Social and Personality Psychology Compass 2 (2008): 10.1111/j.1751-9004.2008.00091.x

What Mistakes Disclose: A Process Dissociation Approach to Automatic and Controlled Processes in Social Psychology

B. Keith Payne*
University of North Carolina, Chapel Hill

Abstract

Process dissociation is a model for separating automatic and controlled contributions to responses in a single task. Although it was developed to separate conscious and unconscious uses of memory, researchers have increasingly found the model useful to answer questions in social psychology. This article reviews process dissociation studies with a focus on social psychology. It emphasizes the model's conceptual definitions of automatic and controlled processing and how those relate to procedures for estimating them. Process dissociation is contrasted with task dissociation, in which automatic and controlled processes are identified with implicit and explicit tasks. Current trends and future directions are identified, including the use of model testing procedures to compare competing theories of how automatic and controlled processes interact.

On May 24, 2005, U.S. President George W. Bush told an audience in Rochester, NY, 'In my line of work you got to keep repeating things over and over and over again for the truth to sink in, to kind of catapult the propaganda.'[1] To his supporters, it was a meaningless verbal gaffe. But to his critics, it was a revealing insight into Bush's approach to truth and communication. These eponymous slips have intrigued psychologists at least since Freud, who believed they were windows into the unconscious. A single incident like this probably reveals more about its interpreters than about the speaker. But what if the speaker made repeated errors that pointed in the same direction? A consistent pattern would make even the most charitable observers suspect something at work 'beneath' the surface. This article describes the process dissociation procedure, a method for uncovering the processes underlying a complex behavior (Jacoby, 1991). My aim is to show how it can turn errors into insights about implicit social cognition.

The procedure offers a tool for separating automatic and intentionally controlled components of thought, but perhaps more importantly, it suggests a more nuanced way of thinking about what automatic and controlled

processes mean. The distinction between automatic and controlled processes has become an important concern across nearly all areas of psychology. And in social psychology, implicit social cognition has been one of the most heavily studied topics of the last decade. There is general agreement in the social psychology literature on the features of automatic processes (they are fast, effortless, efficient, and insensitive to intentions) and controlled processes (they are slow, effortful, costly of cognitive resources, and are readily shaped by intentions; Bargh, 1989; Shiffrin & Schneider, 1977). But there is nonetheless a striking divergence of opinion about how to think about these processes, how to measure them, and how they interact to drive behavior. The process dissociation approach offers a unifying perspective on these issues by asking a few simple questions about what people intend to do and what they do otherwise.

From a process dissociation perspective, understanding the automatic and controlled components driving behavior requires answering three questions. First, what does a person intend to do? That is, what does he or she do when in perfect control? Second, does the person actually do what is intended? The difference between those two questions shows how much control a person has. If President Bush intended to tell the audience that he likes to repeat messages to more effectively spread propaganda (as he actually did), then he was apparently in good control. He most likely intended to say something else, however, which means his control over his language was less than perfect. The first two questions establish intent and control. The last question is: what does a person do when control fails? If the President has frequent slip-ups in which he says something different from what he intends, but his mis-statements are random, then there is no reason to surmise some underlying automatic tendency, only poor control. But if every time he gets confused or tongue-tied, he betrays a plan to advance his policies through half-truths and misinformation, then we learn something about his automatic tendencies. If we know what a person intends to do, whether he actually does it, and also what kinds of unintended errors he makes, we are in a good position to draw conclusions about both automatic and controlled components of behavior. The following pages describe how the procedure was developed to study automatic and controlled uses of memory, and then how it has been expanded to study a variety of topics in social cognition. The final section points to new directions to which a process dissociation approach may lead.

Theoretical Background: Separating Automatic and Controlled Uses of Memory

Jacoby (1991) developed the process dissociation procedure in response to the growing recognition that implicit and explicit memory tests were not 'process-pure'. As an example, in an implicit memory test, subjects might try to complete a word fragment with the first word that comes to mind.

Conscious memory could contaminate the test if a subject thinks back to a studied list for cues to the completion. In an explicit test, subjects might be instructed to intentionally think back to the studied list to complete the word fragment. In this case, unconscious memory could contaminate the test, if a person cannot remember the item but guesses based on the first item that comes to mind (Jacoby, Toth, & Yonelinas, 1993; Schacter, 1987). Jacoby's first innovation was to separate intentional and unintentional influences by placing them in opposition. One study asked subjects to study a list of non-famous names, and either a few minutes or a full day later, to discriminate between famous and non-famous names (Jacoby, Kelley, Brown, & Jasechko, 1989). Subjects were told that none of the names on the studied list were famous; thus, if they consciously remembered studying a name they could reject it as non-famous. But studying the names made them more familiar, an automatic influence that could easily be misattributed to the name being famous. Thus, subjects were more likely to call non-famous names famous when they had been previously studied, but only after a day's delay when recollection for the list was poor. Placing conscious memory and automatic influences of memory in opposition allowed them to be clearly separated.

Opposition procedures like the false fame study provide a conservative test of automatic memory influences because conscious memory works against detecting the automatic influence. When automatic influences are demonstrated, researchers can have confidence that they are unlikely to be contaminated by controlled memory. But for the same reasons, opposition procedures tend to underestimate automatic influences. Jacoby's next innovation was to correct for this underestimation by comparing conditions in which automatic and controlled influences were opposed to conditions in which they worked in concert. In one experiment (Jacoby et al., 1993), words were presented for study with full or divided attention and then later tested with word stems (e.g., water; wat–). In the inclusion test, subjects were instructed to use the stem as a cue for recall of a studied word or, if they could not remember the item, to complete the stem with the first word that came to mind. Subjects could complete a stem with a studied word either because they recollected the word, with a probability R, or because the studied word came automatically to mind (A) as a completion even when they did not recollect $(1 - R)$. The likelihood of completing the item with a studied word in the inclusion condition was thus $R + A(1 - R)$. In the exclusion test, subjects were instructed to complete the stem with the first word that came to mind, but to *not* use a recalled word. That is, they were told to exclude old words and use only new words. In this condition, a stem would be completed with a studied word only if recollection failed and the word came automatically to mind: $A(1 - R)$.

