5a), participants will show the correct response on compatible trials but the incorrect response on incompatible trials. Yet, if automatic processes fail (depicted as *1 – A* in Figure 14.5a), the model assumes a bias in the opposite direction, such that participants will show the incorrect response on compatible trials but the correct response on incompatible trials.

The advantage of formalized dual process models is that they provide mathematical equations that can be used to quantify the relative contributions of distinct processes to performance on tasks in which automatic and controlled processes can work in concert or in opposition. The basic idea is to derive two equations from the proposed structure in the processing tree, one for the probability of correct responses on compatible trials and one for the probability of correct responses on incompatible trials. In Jacoby's (1991) PD model, these equations include two parameters as unknowns: *C,* which is supposed to capture the impact of controlled processes; and *A,* which is supposed to capture the impact of automatic processes. Using the empirically observed probabilities of correct responses on compatible and incompatible trials in a given data set, the particular values of these unknowns can be calculated through simple algebra.

For example, using the graphical depiction of Jacoby's (1991) model in Figure 14.5a, the probability of a correct response on compatible trials should be equal to all processing paths from left to right that lead to a correct response in the "compatible" column. The two paths that produce such a response are *Controlled Process Succeeds*, which can be depicted as *C*, and *Controlled Process Fails* in conjunction with *Automatic Process Succeeds*, which can be depicted as *(1 – C) × A*. Thus, in statistical terms, the probability of a correct response on compatible trials can be described as:

$$p(\text{correct} \mid \text{compatible}) = C + (1 - C) \times A$$

The same logic can be applied to the probability of a correct response on incompatible trials. The two paths that produce such a response are *Controlled Process Succeeds*, which is again depicted as *C*, and *Controlled Process Fails* in conjunction with *Automatic Process Fails*, which is depicted as *(1 – C) × (1 – A)*. On the basis of these processing paths, the probability of a correct response on incompatible trials can be described as:

$$p(\text{correct} \mid \text{incompatible}) = C + (1 - C) \times (1 - A)$$

Through the use of linear algebra, these equations can be solved for *C* and *A*, which allows researchers to quantify the relative impact of automatic and controlled processes. Without going into the details of the mathematical conversion, the controlled process can be quantified algebraically as:

$$C = p(\text{correct} \mid \text{compatible}) - p(\text{incorrect} \mid \text{incompatible})$$

Using the specific number that has been computed for *C*, the automatic process can then be calculated as:

$$A = p(\text{incorrect} \mid \text{incompatible})/(1 - C)$$

For example, if the empirically observed probability of correct responses on compatible trials is 74 and the probability of correct responses on incompatible trials is 46, the resulting estimate for *C* is 20 and the estimate for *A* is 675.[1] Such estimates can be calculated for each participant in a given sample, allowing the use of these estimates as dependent variables in experimental designs or as independent variables in individual difference designs.

Automaticity-Dominating Process Dissociation Model

Even though Jacoby's (1991) PD model has been successfully applied to various tasks in the social-cognitive literature (for a review, see Payne & Bishara, 2009), the model's premise that automatic processes operate only when controlled processes fail does not seem applicable to tasks in which automatic processes play a dominant role despite the potential operation of controlled processes. For example, in the Stroop task, automatic word reading may elicit a tendency to respond on the basis of the semantic meaning of a colored word, which has to be overcome by controlled processes if the word meaning does not match the ink color in which the word is presented. To address this problem, Lindsay and Jacoby (1994) presented a modified variant of Jacoby's (1991) PD model, in which automatic processes are assumed to dominate, such that controlled processes drive responses only if automatic processes fail. This model can again be depicted as a processing tree (see Figure 14.5b). Even though the positions of automatic and controlled processes are reversed in Lindsay and Jacoby's (1994) model, the underlying logic remains the same. If automatic processes drive the response (depicted as *A* in Figure 14.5b), participants will show the correct response on compatible trials but the incorrect response on incompatible trials. If, however, automatic processes fail (depicted as *1 – A* in Figure 14.5b), the final response depends on the operation of controlled processes. If controlled processes succeed (depicted as *C* in Figure 14.5b), participants will show the correct response on both compatible and incompatible trials. However, if controlled processes fail (depicted as *1 – C* in Figure 14.5b), the model assumes incorrect responses on both compatible and incompatible trials. Using the algebraic logic outlined for Jacoby's (1991) PD model, the probability of a correct response on compatible trials can be formalized as follows:

$$p(\text{correct} \mid \text{compatible}) = A + (1 - A) \times C$$

Conversely, the probability of a correct response on incompatible trials can be formalized as follows:

$$p(\text{correct} \mid \text{incompatible}) = (1 - A) \times C$$

On the basis of these equations, the automatic process can be quantified algebraically as follows:

$$A = p(\text{correct} \mid \text{compatible}) - p(\text{correct} \mid \text{incompatible})$$

Using the quantitative estimates for *A*, the controlled process can then be calculated as follows:

$$C = p(\text{correct} \mid \text{incompatible})/(1 - A)$$

Although applications of PD have mostly used Jacoby's (1991) original formulation instead of Lindsay and Jacoby's (1994) modified version, the data analytic strategy of PD has been successfully applied to a wide range of questions in social psychology (for a review, see Payne & Bishara, 2009). Examples of such applications include racial bias in weapon identification (e.g., Payne, 2001), stereotypic biases in memory (e.g., Sherman, Groom, Ehrenberg, & Klauer, 2003), the use of heuristics in social judgment (e.g., Ferreira, Garcia-Marques, Sherman, & Sherman, 2006), and mood effects on automatic evaluations (e.g., Huntsinger, Sinclair, & Clore, 2009).

