personality measures. *European Journal of Psychological Assessment, 21*(2), 100–107.

Sherman, S. J., Presson, C. C., Chassin, L., Rose, J. S., & Koch, K. (2003). Implicit and explicit attitudes toward cigarette smoking: The effects of context and motivation. *Journal of Social and Clinical Psychology, 22*, 13–39.

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics, 69*, 99–118.

Sriram, N., Greenwald, A. G., & Nosek, B. A. (2006). Scale invariant contrasts of response latency distributions. Unpublished manuscript.

Steffens, M. C. (2004). Is the Implicit Association Test immune to faking? *Experimental Psychology, 51*(3), 165–179.

Steffens, M. C., & Buchner, A. (2003). Implicit Association Test: Separating transsituationally stables and variable components of attitudes toward gay men. *Experimental Psychology, 50*(1), 33–48.

Steffens, M. C., & Plewe, I. (2001). Items' cross-category associations as a confounding factor in the Implicit Association Test. *Zeitschrift für Experimentelle Psychologie, 48*, 123–134.

Teachman, B. A., Gregg, A. P., & Woody, S. R. (2001). Implicit associations for fear-relevant stimuli among individuals with snake and spider fears. *Journal of Abnormal Psychology, 110*(2), 226–235.

Teachman, B. A., & Woody, S. R. (2003). Automatic processing in spider phobia: Implicit fear associations over the course of treatment. *Journal of Abnormal Psychology, 112*(1), 100–109.

Teachman, B. A, Gapinski, K. D., Brownell, K. D., Rawlins, M., & Jeyaram, S. (2003). Demonstrations of implicit anti-fat bias: The impact of providing causal information and evoking empathy. *Health Psychology. 22*(1), 68–78.

Teige, S., Schnabel, K., Banse, R., & Asendorpf, J. B. (2004). Assessment of multiple implicit self-concept dimensions using the Extrinsic Affective Simon Task (EAST). Unpublished manuscript.

Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge, MA: Bradford Books.

Wigboldus, D. (2004). Single-target implicit associations. Unpublished manuscript.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review, 107*(1), 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*(2), 262–274.

Wittenbrink, B., Judd, C. M., & Park, B. (2001). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality and Social Psychology, 81*(5), 815–827.

# 7

# Automatic and Controlled Components of Social Cognition: A Process Dissociation Approach

B. KEITH PAYNE and BRANDON D. STEWART

In a famous description of unintended behavior, William James (1890) noted that, "Very absent-minded persons in going to their bedroom to dress for dinner have been known to take off one garment after another and finally to get into bed, merely because that was the habitual issue of the first few movements when performed at a later hour," (p. 115). This kind of absent-mindedness, or something like it, will probably strike most readers as all too familiar. Compare this with a description by the neurologist Francois Lhermitte of a patient nearly a century later. When shown into a bedroom, the patient "immediately began to get undressed. He got into bed, pulled the sheet up to his neck, and prepared to go to sleep," (Lhermitte, 1986, p. 338). We can be sure that this behavior is more exceptional than James' absent-mindedness, because the bedroom belonged to Lhermitte.

The patient had a massive surgical lesion of the left frontal lobe, a brain region critical for strategic planning and control of actions. This and other similar patients suffered from what Lhermitte termed "environmental dependency syndrome." Although their behaviors were coordinated and complex (not simple reflexes), they were under the control of the environment to a striking extent. They behaved in accord with whatever environmental cues caught their eye. In one demonstration, upon noticing that a syringe had been laid out, a patient picked it up and began to give the neurologist an injection. The reader is not told whether the injection was carried out. In another, even more macabre investigation, the patient was led to a table where a pistol and some bullets had been placed. Without pause or comment, he picked up the pistol, pulled back the magazine, and loaded it. At this point we read, "The experiment was then stopped," (p. 338). These were not the momentary lapses of voluntary control that we all experience, but profound absences. Lhermitte described it as a disorder of autonomy.

No less intriguing are the struggles of individuals with alien hand syndrome.

Because of damage to the frontal lobes or the corpus callosum connecting left and right hemispheres, these patients experience "autonomous" actions by one or more limbs. The actions are autonomous in the sense that the patient cannot voluntarily control them, nor do they experience the behaviors (nor even the limb itself) as their own. Like Lhermitte's patients, the "alien" actions are usually triggered by environmental cues. For example, patients may experience a "struggle between the hands as each attempts to answer the telephone," or even more drastically, "one hand tried to turn left when the other hand tried to turn right while driving a car," (Doody & Jankovic, 1992; p. 807). Marchetti and Della Sala (1998) report a patient who, "at dinner, much to her dismay saw her left hand taking some fish bones from the leftovers and putting them into her mouth," (p. 196).

These bizarre conditions seem so strange because they are dissociations between functions that usually work seamlessly together, with intentions reining in the automatic when it goes astray. Dissociations are so informative because so much can be learned about the structure of a system from where its fault lines lie. When a crystalline icicle shatters on the floor it comes apart in remarkably regular patterns, very different from the shattering of an egg. Their inner organization reveals itself in the ways they come apart. Neurologists have been studying dissociations caused by anatomical lesions for more than a century. Neurological dissociations have more recently attracted the attention of social psychologists for their potential to shed light on the control of complex thought processes such as those involved in social behavior (e.g. Bargh, 2005; Beer et al., 2003; Wegner, 2002).

Although anatomical dissociations map out these fault lines in vivid detail, dissociations in the behavior of healthy men and women can be just as informative. One way to find the seams between the intentionally controlled and the uncontrolled is to compare explicit and implicit tasks. Amnesiac patients, for example, show profound deficits when tested using explicit memory tests, which ask the person to intentionally retrieve a memory. Yet when tested with implicit memory tests, their performance shows effects of past experience without the intent to remember or the feeling of remembering (Shimamura, 1986). But it is not just amnesiac patients who show this difference. Normal healthy college students also show dissociations between implicit and explicit memory tests (Jacoby & Dallas, 1981). Factors that affect performance on one kind of test often have no impact on the other (see Roediger & McDermott, 1993). With the help of implicit and explicit tasks, researchers can probe the distinctions between mental processes without relying on rare, unfortunate cases of brain damage. They can infer the inner structures without waiting for them to crack. The fact that healthy men and women show dissociations between intentional and unintentional aspects of behavior forces us to ask about the fault lines in ordinary thought. Do we all have little disorders of autonomy? In some ways, we shall see, the answer is yes.