Assuming that automatic and controlled uses of memory are statistically independent, researchers can then solve for estimates of automatic and

controlled memory processes. The difference between the inclusion (trying to use old words: R + A[1 − R]) and exclusion (trying not to use old words: A[1 − R]) tests provides a measure of recollection. Given that estimate, one can compute the probability of an old word automatically coming to mind: A = Exclusion/(1 − R). Results showed that divided, compared with full, attention during study significantly decreased recollection but had no effect on unconscious influences.

Memory studies with inclusion and exclusion conditions allow automatic and controlled responses to be separated because they answer all three of the questions outlined above. First, the inclusion and exclusion instructions define subjects' intentions; thus, the research knows how subjects intend to respond. Second, the difference between performance on inclusion and exclusion conditions provides an estimate of how much control subjects have over their memory performance; it is the difference between how they intend to respond and how they actually respond. And third, the automatic estimate measures the errors subjects make when control fails.

The process dissociation procedure does not guarantee that the processes labeled as automatic and controlled (e.g., A and R) actually are automatic and controlled, respectively. That is an empirical question that must be validated using manipulations believed to selectively influence one type of process or the other. Across many studies, these procedures have shown that controlled use of memory is strongly affected by variables such as divided attention, depth of processing, speeded responding and aging. In contrast, automatic influences of memory tend to be affected by variables that influence how easily an item comes to mind, such as perceptual fluency, conceptual fluency, and habitual responses (for a review, see Yonelinas, 2002). The process dissociation procedure has been used extensively in the study of human memory over the last 15 years. More recently, researchers in many other fields, and especially in social psychology, have adopted the procedure. Some of the studies have examined automatic and controlled uses of memory as a means to answer social psychological questions, and others have extended the procedure beyond memory altogether.

Automatic and Controlled Memory in Social Cognition

Person memory, the study of how people encode, store, and retrieve information about other people, is a foundational topic in social cognition (Srull & Wyer, 1989; Wyer & Carlston, 1979). Sophisticated theories have been developed to explain how people organize representations of persons, traits, and behaviors in memory (Hastie & Kumar, 1979; Srull & Wyer, 1989). Those theories make detailed predictions about the conditions under which person-based and category-based information will be retrieved. They have not usually distinguished between automatic and controlled

processes, however. Recently, process dissociation has helped characterize the roles of automatic and controlled uses of memory for persons. As an example, Hense, Penner, and Nelson (1995) studied the effects of stereotypes on memory for the traits of older and younger adults. After studying traits that described a series of older and younger individuals, subjects showed a stereotype-consistent memory bias. They preferentially recalled traits such as *slow* about older adults, and traits such as *energetic* about younger individuals. Using inclusion and exclusion instructions similar to those described above, Hense et al. found that conscious recollection was similar for stereotype-consistent and inconsistent traits. Automatic memory, however, was biased in a stereotypical direction, suggesting that the stereotype-consistent memory bias resulted from stereotypical traits coming automatically to mind (see also Sherman, Groom, Ehrenberg, & Klauer, 2003). Studies using subjective reports to study the phenomenology of conscious recollection and automatic familiarity provide converging evidence. Using Tulving's (1985) Remember/Know Distinction, studies have found that expectancy-consistent memory biases were described primarily as 'known', whereas accurate memories were described primarily as 'remembered' (Macrae, Schloerscheidt, Bodenhausen, & Milne, 2002).

Studies like these suggest that stereotype memory biases should be difficult to overcome because they are mediated largely by automatic processes; these processes are difficult to control and deceptive in their phenomenology. Although stereotype biases may not be experienced as 'remembered', the 'know' experience means that subjects have some confidence in their truth. In one study, Payne, Jacoby, and Lambert (2004) examined whether the subjective experiences associated with automatic memory undermined attempts to avoid them. Subjects studied lists of name–occupation pairs in which White-typical and Black-typical names were paired with stereotype-consistent and inconsistent jobs (e.g., basketball player and politician). Later, they were shown the names and asked to identify the correct occupations and rate their confidence in each response. As in past research, subjects showed stereotypical memory biases, and process dissociation analyses showed that the bias was mediated by automatic memory influences, not conscious recollection. Of most interest was the relationship between memory processes and confidence, displayed in Figure 1. Recollection was strongly associated with confidence, but automatic memory biases were unrelated to confidence. That is, subjects' memories were as likely to be influenced by automatic stereotyping when they felt completely confident as when they felt no confidence at all.

A second study gave subjects in one condition the option at test to 'pass' whenever they did not know the answer and encouraged subjects to pass rather than to make mistakes (Payne et al., 2004). This option is akin to the choices everyone sometimes makes in daily life to simply say