Quadruple-Process Model

A shared assumption of the two PD models is that one of the two processes operates only to the extent that the other one fails. For example, in Jacoby's (1991) model, automatic processes are assumed to operate only if controlled processes fail (see Figure 14.5a). Conversely, in Lindsay and Jacoby's (1994) model, controlled processes are assumed to operate only if automatic processes fail (see Figure 14.5b). Conrey, Sherman, Gawronski, Hugenberg, and Groom (2005) argued that these assumptions make the two PD models less suitable to capture situations in which automatic and controlled processes operate simultaneously, as is the case when one process overrides the impact of another process (see also Sherman et al., 2008). To address this limitation, Conrey et al. proposed their quadruple process model (quad model), which includes statistical parameters for four (instead of two) qualitatively distinct processes: (1) the likelihood that an automatic association is activated (described as *association activation* or *AC*); (2) the likelihood that the correct response to the stimulus can be determined (described as *discriminability* or *D*); (3) the likelihood that an automatic association is successfully overcome in favor of the correct response (described as *overcoming bias* or *OB*); and (4) the likelihood that a general response bias (e.g., right-hand bias) drives the response (described as *guessing* or *G*).

The proposed interplay of these processes in the quad model can again be depicted as a processing tree that specifies how their joint operation can lead to correct or incorrect responses on compatible and incompatible trials (Figure 14.6). The most significant component of the model is the assumption that activated associations (reflected in the *AC* parameter) and correct identification of the correct response (reflected in the *D* parameter) can produce two response tendencies that may be congruent or incongruent with each other. If they are incongruent, inhibitory control has to be engaged to suppress the response tendency elicited by activated associations in favor of the correct response (reflected in the *OB* parameter). For example, in an Implicit Association Test designed to measure automatic preferences for whites over blacks, automatic stereotypic associations may elicit a tendency to press the "negative" key in response to a black face. In the so-called incongruent block, this tendency has to be inhibited in favor of the correct "black" response. In this case, the quad model's *AC* parameter reflects the strength of automatic stereotypic associations; the *D* parameter reflects participants' ability to identify

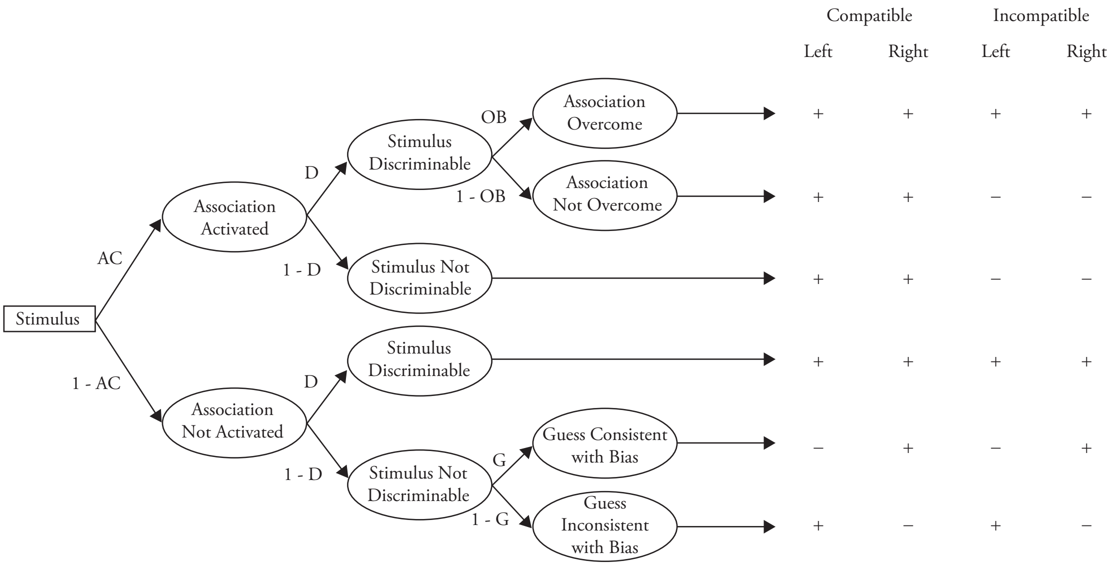

Figure 14.6 The quadruple process model of automatic and controlled processing. (Adapted from Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005. Reprinted with permission.)

the correct response; and the *OB* parameter reflects participants' success in inhibiting automatic stereotypic associations in favor of the correct response. In addition, the quad model's *G* parameter reflects the strength of general response tendencies, such as the tendency to show a right-hand response.