The cases reviewed here are all dissociations between intentional and unintentional processes, each at different levels of analysis. The dissociations become increasingly subtle, but no less intriguing as we zoom in from the macro scale to the micro. Lhermitte's frontal patients showed a dissociation at the level of the *whole person*. Most people act with intentional control most of the time, but these patients did not. Cases of alien hand represent dissociations within a person, but between limbs. One hand is under voluntary control, and the other is not. At a still finer grain are dissociations between tasks. The same person acts differently on the two tasks because the tasks draw on different processes.



We can zoom in further. This chapter focuses on a process dissociation approach – a technique for separating intentional and unintentional contributions to the same behavior performed by the same person at the same time. Imagine that a man is walking toward you on the street. Although you do not recognize him, you have a bad feeling about him and decide to cross the street. If you had been able to intentionally retrieve the fact that you saw him yesterday on a wanted poster, you would have had an even better basis for your decision, and might have hurried faster, or called the police afterward. But even without being able to remember the poster, the vague sense of threat can also guide your response. Here intentional and unintentional forms of learning could both feed into the response, in varying degrees. The fact that you can have one without the other illustrates that they are separable. Just as lesion studies allow dramatic dissociations based on anatomy, the process dissociation approach seeks to separate intentional and unintentional influences, even though they normally operate together.

The goal of this chapter is to overview the logic behind the process dissociation approach, what it measures, and what it does not. We will describe some of the many different topics where process dissociation has been used, and the kinds of insights it can provide for social psychology. Along the way, we will discuss the assumptions that must be met to properly use the procedure, and we will see how this way of thinking about the automatic–controlled distinction compares and contrasts with other prominent approaches. Although the procedure was developed in the context of memory research (Jacoby, 1991; Jacoby, Toth, & Yonelinas, 1993) we will focus on applications outside of pure memory research, emphasizing instead social cognition and behavior. Interested readers are referred to Jacoby (1998) for an overview of the procedure as developed in memory research and Yonelinas (2002) for a theoretical review of dual-process theories of memory using process dissociation and related methods.

## PROCESS DISSOCIATION IN SOCIAL COGNITION

A study by Hense, Penner, and Nelson (1995) marked a point of departure from pure memory research to social memory. We will use this experiment to illustrate how process dissociation can be used to study social memory distortions. Participants were asked to remember a list of traits that described elderly or young individuals. Each trait was stereotypical for either old or young people. After studying the traits, participants were asked to recall the traits that described the older and younger target persons under two sets of instructions. In the *inclusion* condition, participants were asked to respond with the trait they had studied or, if they could not remember the trait, respond with the first word that came to mind. In this condition, responses can be driven by either intentionally retrieved

memories or by automatic forms of memory that cause certain thoughts to come to mind more readily. In the inclusion condition intentional and unintentional forms of memory work in concert. In the *exclusion* condition, participants were asked to respond with a new trait that was *not* studied. If they remembered having studied a trait, they could successfully avoid reporting it. However, if they failed to consciously remember a trait, but it unintentionally came to mind, they would be likely to report it. This condition pits intentional and unintentional forms of memory against each other.

By comparing performance in inclusion and exclusion conditions, the influence of intentional and unintentional uses of memory can be estimated. To the extent that a person responds with the correct trait when they try to, but also withholds it when they try to, memory is under intentional control. Because in this paradigm subjective awareness of the memory is the basis for controlling memory reports, consciousness of the memory can also be inferred. But to the extent that past experience influences performance *regardless* of what participants are trying to do, they are being unintentionally influenced by memory. Jacoby and colleagues have termed this the "logic of opposition." The equations for estimating these influences will be discussed in a later section.

Hense and colleagues found that stereotypical traits had a selective influence on the unintentional use of memory without affecting controlled recollection. Stereotype-consistent traits such as slow and frail came to mind easily and biased memory reports whether they were trying to retrieve them or trying not to retrieve them. Consciously controlled memory was affected by a divided attention task, but this was independent of the stereotyping effect.

This study revealed an important dissociation. Within the single activity of remembering traits, intentional and unintentional forms of memory both fed into responses. Automatic influences reflected participants' own stereotypes. In contrast, consciously controlled memory was affected by divided attention. This study helped to characterize the mechanisms behind stereotypical memory biases that have been known for years (for a review see Stangor & McMillan, 1992). Social psychology has a tradition of emphasizing dissociations between underlying processes that would seemingly go together. Devine's (1989) important demonstration that the automatic activation of stereotypes could be dissociated from their use provides a clear example in the domain of stereotyping. Whereas Devine contrasted results from a priming task with results from self-report measures, the study by Hense and colleagues (1995) contrasted two aspects of a single act of remembering.

Even within the same task, intentional and unintentional uses of memory operate very differently. Using a memory paradigm in which race stereotypes could bias memory, we have explored the ways that components of memory are related to subjective experience (Payne, Jacoby, & Lambert, 2004). We found that people's subjective sense of confidence in their memories was well-attuned to recollection. When they were consciously recollecting the past, they expected to be right, and they usually were. When they had no recollection, they expected to be wrong, and they usually were. But confidence was not at all tuned in to the automatic influences of stereotypes on memory. When conscious memory failed, feelings of confidence gave no clues about whether an automatic bias was at work. This asymmetry of subjective awareness has important consequences for avoiding stereotypical biases.



In one condition of the study we required participants to answer every memory question, whether they remembered the correct answer or not. In the other condition, we instructed participants to answer only if they believed their answer to be correct. What effect did the freedom to choose have on memory reports? That depended on what aspect of memory is looked at. When it came to overall accuracy, the freedom to choose helped memory. Because subjective confidence was well-tuned to recollection, participants were able to avoid answering questions they would get wrong, and they chose to answer those items they were likely to get correct.

But the story was very different when it came to the influence of stereotypes. Memory reports were biased toward stereotype-consistent memory errors. But critically, the pattern of stereotype-consistent bias was just as strong when participants were allowed to choose as when they were required to answer every question. When it came to recollection, as in many areas of life, awareness bestowed control. But when stereotypes came readily to mind, Jamal was likely to become an athlete, and Walter was likely to become a politician independent of intent, independent of subjective experience, and independent of the choice to keep quiet.

These studies moved from basic memory research to social memory biases. But the process dissociation procedure can also be used in contexts completely unrelated to memory. Imagine now that the person approaching on the street looks like a suspect and you are a police officer. The suspect pulls an object from his pocket. What should you do? This was the decision facing four police officers on February 4, 1999 when they confronted Amadou Diallo outside his apartment in New York City. Their decision turned out to be wrong. Diallo was killed despite being unarmed and having nothing to do with the crime the officers were investigating. The case sparked public outrage and charges of racism because Diallo was Black. However, as in so many instances in daily life, there was no "control group" to gauge the impact of race. Since that incident, many similar cases have been reported in the national press. Many inside and outside of law enforcement have asked themselves, "What would I do in that situation?" We have used the process dissociation procedure to study what people actually do in that kind of split-second decision situation.