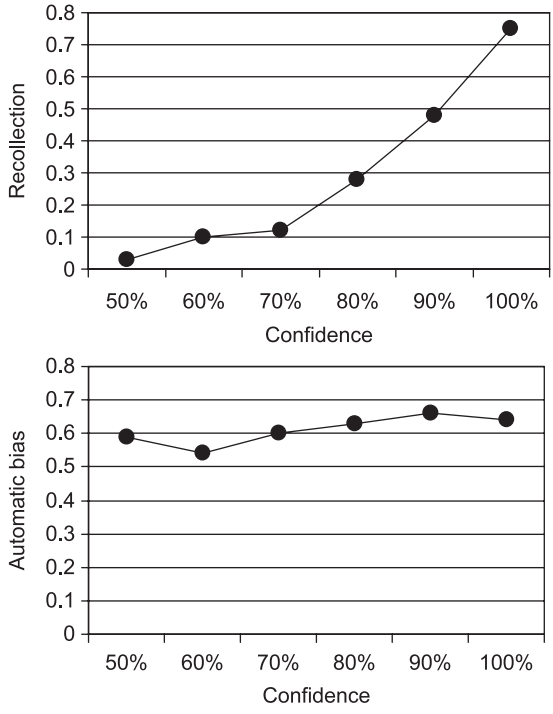

Figure 1 Relationships between subjective confidence and Recollection (top) and Automatic Bias (bottom). Confidence ranged from 50%, (i.e., responding at chance) to 100% (i.e., complete confidence). Adapted from Payne et al. (2004).

nothing when they fear they may say the wrong thing. But like so many times in daily life, subjects did not know when to hold their tongues. Their memory reports when given the pass option were as biased as when they were forced to respond on every trial.

The descriptions 'automatic' and 'controlled' should not be used as blanket statements, as has been pointed out persuasively by Bargh (1994) and Moors and De Houwer (2006). The features distinguishing them, such as intent, control, awareness, or efficiency, may occur separately or together in various combinations. The studies just described suggest that in the context of memory biases, automatic and controlled components are differentially related to subjective awareness and therefore differentially support control over responding. Stereotypes influence memory through automatic memory processes more than conscious recollection; but subjective experiences such as confidence and vivid remembering are tuned to recollection rather than automatic memory influences. The result is that monitoring and controlling stereotypic memory biases is difficult even when people are well intentioned, motivated to be accurate, and have the opportunity to edit their responses.

Beyond Memory

Although process dissociation was developed to separate memory processes, it represents a general theoretical approach to automatic and controlled processes that can be applied in virtually any field. As the distinction between automatic and controlled processes has become important across many areas of psychology, the process dissociation procedure has been applied more broadly. Nowhere has the automatic-controlled distinction been more central than in studies of implicit attitudes, prejudices, and stereotypes.

Social attitudes and stereotyping

The steep decline in racial prejudice over the last half century has been so dramatic that some observers in the 1980s began to suspect that it was too good to be true. In nationally representative surveys, for example, the proportion of White Americans supporting racially integrated schools rose from 32% in 1942 to 90% by 1982; the proportion saying that they would vote for a qualified Black presidential candidate rose from 37% in 1958 to 81% in 1983 (Schumann, Steeh, & Bobo, 1985). But whereas survey responses were becoming more egalitarian, studies using subtle unobtrusive measures continued to show evidence of racial bias (Crosby, Bromley, & Saxe, 1980). An influential explanation holds that White Americans' consciously controlled responses have become more egalitarian in response to changing social norms, whereas their automatic responses have remained relatively negative (Devine, 1989).

Studies pursuing this idea have flourished in the last decade with the development of implicit tests, whose aim is to measure automatically activated attitudes. Implicit tests measure attitudes indirectly, often using reaction times or behavioral responses rather than relying on self-report (e.g., Fazio, Jackson, & Dunton, 1995; Greenwald, McGhee, & Schwartz, 1998; Payne, Cheng, Govorun, & Stewart, 2005; Wittenbrink, Judd, & Park, 1997). Across hundreds of studies performed using various implicit attitude tests, there is one clear commonality: implicit tests of race attitudes and stereotypes tend to show different results than explicit tests. They tend to be only weakly correlated (Fazio & Olson, 2003; Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005) and they sometimes predict different behaviors or predict behaviors under different conditions (e.g., Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Lambert, Payne, Ramsey, & Shaffer, 2005). Many researchers have concluded from these findings that implicit cognitions are very different than explicit cognitions. In some theories, the divergence results because implicit tests tap automatic responses whereas explicit tests tap consciously edited responses (Fazio et al., 1995). In other theories, it is because implicit tests tap unconscious attitudes, whereas explicit tests tap conscious attitudes (Greenwald &

Banaji, 1995; Wilson, Lindsey, & Schooler, 2000). Both kinds of theories, however, have relied on a *task dissociation* approach that equates automatic processes with implicit tests and consciously controlled processes with explicit tests. That is, these theories assume the tests are process pure, in that implicit tests exclusively reflect automatic processes and explicit tests reflect consciously controlled processes. In many cases, the assumption may be unrealistic.

Consider all the ways that implicit and explicit tests differ. Explicit attitude tests are usually surveys in which subjects read sentences and decide how they feel about items and formulate their answers on a numeric scale. The items are sometimes as abstract as policies like affirmative action or busing. In an implicit test such as affective priming, subjects see a picture or word flashed briefly as a prime, followed by a pleasant or unpleasant word and they must classify the word as good or bad (Fazio et al., 1995). The metric is the reaction time to classify the word, depending on the prime. In the implicit association test, subjects sort two sets of items into four combined categories (Greenwald et al., 1998). For example, they may see pictures of Black and White faces and good and bad words, and sort each into the categories 'Black or good', 'White or bad', 'Black or bad', and 'White or good'. Again, the measure of interest is reaction time. To be sure, people can control their responses on a questionnaire more easily than they can control their reaction times on these implicit tests. But that is only one of many differences between the tests. They also differ in the concreteness of the items that subjects evaluate, the ways that attitude objects are represented (e.g., sentences vs. pictures), the complexity of the tasks, the metric (e.g., Likert scales vs. reaction times), and more. If an implicit test and an explicit test show different results, or if they fail to correlate strongly, is it because of differences between implicit and explicit cognition? Or is it because of any of these other differences? It is impossible to tell because a task dissociation approach confounds these features with the implicit-explicit distinction (Payne, Burkley, & Stokes, 2008).