Although the higher number of processes increases the complexity of the mathematical equations in the quad model, the derivation of these equations is equivalent to the two PD models. For example, if the correct response for incompatible trials in a given task is defined as the right key, the three paths that lead to a correct response in the quad model's processing tree (see Figure 14.6) produce the following equation:

$$p\ (\text{correct} \mid \text{incompatible-right}) = [AC \times D \times OB] + [(1 - AC) \times D] + [(1 - AC) \times (1 - D) \times G]$$

The same logic applies to the derivation of the equations for the observed probabilities of correct responses on the other types of trials. Yet, an important difference between the quad model and PD models is that the latter produce two equations with two unknowns that can be solved through linear algebra. In contrast, the quad model entails more equations than unknowns. Consequently, parameter values cannot be calculated directly through linear algebra, but have to be estimated through alternative procedures. The procedure employed by the quad model is multinomial modeling (for a review, see Batchelder & Riefer, 1999), which uses maximum likelihood statistics to identify parameter estimates that minimize the discrepancy between the empirically observed probabilities of correct and incorrect responses and the corresponding probabilities that are predicted through the equations of the model (for more details regarding the statistics underlying the quad model, see Conrey et al., 2005).

Even though the conceptual relation between the quad model's parameters to the parameters of the two PD models is still under debate (Payne & Bishara, 2009; Sherman, Klauer, & Allen, 2010), the quad model's integration of multiple distinct processes provides a more fine-grained analysis compared with traditional dual process theories (Sherman, 2006). In a technical sense, the quad model is thus better described as a *multiple process model* rather than a *dual process model*, although it retains the emphasis on automatic and controlled processes. So far, the quad model has demonstrated its potential mostly in the analysis of data obtained with implicit measures, which often remain ambiguous as to whether a given effect is driven by differences in automatic associations or by any of the other three processes proposed by the model. Whereas in some cases, previous interpretations turned out to be accurate, interpretations of other findings had to be revised, such that effects that have been attributed to differences in automatic associations turned out to be driven by other processes, such as differences in overcoming the biasing influence of automatic associations (for a review, see Sherman et al., 2008).

Criticism of Formalized Dual Process Theories

Compared with phenomenon-specific and generalized dual process theories, formalized models have a unique advantage, in that they provide mathematical implementations that allow researchers to quantify the relative contributions of the proposed processes to a given task. At the same time, there have been controversies surrounding the proper interpretation of the obtained parameter estimates (Payne & Bishara, 2009; Sherman et al., 2010). An important issue—which applies equally to the two PD models as well as the quad model—is that formalized models are often assumed to provide direct access to automatic and controlled processes through the nature of their algebraic logic. For example, researchers using Jacoby's (1991) PD model often assume that the two processes captured by the model are *defined* as automatic or controlled through the basic logic of the model. In part, this assumption has its roots in misleading depictions of the parameter estimates as *automatic* (using the acronym *A*) and *controlled* (using the acronym *C*). Such depictions are problematic for at least two reasons. First, as outlined in the initial sections of this chapter, the different features of automaticity and control do not necessarily covary (Bargh, 1994; Moors & De Houwer, 2006), which makes generic descriptions of parameters as automatic or controlled misleading, as long as it is not specified in which particular sense the captured process is supposed to be automatic or controlled. Second, and more seriously, generic descriptions of parameters as automatic and controlled conflate the nature of a given process (i.e., what is the process doing?) with its operating conditions (i.e., under which conditions does the process operate?). After all, any claims about the conditions under which a process is operating represent empirical assumptions that have to be tested as such, and this empirical work cannot be delegated to the application of a mathematical procedure.

To illustrate this issue, consider Payne's (2001) application of Jacoby's (1991) PD model to a sequential priming task designed to investigate stereotypic biases in weapon identification. On each trial of the task, participants are briefly presented with either a black or a white face prime, which is immediately followed by a target picture showing either a gun or a harmless object. The target picture is quickly replaced by a black-and-white pattern mask, and participants' task is to indicate whether the target picture showed a gun or a harmless object. The common result is that harmless objects were more frequently misidentified as guns when the face prime was black than when it was white, whereas guns were more frequently misidentified as harmless objects when the face prime was white than when it was black (for a review, see Payne, 2006). Using Jacoby's (1991) PD model, Payne (2001) calculated separate estimates reflecting participants' ability to identify guns and harmless objects (reflected in the model's *C* parameter) and stereotypic biases in guessing the nature of the target stimulus if participants were unable to identify the stimulus (reflected in the model's *A* parameter). Importantly, whether or not the two processes—identification of the target object and stereotypic bias in guessing the nature of the target—operate unintentionally, efficiently, unconsciously, and uncontrollably is an empirical question about the conditions under which the two processes operate; it is not an inherent feature of the two processes per se.