To see how process dissociation can be informative here, consider the different possible scenarios that might confront the officer. In one scenario, a Black suspect pulls a gun. Here, the "correct" response (at least for the purposes of our thought experiment) is to "shoot." That response might come about from two routes. One is an intentionally controlled response, in which you shoot because you mean to shoot. The second route is an unintentional or automatic response evoked by race stereotypes about the suspect. Because intentional and unintentional processes are working together here, we cannot tell them apart. Now consider a different scenario, in which the Black suspect holds only a wallet (as, in fact, Diallo held). Here an intentionally controlled response would be to not

shoot. But an automatic response based on racial stereotypes might still lead one to pull the trigger. To the extent that a person systematically shoots despite not intending to, we know that an unintentional process is at work. But to the extent that a person shoots when they mean to, and not otherwise, we know that they are in control.

Our laboratory has conducted a number of studies examining these sorts of scenarios. We have used a simple procedure in which the faces of Black and White individuals are flashed on a computer screen just before pictures of handguns and hand tools are shown. All of the pictures are presented long enough to see clearly and all of the items are easy to identify. We show pairs of objects on the screen and ask participants to respond by pressing a "gun" key or a "tool" key as quickly as possible. Across many studies, we have found a clear and consistent tendency to mistakenly respond "gun" when a Black face is flashed. That tendency increases as participants are rushed to respond faster and faster (Payne, 2001; Payne, Lambert, & Jacoby, 2002).

The similarity of this task to other implicit measures makes it tempting to think of the bias as an "automatic effect." That way of thinking about it is consistent with the task dissociation approach, in which implicit tasks are identified with automatic processes and explicit tasks are identified with controlled processing. But that way of thinking overlooks the possibility of both automatic and controlled processes feeding into responses.

To reveal the distinct processes seamlessly guiding responses, we compared responses when automatic and controlled processes were acting in concert versus when they were opposed. When both automatic and controlled processes pointed to the "gun" response (an inclusion condition), the probability of a gun response was .75. This can be formalized as the sum of controlled processing and automatic processing when control fails: Control + Automatic × (1 – Control). Because the values here are probabilities, the absence of an event can be easily expressed as (1 – the probability of that event). On Black–tool trials where automatic stereotyping would lead to a "gun" response but controlled responding would lead to a "tool" response, participants still responded "gun" with a probability of .37. This represents the tendency for an automatic bias to drive responses in the absence of control: Automatic × (1 – Control).

To estimate how much of this behavior was due to intentional control, we took the difference in performance when both automatic and controlled processes favored a response, versus when automatic but not controlled processes favored it. This gives a control estimate of .38 (= .75 – .37).[1] It is important to notice that control, as conceptualized here, is not a reaction to a stereotypical thought. It is not thought suppression or an after-the-fact editing of responses. It is the ability to focus attention, thought, and action on goal-relevant behaviors independent of automatic distractions. If control were a perfect 1.0, actions would be determined completely by intentions. In this experiment the value was much lower, allowing for other factors to unintentionally influence behavior. With simple constraints like speeded responding, it is not difficult to see momentary disorders of autonomy crippling good intentions.

To estimate the automatic effect of stereotyping we looked at how often participants responded in line with the stereotype even when they intended to respond otherwise (.37). Although this value reflects an unintentional process, it is an underestimate because it is the joint probability that an automatic process was at work *and* that control failed. The more control a person exerts, the more this value underestimates the automatic stereotyping effect. Under the assumption that automatic and controlled processes are independent, we can correct for this underestimation, dividing by the probability that control has failed (1 – .38 = .62). This yields an automatic estimate of .60.



Using this procedure, we found that requiring fast responses dramatically reduced control, compared to a group that responded at their own pace. But fast responding had no effect on the automatic process, as we would expect because automatic processes operate quickly and require little capacity. On the other hand, the automatic estimate was affected by the Black versus White faces. The automatic tendency toward "gun" responses was higher on Black compared to White trials.

These dissociations are important because they enhance our ability to answer questions about how unintended racial biases influence people. When studying how stereotypes have their effects, a commonly asked question is, "is it an automatic or a controlled effect?" And a common way to test that question is to impose a cognitive load, to rush responding, or look for motivational differences. If cognitive load, rushed responding, or low motivation interferes with the effect, it is inferred to be resource dependent and therefore likely controlled. In contrast, if these variables have no effect (or increase the stereotyping effect), it is inferred that the effect is automatic, because it is not dependent on the investment of cognitive resources.

Even a relatively "simple" behavior is complex. It is difficult to find any behavior that does not include some amalgam of processes with automatic features and controlled features (Bargh, 1989). It therefore becomes important to separate complex behaviors into more basic components. The process dissociation approach shifts the question from "automatic *or* controlled?" to "what combination of automatic *and* controlled?" The focus changes from labeling a phenomenon to taking apart its component processes. Importantly, process dissociation also provides a measurement model for quantifying and summarizing those components.

This new question becomes particularly interesting when automatic and controlled components behave in different ways. We have uncovered several variables that all impact or correlate with people's judgments in the weapon scenarios. Each of these selectively affects either automatic stereotyping or intentional control. For instance, we recently investigated the effects of self-regulation depletion on stereotyping (Govorun & Payne, 2006). Based on prior findings that exerting self-control in one domain reduces self-regulation in a subsequent context (Muraven & Baumeister, 2000), we predicted that the depletion group would show reduced intentional control. We assigned one group to perform a boring but attention-demanding Stroop color naming task for a continuous 15 minute period. Following this tedious task, participants completed the weapon identification task as described above. The control group performed the Stroop task for only 30 seconds,

and then went on to the weapon task unfatigued. As predicted, the depleted group showed poorer control, but no differences in automatic stereotyping. For individuals with a stereotypical automatic bias, this reduction in control resulted in more stereotypical false "gun" responses. For these individuals, automatic stereotypes were left unopposed by intentional control.

Compare that manipulation with another one intended to influence stereotyping, a blatant warning that the weapon identification task measures racial stereotyping and that they should be careful to avoid stereotyping. Ironically, warned participants showed a more stereotypical pattern of mistakes than a control group (Payne, Lambert, & Jacoby, 2002). Was this because the ability to control responses was diminished, or because the warning made race accessible to the point that it increases the automatic impact of stereotypes? Our results suggest the latter – a blatant warning increased stereotyping via increasing automatic bias, having no effect on control.