In contrast to the task dissociation approach, the process dissociation approach overcomes this problem by separating processes within the same task rather than comparing two very different tasks. Consider a study conducted soon after the highly publicized death of Amadou Diallo, who was mistakenly shot by New York City police officers who mistook the wallet in his hand for a gun (Payne, 2001). Because Diallo was unarmed and Black, some critics alleged that race biased the officers' use of force. But as in President Bush's verbal mistake, more than a single error is needed to draw any conclusions. The question in the 2001 study was whether this kind of error – systematically mistaking a harmless object for a weapon – can reveal unintended influences of racial attitudes and stereotypes. Subjects distinguished between guns and harmless hand tools that were flashed briefly, but perceptibly, on a computer screen. Immediately

preceding each object was a Black or White male face that served as a prime. This 2×2 design creates conditions in which intentional responding to the target items and automatic influences of racial stereotypes are in concert (an inclusion condition), and in opposition (an exclusion condition). For example, when the prime was Black and the target was a gun, subjects could correctly respond 'gun' either by intentionally controlled (C) detection of the gun, or by an automatic stereotypical response (A) when control failed $(1 - C)$, with $C + A(1 - C)$ representing the likelihood of a correct 'gun' response. In contrast, when the prime was Black and the target was a tool, subjects would incorrectly respond 'gun' when controlled detection failed, but automatic stereotyping favored the gun response: $A(1 - C)$. The degree of intentional control can be solved by taking the difference between 'gun' responses in the inclusion and exclusion conditions. Given that estimate, the degree of automatic bias can be solved by dividing stereotypical false 'gun' responses by failures of control $(1 - C)$.

What do the process estimates mean in this context? This experiment controlled subjects' intentions via the task requirements to distinguish guns from tools. For that reason, the ability to accurately discriminate between target items provides an index of how much control subjects have over their responses. The Control estimate indexes how well subjects carried out their intentions by distinguishing between target objects. In contrast, subjects do not intend to be influenced by the racial primes. The Automatic estimate indexes how much these unintended influences biased their responses. Automatic and controlled processes can be estimated within the same task, avoiding all the problems involved in equating processes with tasks.

The results of the study showed that the race primes did indeed bias responses, as a tool was more likely to be mistaken for a gun when it was primed with a Black face than a White face. But more importantly for present concerns, the process estimates successfully separated automatic and controlled components of responses. Requiring subjects to respond quickly sharply reduced the controlled component, a well-established characteristic of controlled processing. But the race primes did not affect the controlled component. In contrast, the race primes affected the automatic component, but response speed did not. The automatic component also correlated in systematic ways with other measures of racial attitudes, increasing confidence in the validity of the estimate as an index of automatic racial bias.

The findings in the weapon bias studies have been replicated and extended in ways that document several aspects of automaticity and control. In one study, depleting self-regulation resources by having subjects complete hundreds of trials of a difficult cognitive task reduced the controlled component in the weapon bias, but not the automatic component (Govorun & Payne, 2006). In another, the automatic component correlated with implicit measures of race bias, whereas the controlled component

correlated with other measures of cognitive control such as an anti-saccade task, in which subjects must effortfully direct attention away from a character flashing on the screen that would automatically attract attention (Payne, 2005). It is worth emphasizing again that features defining automaticity and control do not necessarily co-occur. The evidence suggests that automatic bias in the weapon bias is fast and efficient, relatively insensitive to intent, and difficult to control. Not much is known about how it relates to awareness. The evidence suggests that the controlled component is sensitive to conscious intentions (i.e., task goals), and it is inefficient and effortful in the sense that time and focused attention are needed.

These studies depict automatic and controlled processes very differently than a task dissociation approach would. By identifying processes with tasks, the task dissociation approach assumes any variable that affects an implicit test affects automatic processing, and variables that affect an explicit task affect controlled processing. Given all of the differences between tasks that are confounded with the automatic-controlled distinction, this mapping creates a number of problems. The studies described here show that both automatic and controlled components of responses can be affected in different ways within the same task, overcoming these confounds.

Given that process dissociation estimates have been validated in several studies as useful measures of automatic and controlled components of responses, they can also be used to explain findings that would otherwise be puzzling. For instance, Lambert et al. found that racial bias in an impression formation task surprisingly increased when subjects were told that after the task they would share and discuss their responses with other research subjects (Lambert, Cronen, Chasteen, & Lickel, 1996). The finding was counter-intuitive because most theories of modern prejudice assume that people minimize prejudice when they expect their responses to be observed. To understand the mechanisms behind this effect, a subsequent study replicated the effects, replacing the impression formation task with a weapon bias task (Lambert et al., 2003). As in the previous study, subjects made more stereotypical responses when they anticipated their responses would be public, compared with private. Process dissociation analyses helped shed light on why. Subjects who anticipated having to share their responses showed poorer control over their responses, but no difference in automatic influences of stereotypes. In the anticipated public setting, subjects had the same degree of automatic influence, but it was less opposed by controlled responding, thus producing a net increase in stereotypical errors. Impairments in control were most pronounced for subjects who were anxious about the impending discussion. Together, these findings show how process dissociation analyses can be used to explain a counter-intuitive finding that would otherwise be puzzling.

The process dissociation approach also helps link concepts of automatic and controlled processing in social cognition to the ways that the terms

are understood in cognitive and neuro-scientific studies. Cognitive control in these fields is measured using tasks such as the anti-saccade task or the Stroop task, in which subjects see color words written in colored fonts and must name the font color and ignore the word meaning. According to a neuro-cognitive model, control requires monitoring for conflict between goals and responses, and directing attention to prevent or overcome automatic responses (e.g., Botvinick, Braver, Barch, Carter, & Cohen, 2001). These processes are associated with specific patterns of neural activity and with specific brain regions. They do not map on well to the type of 'control' measured by self-report attitude questionnaires, but they map on to cognitive control as assessed by process dissociation. In a series of studies, Amodio et al. studied event-related brain potentials that have previously been linked to conflict monitoring and cognitive control. They found that the same potentials were associated with process dissociation estimates of control in a weapons task, and that they were related to individual differences in motivations to control prejudice (Amodio, Devine, & Harmon-Jones, 2008; Amodio et al., 2004).