Addressing this concern, Payne and Bishara (2009) have argued that applications of PD models to particular tasks usually involve one feature of automaticity by definition, whereas claims about other features represent empirical hypotheses that have to be tested as such. With regard to weapon identification, for example, Payne and Bishara have argued that *C* depicts an intentional process and *A* an unintentional process, with their (un)intentionality being defined a priori through the basic structure of the task. However, this conceptualization stands in contrast to research showing that estimated values of *A* increased when participants were instructed to intentionally use race as a cue in identifying the target object (Payne, Lambert, & Jacoby, 2002). Applied to Payne and Bishara's argument, this result leads to the paradoxical conclusion that a process that is *defined* as unintentional can be intentional. Describing the process captured by the *A* parameter as stereotypic bias without reference to intentionality (or other features of automaticity) avoids this problem because stereotypic biases can operate either intentionally or unintentionally. Similar considerations apply to the interpretation of the *C* parameter. In the weapon identification task, this parameter reflects the identification of the target object, which may or may not be intentional, conscious, resource dependent, and controllable. The bottom line is that formalized dual process (and multiple process) theories are very well suited to quantifying the contribution of qualitatively distinct processes (e.g., identification of target object, stereotypic bias). However, any claims about the *operating conditions* of these processes are empirical assumptions that have to be tested as such. The mere application of mathematical procedures cannot replace this empirical work.

Outlook

Dual process theories play a central role in social psychology, and the amount of research that has been stimulated by these theories is simply enormous. In addition to reviewing the core assumptions of the most prominent theories, we have tried to provide a historical perspective by describing the development of dual process theorizing from the emergence of phenomenon-specific theories in the 1980s to the recent advances made by generalized and formalized dual process theories. An interesting question is where dual process theorizing will go from here. In the final sections of this chapter, we offer some conceptual considerations regarding the current state of theorizing and discuss how some limitations of current theories could possibly be overcome.

Models Versus Theories

In line with the common usage of terminology in social psychology, we have used the terms *theory* and *model* interchangeably. Yet, many philosophers of science consider theories and models as conceptually distinct. Whereas theories are usually regarded as sets of well-specified *if–then* conditionals that link two or more concepts in a particular manner, models are regarded as nominal descriptions that provide a conceptual frame of reference. Even though the boundaries between theories and models may become somewhat blurry when the distinction is applied to theorizing in social psychology, it seems useful to keep the distinction in mind when comparing different types of dual process theories. For example, phenomenon-specific dual process theories seem closer to the notion of *theory*, in that they include specific *if–then* conditionals about links between psychological concepts. In contrast, many (though not all) of the reviewed generalized dual process theories seem closer to the notion of *model*, in that they primarily include nominal descriptions. To the extent that these models lack clearly specified *if–then* conditionals, their descriptive classifications do not imply any predictions that could be confirmed or disconfirmed. Instead, their functional value lies in their integrative capacity as frames of reference to describe observed phenomena. Similar considerations can be applied to formalized dual process theories, which also seem closer to the notion of *model*. To be sure, formalized models are extremely valuable in providing quantifications of qualitatively distinct processes. However, these models usually do not include any *if–then* conditionals about these processes that could be tested empirically. Of course, it is entirely possible to use formalized models to conduct more stringent tests of predictions made by other theories that include *if–then* conditionals. Yet, any such predictions are extrinsic to these models, in that they are not logically derived from their core assumptions.

These considerations have important implications for dual process theorizing. During the past decade, generalized dual process *models* have clearly proved their usefulness as integrative frameworks by providing conceptual links between phenomena and research findings that have not been related before. Similarly, formalized dual process *models* provide a powerful tool, not only to disentangle, but also to quantify the contributions of qualitatively distinct processes to overt behavioral responses. Yet, what is still missing in both kinds of models is a conceptual integration of the empirically confirmed *if–then* conditionals that have contributed to the rise of phenomenon-specific dual process *theories* in 1980s and 1990s. In addition, it seems desirable to formulate new *if–then* conditionals that allow researchers to derive novel predictions. Thus, an important task for future theorizing is to combine the unique characteristics of the three dual process approaches: (1) the predictive power of phenomenon-specific dual process theories, (2) the integrative capacity of generalized dual process models, and (3) the methodological advantages of formalized dual process models.

How Many Processes Are There?

A final question that deserves attention in a chapter on dual process theories is: How many processes are there? During the past decade, this question has sparked controversial debates, with some theorists arguing for the classic dual process distinction (e.g., Deutsch & Strack, 2006) and others endorsing either single process (e.g., Kruglanski et al., 2006) or multiple process alternatives (e.g., Sherman, 2006). In evaluating the arguments that have been raised in these debates, it is important to note that existence claims—such as 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). In other words, we cannot test empirically if there are one, two, or multiple processes. Still, researchers can make decisions about the usefulness of ontological claims by empirically testing assumptions *about* the proposed processes. To the extent that the predictions implied by a particular class of theories are confirmed, their underlying ontological claims are typically regarded as justified. Yet, if these predictions are continuously disconfirmed, it seems likely that researchers will at some point reject the underlying ontological claims. Note, however, that in these cases it is not the existence claim itself that is confirmed or disconfirmed, but the assumptions that are made about the proposed entities.