From a distance, the warning study and the ego depletion study appear very similar. Compared to control groups, both warning about race stereotypes and depleting self-regulatory strength increased stereotyping as measured by errors. But these two superficially similar effects were driven by different mechanisms (see also Lambert, Payne, Jacoby, Shaffer, Chasteen, & Khan, 2003). In the warning study, stereotyping increased by boosting the automatic activation of stereotypic associations. In the depletion study, stereotyping increased by reducing control over behaviors.

The two studies we just described suggest a rather pessimistic outlook, because both manipulations increased stereotyping rather than decreasing it. It is worth considering, from both practical and scientific points of view, how stereotyping can be decreased below the baseline level of our control participants. Process dissociation allows us to track how the automatic and controlled components of behavior mediate that change. Stereotypical responding could be reduced either by reducing automatic bias or by increasing intentional control.

A recent study tested the utility of concrete action plans in overcoming the automatic impact of race stereotypes (Stewart & Payne, 2006). Previous work has shown that concrete action plans linking a specific environmental cue to an action can help people carry out their intentions more effectively (Gollwitzer, 1999). When applied to stereotypical weapon judgments, that idea might take the form of "when I see a Black person, I will respond 'tool'." However, this kind of plan would just replace one bias with another. We wanted to find out whether the action phase could be used to generate a thought that would counteract the influence of the stereotype without creating a new bias. To that end, we asked participants in one condition to form a plan so that whenever they saw a Black person, they would "think safe." In fact, that simple plan reduced the effect of stereotypes, compared to a control group who were asked to "think quickly" when they saw a Black person.

Did this simple treatment reduce stereotyping by increasing control or reducing automatic influences? Across three studies, we found that the thought-plan reduced the automatic influence of race, without altering intentional control. Moreover, the plan took effect in the first several trials, suggesting that it was very efficient and did not require extensive practice as other methods of altering automatic biases may (e.g. Kawakami et al., 2000).



## SOME WAYS TO THINK ABOUT AUTOMATIC AND CONTROLLED ESTIMATES

The meanings of the estimates generated by process dissociation depend on the kind of behavior that is being analyzed. In all of the studies reviewed so far, the automatic estimates represented unintentional influences of stereotypes. But as we moved from memory studies to perceptual judgments, the meanings of the controlled estimate changed more dramatically. In the memory studies, the controlled component was recollection: a consciously controlled use of memory. Having full access to the context and the details of an event allowed participants to use or not to use whatever came to mind. These shifts in the interpretations of process estimates are unavoidable, because people control their behavior in different ways depending on what they are doing.

In the weapon judgment studies, what is the best way to characterize the processes underlying intentional control? Concretely, the controlled component reflected the ability to respond based on one set of information (the features of the target items) and not another (racial stereotypes). Seen in this light, the weapon task is similar to other compatibility tasks such as the Stroop color naming task. In that task, subjects try to name the ink colors of words while ignoring the word itself. When the word is a color word, it becomes very difficult to name an ink if it is incompatible with the word meaning (e.g. the word red in green ink). In fact, in the Govorun and Payne study, performance on the Stroop task was significantly correlated with the controlled (but not the automatic) estimate from the weapons task.

### *Executive Function and the Control of Bias*

The Stroop task and similar interference tasks are usually understood as measures of executive control – the ways people direct their information processing and actions to keep them consistent with their goals. Executive control is believed to include subprocesses including selection of relevant information, inhibition of interfering information, and maintenance of the currently pursued goal (Baddeley, 1986; Kane, Bleckley, Conway, & Engle, 2001). The analogy between the weapon judgment scenario and the Stroop task suggests that control in the weapons task has something to do with devoting selective attention to one stream of inputs, while blocking out another.

A recent study supports this suggestion. Payne (2005) used an antisaccade task to test the idea that control in the weapons task relies on the executive processes of selective attention. The antisaccade is a well-established measure of attentional control often used in cognitive and neuroscience studies (Everling & Fischer, 1998). Think back to the last time you were in the middle of a talk or lecture, and someone mistakenly opened the door only to sheepishly realize that they were in

the wrong place. Try as we might to concentrate on the matters at hand, it is incredibly difficult not to turn and look at the interloper. The reason is that the orienting reflex which compels us to attend to new items in our surroundings is highly automatic. The ability to intentionally override such automatic reflexes has been used to measure executive control. The antisaccade task asks participants to avoid looking at an item that flashes abruptly on a computer screen. A distracting item (e.g. a red circle) appears on one side of the screen, and a target (e.g. a letter to be identified) flashes on the other side a fraction of a second later. Looking away from the distracter will enhance identification of the target, but looking at the distracter will interfere.

Research from our lab showed that performance on an antisaccade task was correlated with the controlled component, but not the automatic component, in the weapon judgment task (Payne, 2005). In contrast, the automatic estimate was correlated with two measures of implicit race attitudes, the implicit association test and evaluative priming. This dissociation reveals two very different kinds of processes that normally blend imperceptibly together as people make a single decision. The answer to the question "how would I act in that situation?" has at least two parts. The first depends on a person's automatic reactions to Black individuals. The second depends on the person's ability to engage executive control, that is, to keep their thoughts and actions on track rather than being swayed by accessible but inappropriate information.

These studies illustrate a range of processes that might be measured using process dissociation methods. Other research has found creative applications of the procedure outside of both memory and stereotyping. As one illustration, Fitzsimons and Williams (2000) used a modification of the process dissociation procedure to investigate the *mere measurement effect*. The mere measurement effect is the finding that simply asking a person about how likely they are to perform a behavior in the future actually increases the likelihood that they will perform that behavior (Morwitz, Johnson, & Schmittlein, 1993; Sherman, 1980). Fitzsimons and Williams (2000, study 1) asked one group of participants how likely they would be to choose a new brand of candy bar, whereas the control group was not asked about the candy bar. The mere measurement effect suggests that the group who was asked would be more likely to choose the candy bar than the control group. In addition to the intent question, participants were given information suggesting that they were more or less likely to actually receive the candy bar if they chose it. This manipulation was intended to manipulate the self-interest of participants. A rational (in the sense of self-interested) analysis would suggest that participants should be more likely to choose the candy bar when informed that they were likely to get the candy bar. By crossing the measurement of intent with the self-interest information, this study created conditions in which the mere measurement effect was congruent with self-interest, and conditions in which it was incongruent with self-interest.

This study showed that, indeed, participants asked about their likely choice more often chose the candy bar. Using a modified model based on the logic of process dissociation, these researchers separated two components of the effect. One component reflected how strongly the intent question influenced choice regardless of self-interest (akin to the automatic component in the other studies discussed here). The other component reflected the extent to which choices were guided by the self-interest. This is related to the controlled component we have discussed, but because the procedure pitted self-interest against the mere measurement effect, this component reflected self-interested decision-making. The intent question influenced choices largely irrespective of whether it was consistent with self-interest or not, an effect driven by the automatic component.