These studies help integrate control processes as studied in implicit social cognition with the kinds of control studied widely in other fields. More generally, process dissociation advances a broad construal of 'control' as acting in line with intentions and 'automatic influences' as those that drive behavior when control fails. Within social psychology, the procedure has been applied most often to questions of social memory, attitudes, and stereotyping. But recently, several studies have used process dissociation to test dual process theories of judgment and decision making, with encouraging results.

Decision making

Dual-process theories of decision making hold that people sometimes make decisions according to logical reasoning and sometimes through other means such as intuition, emotion, or heuristics (e.g., Kahneman & Frederick, 2002; Sloman, 1996; Stanovich & West, 2000). Studies often pit outcomes against each other to demonstrate separable processes. As an example, subjects in one study were asked to try to win money by drawing a red jellybean from a jar of mostly white jellybeans, without seeing what they were choosing (Denes-Raj & Epstein, 1994). Subjects could pick from a small bowl containing 1 red and 9 white jellybeans, or a large bowl containing 9 red and 91 white jellybeans. Although subjects knew the probability of winning was greater when picking from the smaller bowl (i.e., 0.10 versus 0.09), a majority of subjects nonetheless preferred to pick from the larger bowl with a greater absolute number of red jellybeans. Studies that pit logic against intuition demonstrate that people sometimes forgo the logical option for an intuitive choice. But this kind of procedure cannot measure the influences of logic and intuition

separately. The careful reader may have noticed that this design amounts to one half of the process dissociation procedure; it is an exclusion condition. If subjects respond based on logic they will make the normatively correct response. If logic fails and they respond based on intuition, they will make an intuitive 'error' (the scare quotes are meant to recognize that in many decision making tasks, there is no correct or incorrect response, but there are often normative benchmarks such as formal probability calculations, consistency, and so on).

A series of studies applying the process dissociation procedure to decision making showed that adding an inclusion condition helps separate rule-based reasoning from heuristic influences (Ferreira, Garcia-Marques, Sherman, & Sherman, 2006). Using several well-known heuristics and biases, these researchers created exclusion conditions similar to the problems typically used, in which rule-based reasoning and heuristics led to conflicting responses. As an example, following the jellybean selection task of Denes-Raj and Epstein (1994) subjects chose between an option with 2 winners out of 10, or an option with 19 winners out of 100. The intuitive appeal of the larger absolute number leads people to choose the 19 out of 100 option, but calculating the probabilities favors the 2 out of 10 option. They also added inclusion conditions in which rule-based and heuristic strategies led to the same response. For example, subjects could choose between an option with 2 winners out of 10 or an option with 21 winners out of 100. Here, both absolute numbers and probabilities favor the 21 out of 100 option. Applying the process dissociation procedure allowed them to estimate the use of rule-based reasoning and heuristic inferences. As expected, based on dual process theories, dividing attention while subjects made decisions reduced estimates of rule-based reasoning, but did not affect the use of heuristics. In another study, some subjects were encouraged to use a heuristic strategy by completing several initial problems in which heuristics led easily to the best answer. Compared with a control condition, the group encouraged to use heuristics showed greater estimates of heuristic processing, but no differences in rule-based reasoning.

Process dissociation methods have been used to separate logical reasoning and heuristic influences in several other decision-making contexts (Bishara, 2005; Fitzsimons & Williams, 2000; Liu & Johnson, 2005). For example, the anchoring heuristic occurs when arbitrary numbers bias numerical judgments (Tversky & Kahneman, 1974). In one study, real estate agents estimated the value of a real house after seeing all the relevant pricing information, inspecting the house, and seeing the listing price (Northcraft & Neale, 1987). Experimenters manipulated the listing price to create a high or low anchor value, with dramatic effects on the agents' estimates of the house's value. Bishara (2005) used process dissociation to separate anchoring effects into two components: controlled retrieval of information, and inferences based automatically on anchor values.

Logical or rule-based reasoning has frequently been contrasted with heuristic or intuitive decision making as polar opposites. The implication is a zero-sum relationship, in which increasing the use of rules must reduce the use of heuristics. The dissociations shown in these studies suggest instead that rule-based reasoning and heuristic inferences can make independent contributions to decision making. Given certain assumptions, the process dissociation procedure can be used to separate those contributions. The next section considers those assumptions and alternative approaches to consider if they are violated.

Assumptions and Alternatives

The equations used to derive process dissociation estimates rest on some assumptions that must be met for the estimates to be accurate. The procedure assumes, first, that automatic and controlled processes are statistically independent. Independence is important because estimates are based on multiplying and dividing probabilities, and joint or conditional probabilities can only be computed by simple multiplication or division if they are independent. A second assumption is that automatic and controlled processes exert the same degree of influence in inclusion and exclusion conditions. These assumptions have been controversial in some cases; thus, researchers using or interpreting process dissociation results should be aware of the issues and the context surrounding them. The assumptions have been discussed at length elsewhere (see Curran & Hintzman, 1995; Graf & Komatsu, 1994; Jacoby, Begg, & Toth, 1997; Jacoby & Shrout 1997). In the interest of brevity, my aim here is to place debates about process dissociation's assumptions in a larger context. It is important to recognize that any model makes assumptions. As an example, signal detection theory is commonly used to separate components underlying recognition memory, perception, and other kinds of performance. The model separates decisions into two components: sensitivity to evidence and response bias (how lenient or strict one's decision criterion is; MacMillan & Creelman, 1991). This commonly used model also assumes that its two processes – sensitivity and response bias – are independent. The model also makes other assumptions about underlying distributions and variances. Even commonly used inferential statistics such as analysis of variance (ANOVA) and regression models make assumptions such as normal distributions and interval scaling. Process dissociation's assumptions are not different in kind from these other models.