Evaluated from this point of view, dual process theories have fared very well, which explains their dominant role in social psychology. At the same time, it is essential to reiterate the quest for novel predictions, which still represent the hallmark of theory evaluation. As noted in our earlier discussion of generality–specificity trade-offs, theories with a high degree of generality are certainly functional in their capacity to explain a wide range of phenomena with a parsimonious set of basic principles. At the same time, highly integrative theories often come at the cost of specificity, which is essential for the derivation of novel predictions (Quine & Ullian, 1978). Even though this concern has been raised in particular against generalized dual process theories (e.g., Keren & Schul, 2009), it is important to note that it equally applies to their single process alternatives, whose high levels of generality make them capable of explaining almost any observable outcome in a post hoc fashion. Yet, their capacity for making predictions often seems rather weak, in that it is difficult to specify in an a priori fashion what outcome one would expect for a particular set of conditions. Thus, what seems essential to evaluate any type of ontological premise—be it in single process, dual process, or multiple process theorizing—is the formulation of precise *if–then* conditionals that generate novel predictions. Ideally, such *if–then* conditionals should be formulated in a manner that specifies not only which events are implied by the theory but also which events should not happen according to the theory. After all, the informative value of a scientific theory increases with the number of observations that are "prohibited" by the theory, not with the number of observations that are "allowed" (Popper, 1934). Thus, an important task for the refinement of dual process theorizing is the simultaneous consideration of generality and specificity to maximize both their explanatory and their predictive power.

Conclusion

During the past three decades, dual process theorizing has exerted an overwhelming impact on research in social psychology. Despite the reviewed criticism of dual process theorizing, it seems highly unlikely that this influence will dissipate in the near future. After all, the enormous body of research that has been inspired by dual process theories speaks for itself, and it is difficult to imagine how a posteriori explanations of alternative accounts could lead to a full replacement of the dual process idea. Yet, many of the concerns that have been raised against dual process theories seem justified on conceptual grounds. To the extent that dual process theorists take these concerns seriously, theorizing in social psychology can only become stronger.

Author Note

Preparation of this chapter has been supported by the Canada Research Chairs Program and the Social Sciences and Humanities Research Council of Canada. We are grateful to Jeff Sherman for valuable comments on an earlier version of this chapter. Please address correspondence to

Bertram Gawronski, Ph.D., Department of Psychology, The University of Western Ontario, Social Science Centre, London, Ontario N6A 5C2 Canada; e-mail: bgawrons@uwo.ca.

Notes

1. Note that p(correct | compatible) = 1 – p(incorrect | compatible). Correspondingly, p(correct | incompatible) = 1 – p(incorrect | incompatible).

References

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Asch, S. E., & Zuckier, H. (1984). Thinking about persons. *Journal of Personality and Social Psychology*. *46*, 1230–1240.

Back, M. D., Schmukle, S. C., & Egloff, B. (2009). Predicting actual behavior from the explicit and implicit self-concept of personality. *Journal of Personality and Social Psychology*, *97*, 533–548.

Banse, R., Gawronski, B., Rebetez, C., Gutt, H., & Morton, J. B. (2010). The development of spontaneous gender stereotyping in childhood: Relations to stereotype knowledge and stereotype flexibility. *Developmental Science*, *13*, 298–306.

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. (1997). The automaticity of everyday life. In R. S. Wyer, Jr. (Ed.), *Advances in social cognition* (Vol. 10, pp. 1–61). Mahwah, NJ: Erlbaum.

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin and Review*, *6*, 57–86.

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.

Brewer, M. B., & Harasty-Feinstein, A. S. (1999). Dual processes in the representation of persons and social categories. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 255–270). New York: Guilford Press.

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.

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., Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic processing within and beyond the persuasion context. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 212–252). New York: Guilford Press.

Chaiken, S., & Maheswaran, D. (1994). Heuristic processing can bias systematic processing: Effects of source credibility, argument ambiguity, and task importance on attitude judgment. *Journal of Personality and Social Psychology*, *66*, 460–473.

Chaiken, S., & Trope, Y. (Eds.). (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.

Chun, W. Y., & Kruglanski, A. W. (2006). The role of task demands and processing resources in the use of base-rate and individuating information. *Journal of Personality and Social Psychology*, *91*, 205–217.

Chun, W. Y., Spiegel, S., & Kruglanski, A. W. (2002). Assimilative behavior identification can also be resource dependent: A unimodel perspective on personal-attribution phases. *Journal of Personality and Social Psychology*, *83*, 542–555.

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.

D'Agostino, P. R., & Fincher-Kiefer, R. (1992). Need for cognition and the correspondence bias. *Social Cognition*, *10*, 151–163.