From the studies reviewed here, it is clear that the process dissociation procedure is not limited to measuring recollection and familiarity in the context of memory research where it was developed. Although extensions beyond memory research are a relatively new endeavor, the procedure is flexible, and can be adapted to any number of topics. The key to this flexibility is that process dissociation represents a general framework for thinking about intentional and unintentional processes. The basic logic of placing intended and unintended influences in concert and in opposition in order to disentangle them can be implemented across many, many domains.

It is important to note that the meaning of the processes measured depends entirely on the task being studied. Phrased another way, the meaning of the estimates depends on what processes are placed in concert and in opposition with each other. If implicit and explicit influences of memory are arranged in this way, the procedure can yield estimates of implicit and explicit memory. If automatic stereotyping and executive control are arranged in this way, the procedure can estimate these processes, and so on. As these contrasts illustrate, the logic of opposition creates a conceptual and methodological way to think about many different kinds of factors that might influence people either intentionally or regardless of intent. Nevertheless, it is often tempting to think of the automatic and controlled components as having fixed meanings, based on other prominent process distinctions that have been made in social cognition. In the following section we compare the process dissociation approach to some of these commonly invoked dimensions of automatic and controlled processing, highlighting the similarities and differences.

## Invariants and Particulars

Given all the different uses to which one might put the procedure, one might wonder if there is anything that the different automatic or controlled processes have in common (see Moors & De Houwer, Chapter 1 of this volume). Yet there appear to be certain properties consistently attached to controlled components, and certain properties attached to automatic components, across the many different domains. For example, the controlled component in memory studies requires attention, and is disrupted by distraction (Hense et al., 1995; Sherman et al., 2003). The controlled component in weapon identification studies requires time, and is disrupted by rushed responding (Payne et al., 2002). Control in the weapon studies is also influenced by motivations. Higher motivation to control prejudice is associated with more intentional control during judgments (Payne, 2005; Amodio et al., 2004). In contrast to these results for intentional control, these same studies

show that the automatic estimate appears unaffected by divided attention, occurs rapidly, and is not associated with motivations. Readers may have noticed that these are the defining characteristics usually invoked to distinguish automatic and controlled processing in general (Bargh, 1989; Posner & Snyder, 1975; Shiffrin & Schneider, 1977).

This is no accident. The fact that these properties remain attached to automatic and controlled estimates regardless of the specific task suggests something interesting. It suggests that regardless of *what* one intends to do, it is the constraining of behavior to intent that requires resources, time, and motivation. On the other hand, it is being pushed along regardless of intent that is quick, effortless, and easy, perhaps regardless of what force is doing the pushing.

In this light, the relationship between the process dissociation framework and other commonly studied forms of automatic and controlled behavior becomes clearer. To bring this relationship more sharply into focus, we will consider as examples two dual-process distinctions prominent in social psychology. The first is deliberative reasoning versus shallow heuristic-based reasoning. The second is the distinction between implicit and explicit attitudes. These distinctions form the crux of more than one prominent dual-process theory aimed at explaining why, when, or how automatic and controlled aspects of cognition combine to drive behavior. Each distinction could potentially be illuminated by a process dissociation approach, once it is understood what assumptions must be made and how process estimates could map on to these distinctions.

The contrast between deliberative reasoning and shallow inferences based on heuristics is at the heart of many dual-process theories (Chaiken, 1980; Fazio, 1990; Petty & Cacioppo, 1986). The process dissociation estimates discussed thus far in the context of memory and weapon identification studies seem far removed from this distinction. However, the conceptual approach can be applied here just as in these other areas of study. Perhaps the most closely related research is the work of Fitzsimons and Williams (2000) described above. In their studies they placed self-interest and the mere measurement effect in concert and in opposition to tease apart their separate contributions. It is easy to imagine other situations where careful thought would lead to one kind of outcome, and heuristics would lead to another. Here the two processes would be set in opposition. It is just as easy to imagine an experimental arrangement in which careful thought and heuristics would lead to the same response, thereby setting up an in-concert condition.

In fact, this kind of paradigm is routinely used in studies of attitudes and persuasion, although it may not be framed in these terms. A common procedure for studying the processes of attitude change is to manipulate argument strength. The assumption is that when people are thinking carefully they will be more persuaded by strong than weak arguments. This manipulation is sometimes crossed with a heuristic or cue, such as the number of arguments used or the attractiveness of the source making the arguments. The assumption is that when people are processing shallowly, they are more likely to be persuaded by these simple cues.

When argument strength and heuristic cues are fully crossed, the design creates some conditions in which both deliberate thinking and heuristic thinking would lead to agreement or disagreement with the message (i.e. the strong argument, positive cue cell and the weak argument, negative cue cell). Also created are conditions in which deliberative and heuristic thinking would lead to different outcomes (i.e., the strong argument, negative cue cell and the weak argument, positive cue cell). Because deliberative thinking and heuristic-based thinking can be arranged in this way, they can potentially be separated and quantified using the process dissociation approach. One estimate would represent the contribution of deliberate reasoning and the other would represent the contribution of heuristic reasoning. As in any new application of a model, validation tests would need to be performed to test whether the assumptions of the process dissociation procedure match the properties of the attitude change paradigm. The broader point is that process dissociation does not fix the meanings of the automatic and controlled estimates. It is instead a way to think about and quantify the contributions of different processes, as needed to answer specific questions.



In the studies of race biases in memory and perceptual identification described in the previous section, the automatic estimate can be said to reflect implicit attitudes or stereotypes. Implicit attitudes are commonly contrasted with explicit attitudes, which are the attitudes people overtly express when directly asked. However, in the studies described, automatic biases were contrasted not with explicit attitudes, but with the ability to intentionally control responses. By this approach an implicit attitude is the evaluation that drives responses when intentional control fails. For some purposes, however, researchers wish to separate implicit attitudes from explicit attitudes.

The typical way of separating implicit and explicit attitudes is by comparing an implicit measure and an explicit measure. If an outcome behavior correlates with an implicit measure but not an explicit measure, it is said to be the product of implicit or automatic processes. If a behavior correlates with an explicit measure but not an implicit measure, it is said to be the product of controlled processes. However, many different factors vary between implicit and explicit measures beyond automaticity and control. For example, how should one compare reaction times in word pairings to Likert scales? Even if they are standardized to the same scale, should they be interpreted in the same ways? What if one measure is more reliable or sensitive than the other? Early findings that implicit and explicit measures of attitudes toward the same topics tended to correlate weakly, if at all, led to a great deal of theorizing about whether implicit and explicit attitudes represent separate constructs (e.g., Fazio & Olson, 2003; Wilson, Lindsey, & Schooler, 2000). In our view, there are many reasons to expect the relationship between implicit and explicit measures to be weakened, even if they are tapping the same construct. Before we are able to fully address this issue of single versus separate constructs, it will be important to deal with other factors such as reliability (which tends to be lower for implicit measures; see Cunningham, Preacher, & Banaji, 2001) and the difficulties of comparing vastly different types of behavior on implicit versus explicit tasks.