There are a few steps researchers can take to avoid violating assumptions. First, the assumptions may be violated when performance is very high or very low; researchers should avoid 'ceiling' and 'floor' effects by using tasks that are neither too easy nor too difficult. Second, although the assumptions cannot be directly proven, the plausibility of assumptions in a given paradigm can be tested experimentally. The most common way

to test the independence assumption is to arrange experiments in which automatic and controlled influences are expected, based on a theory, to show separate effects (that is, a dissociation). Divided attention, for example, may be predicted to reduce controlled estimates, whereas accessible information may be predicted to influence automatic estimates. If the assumptions are badly violated, dissociations would be difficult to detect. That is because if automatic and controlled influences were highly correlated then a variable that affects one would likely affect the other. All of the studies reviewed here have found such dissociations, supporting the assumptions.

Assumptions may be met to varying degrees, and if they are violated badly, researchers should choose alternative methods that make different assumptions. The choice is between procedures that make different assumptions because no procedure is assumption-free. It is sometimes tempting to think that by relying on task dissociations rather than a quantitative model, researchers can avoid making judgments about assumptions. But task dissociation assumes that tests are process-pure, and that the only difference between implicit and explicit tests is automatic versus controlled processing. Whereas process dissociation's assumptions may sometimes be violated, this assumption almost certainly is. Finally, an alternative approach is to arrange situations so that hypothesized automatic and controlled processes are pitted against each other in ways that should lead to opposing responses while holding other aspects of the design constant (Jacoby et al., 1989; Payne et al., 2007; Payne et al., 2005). Although this opposition procedure alone cannot provide quantitative estimates of automatic and controlled influences, it allows qualitative predictions and avoids unrealistic process-pure assumptions. Because different procedures make different assumptions, the strongest evidence is often converging evidence using multiple approaches.

Future Directions

Researchers in social psychology are using process dissociation techniques more frequently than ever before. Although this research is still developing, a few trends seem clear and likely to continue. One trend is the use of process dissociation in studying ever more topics where dual process theories have been developed. Dual process theories have been successful in integrating and explaining a wide range of findings in social psychology, and many emphasize a contrast between automatic and controlled processes (Chaiken & Trope, 1999). The techniques used to test these theories have been as diverse as the theories themselves, but in most cases, process dissociation could be a useful theoretical framework for thinking about automatic and controlled processes.

At one level of analysis, the meanings of the processes change from one topic to another. The controlled process may represent recollection in a memory theory, reasoning in decision-making theory, and perception of

weapons in another. The automatic process may be familiarity, heuristic inference, or stereotyping. These changes in specific meanings are as they should be because what it means to control thought and behavior in a memory test is different than what it means in a perception test. But at a broader level, process dissociation provides a unifying framework by defining control as the ability to respond in line with intent, and defining automaticity as influences that drive responses when control fails. Dual-process theories have frequently been tested using the task dissociation approach. But researchers are using a process dissociation approach more and more as an alternative to, or in addition to task dissociations. Because process dissociation offers a means of estimating processes within a task, and it offers a relatively simple way to conceptualize automatic and controlled processes, the trend is likely to continue.

Understanding automatic and controlled influences as distinct bases for responding has practical implications that future research should pursue. In some situations, the main cause of an error or bias might be an exceptionally strong automatic impulse, whereas in other cases, the problem might be a particularly poor ability for control. These situations call for different kinds of remedies aimed at either reducing the automatic impulse or boosting control (Stewart & Payne, 2008).

The focus in most of the studies described above was to estimate the influence of particular processes so that they could be studied. Another important set of questions, however, concerns how those processes are related to each other. As an example, dual-process theories in social psychology are often described as a two-step correction sequence. Some idea or impulse is automatically activated: 'Snails – that's disgusting!' and then may be overcome in a controlled editing process – 'I mean, the escargot looks lovely but I think I'll have a salad.' This 'late correction' model as described by Jacoby, Kelley, and McElree (1999) makes controlled responding conditional upon inhibition of an automatic impulse. That relationship contrasts with 'early selection' models in which automatic influences bias responses only when control fails. That is, automatic bias is conditional on control. Most dual-process theories provide a verbal description of how processes interact with each other. They may say, for instance, that multiple processes can be carried out in parallel, or that one process must be completed before another can begin, or that processes are independent but interacting. These verbal descriptions, though, are difficult to distinguish empirically.

Variations on process dissociation models have been developed recently to test these different arrangements. Alternative models can be compared using multinomial process modeling (Batchelder & Riefer, 1999), which generates process estimates, and also statistically tests how well the model fits the data. Figure 2 displays how early selection and late correction models can be represented as multinomial models. Lindsay and Jacoby (1994) found that whereas memory studies are well described by an early