Deutsch, R., Gawronski, B., & Strack, F. (2006). At the boundaries of automaticity: Negation as reflective operation. *Journal of Personality and Social Psychology*, *91*, 385–405.

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.

Duncan, B. L. (1976). Differential perception and attribution of intergroup violence: Testing the lower limits of stereotyping of Blacks. *Journal of Personality and Social Psychology*, *34*, 590–598.

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

Epstein, S. (1973). The self-concept revisited. Or a theory of a theory. *American Psychologist*, *28*, 404–416.

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

Epstein, S., & Pacini, R. (1999). Some basic issues regarding dual-process theories from the perspective of Cognitive-Experiential Self-Theory. In S. Chaiken, & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 462–482). New York: Guilford Press.

Epstein, S., Pacini, R., Denes-Raj, V., & Heier, H. (1996). Individual differences in intuitive-experiential and analytical-rational thinking styles. *Journal of Personality and Social Psychology*, *71*, 390–405.

Erb, H. P., Pierro, A., Mannetti, L., Spiegel, S., & Kruglanski, A. W. (2007). Biased processing of persuasive evidence: On the functional equivalence of cues and message arguments. *European Journal of Social Psychology*, *37*, 1057–1075.

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

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

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

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, *54*, 297–327.

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

Ferreira, M. B., Garcia-Marques, L., Sherman, S. J., & Sherman, J. W. (2006). A dual-process approach to judgment under uncertainty. *Journal of Personality and Social Psychology*, *91*, 797–813.

Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row Peterson.

Fiske, S. T., Lin, M., & Neuberg, S. L. (1999). The continuum model: Ten years later. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 231–254). New York: Guilford Press.

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.

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. (2004). Theory-based bias correction in dispositional inference: The fundamental attribution error is dead, long live the correspondence bias. *European Review of Social Psychology*, *15*, 183–217.

Gawronski, B., Alshut, E., Grafe, J., Nespethal, J., Ruhmland, A., & Schulz, L. (2002). Prozesse der Urteilsbildung über bekannte und unbekannte Personen: Wie der erste Eindruck die Verarbeitung neuer Informationen beeinflusst [Processes of judging known and unknown persons: How the first impression influences the processing of new information]. *Zeitschrift für Sozialpsychologie*, *33*, 25–34.

Gawronski, B., & Bodenhausen, G. V. (2006a). 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. (2006b). Associative and propositional processes in evaluation: Conceptual, empirical, and meta-theoretical issues. Reply to Albarracín, Hart, and McCulloch (2006), Kruglanski and Dechesne (2006), and Petty and Briñol (2006). *Psychological Bulletin*, *132*, 745–750.

Gawronski, B., & Bodenhausen, G. V. (2007). Unraveling the processes underlying evaluation: Attitudes from the perspective of the APE Model. *Social Cognition*, *25*, 687–717.

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., Deutsch, R., LeBel, E. P., & Peters, K. R. (2008). Response interference as a mechanism underlying implicit measures: Some traps and gaps in the assessment of mental associations with experimental paradigms. *European Journal of Psychological Assessment*, *24*, 218–225.

Gawronski, B., Geschke, D., & Banse, R. (2003). Implicit bias in impression formation: Associations influence the construal of individuating information. *European Journal of Social Psychology*, *33*, 573–589.

Gawronski, B., & LeBel, E. P. (2008). Understanding patterns of attitude change: When implicit measures show change, but explicit measures do not. *Journal of Experimental Social Psychology*, *44*, 1355–1361.

Gawronski, B., LeBel, E. P., & Peters, K. R. (2007). What do implicit measures tell us? Scrutinizing the validity of three common assumptions. *Perspectives on Psychological Science*, *2*, 181–193.

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

Gawronski, B., & Sritharan, R. (2010). Formation, change, and contextualization of mental associations: Determinants and principles of variations in implicit measures. In B. Gawronski, & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 216–240). New York: Guilford Press.

Gawronski, B., & Strack, F. (2004). On the propositional nature of cognitive consistency: Dissonance changes explicit, but not implicit attitudes. *Journal of Experimental Social Psychology*, *40*, 535–542.

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.

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

Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin*, *117*, 21–38.

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*, *54*, 733–740.

Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge, UK: Cambridge University Press.

Greeley, A. M., & Sheatsley, P. B. (1971). Attitudes toward racial integration. *Scientific American*, *225*, 13–19.

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.

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

Grumm, M., Nestler, S., & von Collani, G. (2009). Changing explicit and implicit attitudes: The case of self-esteem. *Journal of Experimental Social Psychology*, *45*, 327–335.

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133–168). New York: Guilford Press.

Hoefling, A., Likowski, K. U., Deutsch, R., Häfner, M., Seibt, B., Mühlberger, A., Weyers, P., & Strack, F. (2009). When hunger finds no fault with moldy corn: Food deprivation reduces food-related disgust. *Emotion*, *9*, 50–58.

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

Hugenberg, K., & Bodenhausen, G. V. (2003). Facing prejudice: Implicit prejudice and the perception of facial threat. *Psychological Science*, *14*, 640–643.