Because of these limitations to the *task* dissociation approach, we have pursued a complementary approach within the *process* dissociation framework. It may be possible to design experiments in which intentional and unintentional

contributions estimated from the same behavior represent explicit and implicit attitudes. If so, this would allow intentional and unintentional aspects of attitudes to be compared on the same scale and within the same task, overcoming the problems of comparing across radically different measures. We next describe a newly developed method for measuring attitudes implicitly within the task dissociation framework, and then explore how that method might be expanded for a process dissociation analysis.

We have known for years that people sometimes misattribute their evaluative reactions from one source to another source (Dutton & Aron, 1974; Schwarz & Clore, 1983). For example, Murphy and Zajonc (1993) showed that flashing pleasant and unpleasant images before presenting an ambiguous Chinese pictograph influenced the way people evaluated the pictographs. In this study, the outcome of interest was not response times, but how pleasant or unpleasant participants found the pictographs. When the prime was pleasant, participants found the pictograph more pleasant; when the prime was unpleasant, participants found the pictograph less pleasant.

Although this effect has been widely known for more than a decade, there is an important implication that has gone unnoticed: This misattribution produces an indirect measure of individuals' attitudes toward the primes. If a particular prime item systematically causes participants to evaluate an ambiguous pictograph positively, it suggests a positive attitude toward the prime item. In the original Murphy and Zajonc (1993) procedure, the primes only affected judgments of the pictographs when they were flashed too briefly to be consciously identified. However, with some modifications to the procedure, we were able to produce strong misattributions even when the primes were plainly visible, and even when participants were blatantly warned against being influenced by them (Payne, Cheng, Govorun, & Stewart, 2005).

Using this modified procedure (which we refer to as an affect misattribution procedure), we found that the kinds of misattributions people made when primed with pictures of George W. Bush and John Kerry strongly predicted their explicitly rated attitudes toward the candidates ($r = .65$) and who they intended to vote for ($r = .58$). We expected high implicit–explicit correspondence in this domain because people are well aware of their political attitudes and quite willing to express them. In this case, whatever variance in the two measures did not overlap may have been caused by the problems of comparing across two very different kinds of measures.

In another study, we replaced the pictures of Bush and Kerry with pictures of White and Black young men who were judged as appearing prototypical of their respective groups. For people who had a negative affective reaction when presented with a Black person, we expected this reaction to be reflected in their judgments of the pictographs. As expected, we found a pattern of in-group favoritism on the task. White subjects showed a strong pattern of misattributions implying more favorable attitudes toward the White photos than the Black photos. In contrast, Black participants showed the opposite pattern. These patterns of misattributions persisted despite blatant warnings against being influenced by the primes. They also correlated with explicit ratings of attitudes toward Blacks and Whites as groups ($r = .58$). However, this relationship was moderated by people's motivations to respond without prejudice. Among those who lacked such motivation, their implicit and explicit measures correlated very strongly. However, among individuals who were highly motivated to avoid prejudice, the relationship was much weaker. This was because highly motivated people expressed highly positive attitudes toward Blacks on the self-report measure, but still showed negativity toward Blacks on the implicit measure.



So far, these results are exactly what one would expect based on previous research comparing implicit and explicit attitude measures, except in two respects. The first is that the correlations are larger than most findings reported. The second is that the misattribution measure produced much higher reliability than most implicit measures, and equal to many explicit measures (average Cronbach's *alpha* = .88). Although the kinds of responses compared in these studies (pleasantness judgments about pictographs and favorability ratings of individuals or groups) are not as discordant as comparing reaction times and rating scales, they were still not directly comparable.

Consider now a modification of the affect misattribution procedure consisting of two phases. One phase is identical to the procedure already described: participants judge pictographs as pleasant or unpleasant and are instructed *to avoid* being influenced by the primes of Black and White photos. In the second phase, participants see the same item pairs, but this time they are instructed to ignore the pictographs and evaluate their reactions to the prime photos themselves as pleasant or unpleasant. The first phase measures attitudes toward the prime items indirectly, through their unintended effects on judgments of the pictographs. The second phase measures attitudes toward the prime items directly. In the first phase, participants are trying not to express any evaluation of the primes. In the second phase, they are trying to express their evaluation of the primes. The stimuli and judgment scale are held constant; the only factor that varies is participants' intentions. With the stimuli and judgment scale held constant, we are in a position to make direct comparisons between the two conditions. We could then use the first phase as an implicit measure of racial attitudes, and the second phase as an explicit measure.

We do not need to stop with this comparison. The second phase of the task is not only an explicit measure of attitudes toward the primes – it is also an *inclusion* condition. The ways people judge the pleasantness may be a product of both intentional and unintentional evaluative influences. The first phase is not only an implicit measure, but it is also an *exclusion* condition. Judgments of the pictographs are only influenced by evaluations of the primes when people's intentions to avoid their influence fail. By comparing these two conditions, we can estimate how much control each person has over whether they express a particular evaluation, and we can estimate what evaluation is revealed when control fails.

As part of validating this procedure, we have to validate some assumptions that are made when carrying out the process dissociation analysis. In the next section, we outline those assumptions that are made by any application of process dissociation, and place them in the context of other models and alternative assumptions that may be made.

### Assumptions and Alternatives

All mathematical models make assumptions in order to relate actual data to formal equations. When applying the process dissociation framework in a new context, it is important to be clear about those assumptions. One assumption of process dissociation is that the controlled and automatic processes at work exert similar influences in inclusion and exclusion conditions. In other words, the two processes should exert as much influence together in the inclusion condition as they exert against each other in the exclusion condition. It is important to avoid the misinterpretation that automatic and controlled estimates should be numerically equal across experimental conditions, or to each other. If the estimates logically had to equal some particular value, we would not need to do the experiment or compute the estimates from data. Instead, it means that the automatic and controlled processes in question play the same roles in inclusion and exclusion conditions.