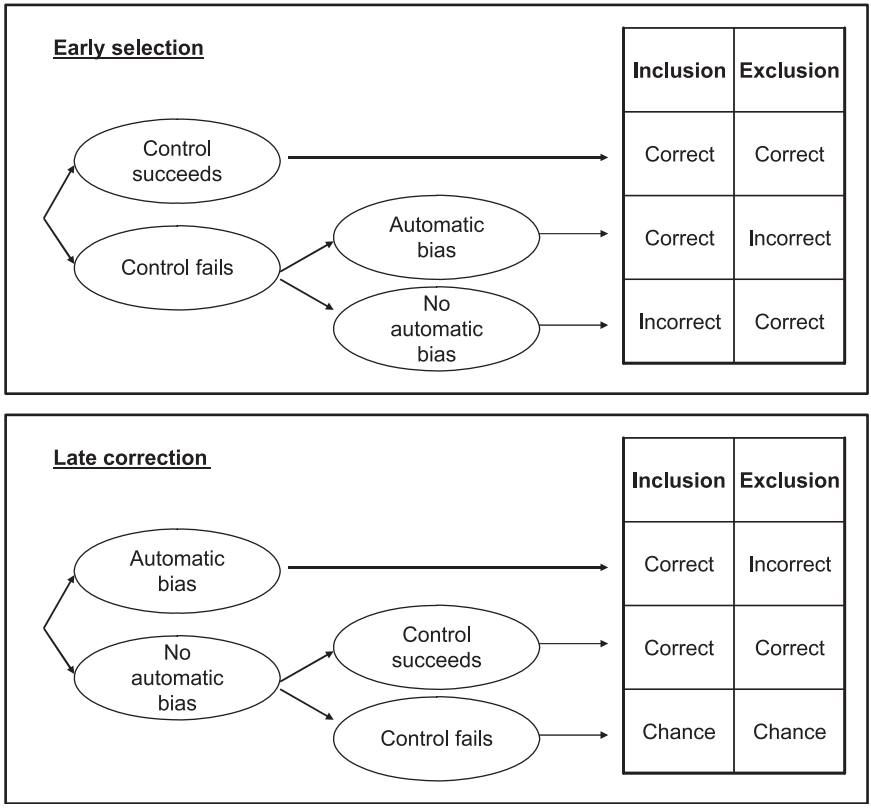

Figure 2 Early selection model (top) and late correction model (bottom) represented as multinomial models.

selection model, performance on the Stroop color-naming task is better described by a late correction model. In that model, only when automatic word reading was inhibited did controlled color naming drive responses. When results from the weapon identification task are compared using these models, results consistently favor the early selection model (Payne, Jacoby, & Lambert, 2005). Only when a person is unable to control responses does automatic stereotyping drive responses. Comparing models allows a test of which process is conditional upon the other.

More elaborate models can also be developed. Jacoby et al. (2005), for example, developed a memory model that included two automatic components in addition to controlled recollection to account for age differences in memory interference. In one process, misleading information can automatically capture responses. If not captured, recollection can proceed, but if recollection fails, subjects respond on the basis of automatically accessible information. This model combines early selection and late

correction models. Conrey et al. (2005) developed a model for implicit attitude and stereotyping tasks that includes four processes: an automatic activation process, a controlled discrimination process in which a person distinguishes between correct and incorrect responses, a controlled process in which people overcome conflicts between automatically activated and correctly discriminated responses, and a guessing process that drives responses when both automatic activation and controlled discrimination fail.

As the number of processes posited by a model increases, researchers have to weigh the benefits of additional explanatory power against the costs of added complexity and lost parsimony (Bishara & Payne, forthcoming; Payne & Jacoby, 2006). But by comparing models, multinomial models offer a powerful way to test theories about how automatic and controlled processes interact. Future research could more fully integrate the kinds of processes identified in these formal models (e.g., capture, discrimination, and accessibility bias) with the intuitive processes described in verbal models (e.g., implicit attitudes and correction). Doing so would increase the precision of verbal theories and the generality of formal models. These models have come a long way from Freud's use of verbal slips to reveal the unconscious, but they still estimate automaticity from errors. Combined with the tools of process dissociation and related models, mistakes disclose both what lies beneath the surface of consciousness, and how people try to maintain control.

Short Biography

Keith Payne's research focuses on social cognition – the processes underlying thinking and feeling about people and socially important issues. He is especially interested in the unintended and the unconscious. Here are some of the questions that motivate his research.

Why do people sometimes act in prejudiced ways even when they intend to be fair? How do people set aside unwanted biases and act the way they want to? And what happens when that process breaks down? Do you have any opinions or beliefs that you don't know about? If so, how can we measure them? And if we did, what would they really mean? Payne studies these questions using a variety of methods, ranging from behavioural experiments to mathematical modeling to neuroimaging. Before coming to the University of North Carolina at Chapel Hill, where he currently teaches, he taught at Ohio State University. He holds a PhD in Psychology from Washington University in St. Louis.

Endnotes

* Correspondence address: Department of Psychology, Campus Box # 3270 Davie Hall, University of North Carolina, Chapel Hill, Chapel Hill, NC 27599, USA.
Email: payne@unc.edu

[1] http://www.whitehouse.gov/news/releases/2005/05/20050524-3.html

References

Amodio, D. M., Devine, P. G., & Harmon-Jones, E. (2008). Individual differences in the regulation of intergroup bias: The role of conflict monitoring and neural signals for control. *Journal of Personality and Social Psychology*, **94**, 60–74.

Amodio, D. M., Harmon-Jones, E., Devine, P. G., Curtin, J. J., Hartley, S. L., & Covert, A. E. (2004). Neural signals for the detection of unintentional race bias. *Psychological Science*, **15**, 88–93.

Bargh, J. A. (1989). Conditional automaticity: Varieties of automatic influences in social perception and cognition. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended Thought* (pp. 3–51). New York, NY: Guilford Press.

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

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

Bishara, A. J. (2005). *Anchoring is no illusion*. Unpublished doctoral dissertation, St. Louis, MO: Washington University.

Bishara, A. J., & Payne, B. K. (forthcoming). Multinomial process tree models of control and automaticity in weapon misidentification. Manuscript submitted for publication.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, **108**, 624–652.

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

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

Crosby, F., Bromley, S., & Saxe, L. (1980). Recent unobtrusive studies of Black and White discrimination and prejudice: A literature review. *Psychological Bulletin*, **87**, 546–563.

Curran, T., & Hintzman, D. L. (1995). Violations of the independence assumption in process dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21**, 531–547.