Hugenberg, K., & Bodenhausen, G. V. (2004). Ambiguity in social categorization: The role of prejudice and facial affect in racial categorization. *Psychological Science*, *15*, 342–345.

Huntsinger, J. R., Sinclair, S., & Clore, G. L. (2009). Affective regulation of implicitly measures stereotypes and attitudes: Automatic and controlled processes. *Journal of Experimental Social Psychology*, *45*, 560–566.

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

Jones, E. E. (1990). *Interpersonal perception*. New York: Freeman.

Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. *Advances in Experimental Social Psychology*, *2*, 219–266.

Jones, E. E., & Harris, V. A. (1967). The attribution of attitudes. *Journal of Experimental Social Psychology*, *3*, 1–24.

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

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press.

Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska symposium on motivation* (pp. 192–240). Lincoln, NE: University of Nebraska Press.

Kelley, H. H. (1972). Causal schemata and the attribution process. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 151–174). Morristown, NJ: General Learning Press.

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.

Kruglanski, A. W., & Ajzen, I. (1983). Bias and error in human judgment. *European Journal of Social Psychology*, *13*, 1–44.

Kruglanski, A. W., & Dechesne, M. (2006). Are associative and propositional processes qualitatively distinct? Comment on Gawronski and Bodenhausen (2006). *Psychological Bulletin*, *132*, 736–739.

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., Pierro, A., Mannetti, L., Erb, H. P., & Chun, W. Y. (2007). On the parameters of human judgment. *Advances in Experimental Social Psychology*, *39*, 255–303.

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

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.

Krull, D. S., & Dill, J. C. (1996). On thinking first and responding fast: Flexibility in the social inference process. *Personality and Social Psychology Bulletin*, *22*, 949–959.

Krull, D. S., & Erickson, D. J. (1995). Judging situations: On the effortful process of taking dispositional information into account. *Social Cognition*, *13*, 417–438.

Kunda, Z., Miller, D. T., & Claire, T. (1990). Combining social concepts: The role of causal reasoning. *Cognitive Science*, *14*, 551–577.

Kunda, Z., & Sherman-Williams, B. (1993). Stereotypes and the construal of individuating information. *Personality and Social Psychology Bulletin*, *19*, 90–99.

Leyens, J. P., & Yzerbyt, V. Y. (1992). The ingroup overexclusion effect: The impact of valence and confirmation on information search. *European Journal of Social Psychology*, *22*, 549–569.

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

Lieberman, M. D. (2007). Social cognitive neuroscience: A review of core processes. *Annual Review of Psychology*, *58*, 259–289.

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

Lindsay, D. S., & Jacoby, L. L. (1994). Stroop process-dissociations: The relationship between facilitation and interference. *Journal of Experimental Psychology: Human Perception & Performance*, *20*, 219–234.

Maheswaran, D., & Chaiken, S. (1991). Promoting systematic processing in low-motivation settings: Effects of incongruent information on processing and judgment. *Journal of Personality and Social Psychology*, *61*, 13–25.

Maheswaran, D., Mackie, D. M., & Chaiken, S. (1992). Brand name as an heuristic cue: The effects of task importance and expectancy confirmation on consumer judgments. *Journal of Consumer Psychology*, *1*, 317–336.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex? Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*, 419–457.

McConahay, J. B. (1986). Modern racism, ambivalence, and the modern racism scale. In J. D. Dovidio, & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 91–125). Orlando, FL: Academic Press.

Metcalfe, J., & Mischel, W. (1999). A hot/cool-system analysis of delay of gratification. Dynamics of willpower. *Psychological Review*, *106*, 3–19.

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

Neumann, R., Förster, J., & Strack, F. (2003). Motor compatibility: The bidirectional link between behavior and emotion. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 371–391). Mahwah, NJ: Erlbaum.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research* (pp. 1–18). New York: Plenum.

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

Olson, M. A., & Fazio, R. H. (2006). Reducing automatically activated racial prejudice through implicit evaluative conditioning. *Personality and Social Psychology Bulletin*, *32*, 421–433.

Olson, M. A., & Fazio, R. H. (2009). Implicit and explicit measures of attitudes: The perspective of the MODE model. In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.), *Attitudes: Insights from the new implicit measures* (pp. 19–63). New York: Psychology Press.

Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology*, *81*, 181–192.

Payne, B. K. (2006). Weapon bias: Split second decisions and unintended stereotyping. *Current Directions in Psychological Science*, *15*, 287–291.

Payne, B. K. (2008). What mistakes disclose: A process dissociation approach to automatic and controlled processes in social psychology. *Social and Personality Psychology Compass*, *2*, 1073–1092.

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., Lambert, A. J., & Jacoby, L. L. (2002). Best laid plans: Effects of goals on accessibility bias and cognitive control in race-based misperceptions of weapons. *Journal of Experimental Social Psychology*, *38*, 384–396.

Peters, K. R., & Gawronski, B. (2011). Mutual influences between the implicit and explicit self-concepts: The role of memory activation and motivated reasoning. *Journal of Experimental Social Psychology*, *47*, 436–442.