The second assumption, which has been discussed more widely, is that automatic and controlled processes are independent of each other (for discussions of the independence assumption in memory research, see Curran and Hintzman, 1995; 1997; Jacoby, Begg, & Toth, 1997; Jacoby & Shrout, 1997). Whether this assumption is met depends on the experimental paradigm that is being used. In some cases, automatic and controlled processes could be positively or negatively correlated with each other, which would violate the independence assumption. Because we cannot directly observe the processes, we must indirectly test whether the processes are likely to be independent or dependent. The most common way to do this is to look for dissociations, or selective effects on one or both estimates. The logic is that if automatic and controlled processes are independent, then it should be relatively easy to find variables that affect one but not the other. If the independence assumption is badly violated, then automatic and controlled processes would strongly covary with each other. As a result, it would be difficult to find variables that affect one process without affecting the other.

Much of our work has been focused on examining selective effects on automatic and controlled components in the weapon identification procedure. For example, Payne (2001) found a double dissociation between the two processes. Prime pictures of Black and White faces affected the automatic component but not the controlled component. In contrast, speeded responding affected the controlled component but not the automatic component. Further, racial attitudes correlated selectively with the automatic component. Lambert and colleagues (2003) found that anxiety over an impending public discussion reduced the controlled component without affecting the automatic component. The study described above by Govorun and Payne (2006) showed that ego depletion influenced the controlled component but not the automatic component. Finally, the study by Stewart and Payne (2006) described above showed that specific action plans could affect the automatic component without changing the controlled component.

These dissociations would not be expected if the independence assumption were violated in the weapon identification task. However, for some tasks, or under some conditions, it is always possible to violate one's assumptions. Researchers will be most familiar with these considerations in the context of common statistical tests. It is widely understood that different statistical tests make different assumptions. For example, analysis of variance (ANOVA) assumes a dependent variable that is at least an interval scale, a normally distributed dependent variable, and homogeneity of variance across different conditions, among other things. If an assumption is violated slightly (e.g., a slightly skewed distribution) the resulting biases are usually small. If an assumption is violated badly (for example, distributions are heavily skewed) it is often a good idea to choose a different test that does not depend on the problematic assumption. Just as the failure of an assumption in a particular study does not invalidate the ANOVA technique in general, studies showing that an assumption of process dissociation has been violated do not invalidate the general method. Instead, other methods may be more appropriate in a particular context.



Several other methods are sometimes used as alternatives to process dissociation. These include the task dissociation method (comparing explicit and implicit measures), signal detection theory, and multinomial models. Although they may not be explicitly stated, each of these approaches also involves assumptions which may be violated. Let us take first the task dissociation method. Although this method does not use a mathematical model, it still makes some assumptions. By using an implicit task to measure automatic or unconscious processes, and an explicit task to measure controlled or conscious processes, the task dissociation approach makes the tacit assumption that each measure is process-pure. That is, one assumes that the measures differ only on the dimension of interest to the researcher. If the two tasks differ in ways other than the explicit/implicit dimension, then any different results on explicit versus implicit tasks could be because of those other (confounded) features.

The psychological processes behind implicit tasks (such as reaction times to classify words) and those behind explicit tasks (such as endorsing complex propositional statements) are very different. As a result, the assumption that implicit and explicit tasks differ only on the dimension of interest is not likely to be commonly met. Both process dissociation and task dissociation methods make assumptions to relate observed data to unobserved theoretical ideas. In the case of process dissociation, those assumptions are made explicit, whereas in the task dissociation method they often remain unstated.

A second alternative approach is signal detection theory (SDT). Signal detection theory assumes that perceivers are natural statisticians, who make decisions about world events in the way that researchers decide whether to reject a null hypothesis (Tanner & Swets, 1954). A decision about what one is perceiving or how to respond is treated as a problem of detecting a signal in a noisy environment. Perceivers have a certain amount of evidence, and they select a criterion (similar to the conventional use of $p < .05$ in psychology research) that marks off how strong the evidence has to be before they will accept that a signal is present. Given a pattern of correct responses and errors, signal detection theory can separate *sensitivity* (the ability to discriminate when a signal is actually present or absent) from *bias* (a tendency to respond as if a

signal is present whether it is or not). Signal detection theory is mute on issues of automatic versus intentionally controlled behavior, and its development predated the current interest in automaticity. Nonetheless, signal detection analyses are sometimes preferred to process dissociation on the belief that SDT makes fewer assumptions.

That belief is mistaken. Signal detection theory makes some of the same assumptions as process dissociation, and some that are different. For instance, SDT also makes an independence assumption. It assumes that sensitivity and bias are independent in the same way that process dissociation assumes that controlled and automatic components are independent. Signal detection also assumes normal distributions of evidence strength, and equal variances. Beyond statistical assumptions, signal detection makes substantive assumptions about the way humans process information. For instance, it assumes that decisions are made on the basis of a single continuum of evidence. There is no allowance for qualitatively different kinds of evidence. Process dissociation, in contrast, treats intentional control and automatic biases as qualitatively different processes feeding into behavior. Like ANOVA, both models make assumptions that may be more or less suitable in a given context.

Finally, multinomial models have a great deal in common with process dissociation. A multinomial model posits a branching tree of unobserved cognitive processes, leading eventually to behavioral responses (Riefer & Batchelder, 1988). For example, Klauer and Wegener (1998) developed a model to study the effects of stereotypic expectations on memory. In this model, participants attempt to remember whether they witnessed a given action. If so, they attempt to remember who performed the action. If they cannot remember who performed the action, they may remember the social category to which the actor belonged. And if they cannot remember the social category, they may guess the category on the basis of stereotypes, and so on. Using the pattern of correct and incorrect memory responses, a computer algorithm is used to estimate the best-fitting values for each process. In this way, the degree of memory and guessing at each stage can be estimated, and the model can be tested statistically to see how well it fits the data (see also Conrey et al., 2005).

Multinomial models are sometimes presented as alternatives to a process dissociation approach. However, it is probably more accurate to think of process dissociation and multinomial models as two specific cases of a general family of models. Both make similar assumptions. Both are aimed at separating unobservable psychological processes that give rise to observed behavior. Process dissociation uses algebra to estimate the cognitive processes involved, whereas multinomial models use a computer algorithm. However, the relatedness of the models can be seen in the fact that the process dissociation model can be represented and estimated as a multinomial model with two process parameters (automatic and controlled components; Jacoby, 1998; Payne, Jacoby, & Lambert, 2005).

A multinomial model may have any number of parameters, in any number of combinations. This is both a strength and a weakness. It is a strength because it allows flexibility in exploring various theoretical models. However, the more



parameters a model has, the more it is likely to fit any given set of data, even an incorrect one. As the number of parameters increases, the possible ways to combine those parameters increases exponentially (do people attempt to remember social category only after memory for the person fails, or do they first remember the category and then the person?). There are often many different models that could fit the data equally well. It is important when using this approach to have strong a priori theoretical predictions to avoid choosing an arbitrary model or capitalizing on chance.