Denes-Raj, V., & Epstein, S. (1994). Conflict between intuitive and rational processing: When people behave against their better judgment. *Journal of Personality and Social Psychology*, **66**, 819–829.

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

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology*, **33**, 510–540.

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

Fazio, R. H., Jackson, J. R., & Dunton, B. C. (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.

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.

Fitzsimons, G. J., & Williams, P. (2000). Asking questions can change behavior: Does it do so automatically or effortfully? *Journal of Experimental Psychology: Applied*, **6**, 195–206.

Govorun, O., & Payne, B. K. (2006). Ego depletion and prejudice: Separating automatic and controlled components. *Social Cognition*, **24**, 111–136.

Graf, P., & Komatsu, S. (1994). Process dissociation procedure: Handle with caution. *European Journal of Cognitive Psychology*, **6**, 113–129.

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. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, **74**, 1464–1480.

Hastie, R., & Kumar, P. A. (1979). Person memory: Personality traits as organizing principles in memory for behaviors. *Journal of Personality and Social Psychology*, **37**, 25–38.

Hense, R. L., Penner, L. A., & Nelson, D. L. (1995). Implicit memory for age stereotypes. *Social Cognition*, **13**, 399–415.

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin*, **31**, 1369–1385.

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

Jacoby, L. L., & Shrout, P. E. (1997). Toward a psychometric analysis of violations of the independence assumption in process dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **23**, 505–510.

Jacoby, L. L., Begg, I. M., & Toth, J. P. (1997). In defense of functional independence: Violations of assumptions underlying the process–dissociation procedure? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **23**, 484–495.

Jacoby, L. L., Bishara, A. J., Hessels, S., & Toth, J. P. (2005). Aging, subjective experience, and cognitive control: Dramatic false remembering by older adults. *Journal of Experimental Psychology: General*, **134**, 131–148.

Jacoby, L. L., Kelley, C., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology*, **56**, 326–338.

Jacoby, L. L., Kelley, C. M., & McElree, B. D. (1999). The role of cognitive control: Early selection versus late correction. In S. Chaiken & Y. Trope (Eds.), *Dual-process Theories in Social Psychology* (pp. 383–400). New York, NY: Guilford Press.

Jacoby, L. L., Toth, J. P., & Yonelinas, A. P. (1993). Separating conscious and unconscious influences of memory: Measuring recollection. *Journal of Experimental Psychology: General*, **122**, 139–154.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin & D. Kahneman (Eds.), *Heuristics and Biases* (pp. 49–81). New York, NY: Cambridge University Press.

Lambert, A. J., Cronen, S., Chasteen, A. L., & Lickel, B. (1996). Private versus public expressions of racial prejudice. *Journal of Experimental Social Psychology*, **32**, 437–459.

Lambert, A. J., Payne, B. K., Ramsey, S., & Shaffer, L. M. (2005). On the predictive validity of implicit attitude measures: The moderating effect of perceived group variability. *Journal of Experimental Social Psychology*, **41**, 114–128.

Lambert, A. J., Payne, B. K., Shaffer, L. M., Jacoby, L. L., Chasteen, A., & Khan, S. (2003). Stereotypes as dominant responses: On the 'social facilitation' of prejudice in anticipated public contexts. *Journal of Personality and Social Psychology*, **84**, 277–295.

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.

Liu, S. S., & Johnson, K. F. (2005). The automatic country-of-origin effects on brand judgments. *Journal of Advertising*, **34**, 87–97.

MacMillan, N. A., & Creelman, C. (1991). *Detection Theory: A User's Guide*. New York, NY: Cambridge University Press.

Macrae, C. N., Schloerscheidt, A. M., Bodenhausen, G. V., & Milne, A. B. (2002). Creating memory illusions: Expectancy-based processing and the generation of false memories. *Memory*, **10**, 63–80.

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

Northcraft, G. B., & Neale, M. A. (1987). Experts, amateurs, and real estate: An anchoring-and-adjustment perspective on property pricing decisions. *Organizational Behavior and Human Decision Processes*, **39**, 84–97.

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. (2005). Conceptualizing control in social cognition: How Executive Functioning Modulates the Expression of Automatic Stereotyping. *Journal of Personality and Social Psychology*, **89**, 488–503.

Payne, B. K., & Jacoby, L. L. (2006). What should a process model deliver? *Psychological Inquiry*, **17**, 194–198.

Payne, B. K., Burkley, M., & Stokes, M. B. (2008). Why do implicit and explicit attitude tests diverge? The role of structural fit. *Journal of Personality and Social Psychology*, **94**, 16–31.

Payne, B. K., Jacoby, L. L., & Lambert, A. J. (2004). Memory monitoring and the control of stereotype distortion. *Journal of Experimental Social Psychology*, **40**, 52–64.

Payne, B. K., Jacoby, L. L., & Lambert, A. J. (2005). Attitudes as accessibility bias: Dissociating automatic and controlled components. In R. Hassin, J. Bargh & J. Uleman (Eds.), *The New Unconscious*. New York, NY: Oxford University Press.

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measure. *Journal of Personality and Social Psychology*, **89**, 277–293.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **13**, 501–518.

Schumann, H., Steeh, C., & Bobo, L. (1985). *Racial Attitudes in America*. Cambridge, MA: Harvard University Press.

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.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic 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.

Srull, T. K., & Wyer, R. S. (1989). Person memory and judgment. *Psychological Review*, **96**, 58–83.

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

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

Tulving, E. (1985). How many memory systems are there? *American Psychologist*, **40**, 385–398.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, **185**, 1124–1130.

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

Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology*, **72**, 262–274.

Wyer, R. S., & Carlston, D. E. (1979). *Social Cognition, Inference, and Attribution*. Hillsdale, NJ: Erlbaum.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, **46**, 441–517.