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

Petty, R. E., & Cacioppo, J. T. (1984). Source factors and the elaboration likelihood model of persuasion. *Advances in Consumer Research*, *11*, 668–672.
Petty, R. E., & Cacioppo, J. T. (1986). The Elaboration Likelihood Model of persuasion. *Advances in Experimental Social Psychology*, *19*, 123–205.
Petty, R. E., Cacioppo, J. T., & Schumann, D. W. (1983). Central and peripheral routes to advertising effectiveness: The moderating role of involvement. *Journal of Consumer Research*, *10*, 135–146.
Petty, R. E., & Wegener, D. T. (1999). The Elaboration Likelihood Model: Current status and controversies. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 41–72). New York: Guilford Press.
Popper, K. R. (1934). *Logic of scientific discovery*. New York: Basic Books.
Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
Quattrone, G. A. (1982). Overattribution and unit formation: When behavior engulfs the person. *Journal of Personality and Social Psychology*, *42*, 593–607.
Quine, W. V. O. (1960). *Word and object*. Cambridge MA: MIT Press.
Quine, W. V. O., & Ullian, J. S. (1978). *The web of belief* (2nd ed.). New York: McGraw-Hill.
Richards, Z., & Hewstone, M. (2001). Subtyping and subgrouping: Processes for the prevention and promotion of stereotype change. *Personality and Social Psychology Review*, *5*, 52–73.
Ross, L. D. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. *Advances in Experimental Social Psychology*, *10*, 173–220.
Ross, L. D., Amabile, T. M., & Steinmetz, J. L. (1977). Social roles, social control, and biases in social-perception processes. *Journal of Personality and Social Psychology*, *35*, 485–494.
Sagar, H. A., & Schofield, J. W. (1980). Racial and behavioral cues in black and white children's perceptions of ambiguously aggressive acts. *Journal of Personality and Social Psychology*, *39*, 590–598.
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.
Sears, D. O. (1988). Symbolic racism. In P. A. Katz, & D. A. Taylor (Eds.), *Eliminating racism: Profiles in controversy* (pp. 53–84). New York: Plenum.
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.
Sherman, J. W., Groom, C., Ehrenberg, K., & Klauer, K. C. (2003). Bearing false witness under pressure: Implicit and explicit components of stereotype-driven memory bias. *Social Cognition*, *21*, 213–246.
Sherman, J. W., Klauer, K. C., & Allen, T. J. (2010). Mathematical modeling of implicit social cognition: The machine in the ghost. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 156–175). New York: Guilford Press.
Shiffrin, R., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, *84*, 127–190.
Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*, 3–22.
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.
Snyder, M. L., & Frankel, A. (1976). Observer bias: A stringent test of behavior engulfing the field. *Journal of Personality and Social Psychology*, *34*, 857–864.
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.
Taylor, D. G., Sheatsley, P. B., & Greeley, A. M. (1978). Attitudes toward racial integration. *Scientific American*, *238*, 42–49.
Trope, Y. (1986). Identification and inferential processes in dispositional attribution. *Psychological Review*, *93*, 239–257.
Trope, Y., & Alfieri, T. (1997). Effortfulness and flexibility of dispositional judgment processes. *Journal of Personality and Social Psychology*, *73*, 662–674.
Trope, Y., & Cohen, O. (1989). Perceptual and inferential determinants of behavior-correspondent attributions. *Journal of Experimental Social Psychology*, *25*, 142–158.
Trope, Y., Cohen, O., & Alfieri, T. (1991). Behavior identification as a mediator of dispositional inference. *Journal of Personality and Social Psychology*, *61*, 873–883.
Trope, Y., Cohen, O., & Maoz, Y. (1988). The perceptual and inferential effects of situational inducements on dispositional attribution. *Journal of Personality and Social Psychology*, *55*, 165–177.
Trope, Y., & Gaunt, R. (1999). A dual-process model of overconfident attributional inferences. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 161–178). New York: Guilford Press.
Trope, Y., & Gaunt, R. (2000). Processing alternative explanations of behavior: Correction or integration? *Journal of Personality and Social Psychology*, *79*, 344–354.
Tversky, A., & Kahneman, D. (1983). Extensional vs. intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, *90*, 293–315.
Vonk, R. (1999). Effects of outcome dependency on correspondence bias. *Personality and Social Psychology Bulletin*, *25*, 382–389.
Webster, D. M. (1993). Motivated augmentation and reduction of the overattribution bias. *Journal of Personality and Social Psychology*, *65*, 261–271.
Whitfield, M., & Jordan, C. H. (2009). Mutual influences of explicit and implicit attitudes. *Journal of Experimental Social Psychology*, *45*, 748–759.
Wicker, A. W. (1969). Attitudes versus actions: The relationship of verbal and overt behavioral responses to attitude objects. *Journal of Social Issues*, *25*, 41–78.
Wilson, T. D., Lindsey, S. & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, *107*, 101–126.