Despite these caveats, multinomial models as well as task dissociations, signal detection, and process dissociation all provide valuable tools for taking apart the complex patterns in social behavior into simpler, more digestible parts. Social psychologists know a lot about how knowledge structures and prior experience guide our reactions. Those procedures we know a lot about tend to seem harmless, whereas less familiar ones tend to seem more menacing. As we shift from the comfortable ANOVA and task dissociation methods toward signal detection theory, which is less frequently seen in social psychology, many readers may feel less sure. And as we look closely at multinomial models and process dissociation, there will probably be more unease. The methods seem foreign, the assumptions seem difficult. All of these tools make assumptions, and there is variability in how nervous these assumptions make us. But the difficulty in many cases is not so much with the methods and assumptions as with the newness of these tools for social psychologists. The costs of making assumptions must be gauged against the potential gains in knowledge generated by using these tools. In our view, the evidence reviewed here shows that potential to be high.

## CONCLUSION

We began this chapter by exploring the startling ways that brain damage can fragment mental events that normally flow silently together. At a finer grain of analysis, comparison of implicit and explicit tasks reveals the same kinds of fissures in normal healthy persons. Social psychological research on automaticity (Bargh & Ferguson, 2000), willed behavior (Wegner, 2002), and implicit cognition (Greenwald & Banaji, 1995) is humbling, because it suggests that disorders of autonomy may not be such rare conditions after all. They can be captured in little slips and subtle lapses that are made and forgotten every day. We may have a little more in common with some lesion patients than we thought. But how much? Process dissociation helps tally up the lapses, giving a number to something as gossamer as goals, intent, and will.

## NOTE

1. The values of process estimates here differ by .02–.03 from the values reported in Payne (2001) because a statistical correction was applied to the data in that article to correct for extreme values prior to calculating estimates. Here we have used the raw data for the sake of clarity.

## REFERENCES


Amodio, D. M., Harmon-Jones, E., Devine, P. G., Curtin, J. J., Hartley, S. L., & Covert, A. E. (2004). Neural signs for the detection of unintentional race bias. *Psychological Science*, *15*, 225–232.

Baddeley, A. D. (1986). *Working memory*. London: Oxford University Press.

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: Guilford.

Bargh, J. A. (2005). Bypassing the will: Towards demystifying behavioral priming effects. In R. Hassin, J. Uleman & J. Bargh (Eds.), *The new unconscious*. Oxford, UK: Oxford University Press.

Bargh, J. A., & Ferguson, M. J. (2000). Beyond behaviorism: On the automaticity of higher mental processes. *Psychological Bulletin*, *126*, 925–945.

Beer, J. S., Heerey, E. H., Keltner, D., Scabini, D., & Knight, R. T. (2003). The regulatory function of self-conscious emotion: Insights from patients with orbitofrontal damage. *Journal of Personality and Social Psychology*, *85*, 594–604.

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

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.

Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science*, *12*, 163–170.

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.

Curran, T., & Hintzman, D. L. (1997). Causes and consequences of correlations in process dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 496–504.

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

Doody, R. S., & Jankovic, J. (1992). The alien hand and related signs. *Journal of Neurology, Neurosurgery, and Psychiatry*, *55*, 806–810.

Dutton, D. G., & Aron, A. P. (1974). Some evidence for heightened sexual attraction under conditions of high anxiety. *Journal of Personality and Social Psychology*, *30*, 510–517.

Everling, S., & Fischer, B. (1998). The antisaccade: A review of basic research and clinical studies. *Neuropsychologia*, *36*, 885–899.





Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 75–109). New York: Academic Press.

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

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

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist*, *54*, 493–503.

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

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

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

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. (1998). Invariance in automatic influences of memory: Toward a user's guide for the process-dissociation procedure. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*, *24*, 3–26.

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., & Dallas, M. (1981). On the relationships between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, *110*, 306–340.

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

James, W. (1890). *Principles of psychology*. New York: Holt.

Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, *130*, 169–183.

Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S., & Russin, A. (2000). Just say no (to stereotyping): Effects of training in the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology*, *78*, 871–888.

Klauer, K. C., & Wegener, I. (1998). Unraveling social categorization in the "Who said what?" paradigm. *Journal of Personality and Social Psychology*, *75*, 1155–1178.

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

Lhermitte, F. (1986). Human autonomy and the frontal lobes. Part II: Patient behavior in complex and social situations: The "environmental dependency syndrome." *Annals of Neurology*, *19*, 335–343.

Marchetti, C., & Della Sala, S. (1998). Disentangling the alien hand and anarchic hand. *Cognitive Neuropsychiatry, 3*, 191–207.

Morwitz, V. G., Johnson, E., & Schmittlein, D. (1993). Does measuring intent change behavior? *Journal of Consumer Research, 20*, 46–61.

Muraven, M., & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: Does self-control resemble a muscle? *Psychological Bulletin, 126*, 247–259.

Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology, 64*, 723–739.

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 control modulates the expression of automatic stereotyping. *Journal of Personality and Social Psychology, 89*, 488–503.

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

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*. Oxford, UK: Oxford University Press.

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.

Petty, R. E., & Cacioppo, J. T. (1986). *The Elaboration Likelihood Model of persuasion*. New York: Academic Press.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solo (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 55–85). Hillsdale, NJ: Erlbaum.

Riefer, D. M., & Batchelder, W. H. (1988). Multinomial modelling and the measurement of cognitive processes. *Psychological Review, 95*, 318–339.

Roediger, H. L., & McDermott, K. B. (1993). Implicit memory in normal human subjects. In H. Spinnler & F. Boller (Eds.), *Handbook of neuropsychology* (Vol. 8, pp. 63–131). Amsterdam: Elsevier.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology, 45*, 513–523.

Sherman, S. J. (1980). On the self-erasing nature of errors of prediction. *Journal of Personality and Social Psychology, 39*, 340–350.

Sherman, J. W., Groom, C. J., Ehrenberg, K., & Klauer, K. C. (2003). Bearing false witness under pressure: Implicit and explicit components of stereotype-driven memory distortions. *Social Cognition, 21*, 213–246.

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.

Shimamura, A. P. (1986). Priming effects in amnesia: Evidence for a dissociable memory function. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 38*, 619–644.



Stangor, C., & McMillan, D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social developmental literatures. *Psychological Bulletin, 111*, 42–61.

Stewart, B. D., & Payne, B. K. (2006). Counterstereotypical thought plans reduce automatic stereotyping. Manuscript submitted for publication.

Tanner, W. P., & Swets, J. A. (1954). A decision making theory of visual detection. *Psychological Review, 61*, 401–409.

Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge, MA: MIT Press.

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

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46*, 441–451.