15

Attitudes as Accessibility Bias: Dissociating Automatic and Controlled Processes

B. Keith Payne, Larry L. Jacoby, and Alan J. Lambert

In February 1999, four New York Police Department officers ordered West African immigrant Amadou Diallo to freeze as he stood in his apartment's darkened alcove. Diallo raised his wallet in the air, and 41 bullets later, he was dead. Somewhere between the time Diallo pulled out his wallet and the first officer drew his gun, one of the officers made a judgment that their lives were in danger. The incident caused local protests and national controversy, as the public and the courts debated whether that judgment depended on the fact that Amadou Diallo was Black.

We will not speculate on what actually took place in the Diallo case—a question that is unlikely ever to be answered definitively. But the incident provides a compelling way to frame some broader psychological questions we have been studying. For example, can racial prejudice influence the decisions people make in such high-pressure situations? If so, what is the mechanism by which this influence occurs?

We highlight an important distinction that has been made in social cognition research between automatic or implicit attitudes and more controlled, explicit processes as distinct influences on judgments. We describe an experiment in this tradition in which we used separate direct and indirect tests to measure how explicit and implicit attitudes contribute to a social judgment.

Next, we describe a process dissociation approach (Jacoby, 1991) that treats implicit attitudes as a source of "guessing" or "accessibility bias." Our emphasis on accessibility effects builds on the "New Look" movement in perception (e.g., Bruner, 1957; Greenwald 1992, along with accompanying commentaries). The New Look movement held that perception is strongly influenced by expectancies, values, attitudes, and needs. According to Bruner, perception involves categorization, as do other cognitive activities, and thus reflects differences in the accessibility of categories. As we discuss, subsequent

research questioned whether such accessibility effects reflect an influence on people's ability to discriminate real-world, objective differences or, instead, have their effect through an influence on bias, reflected by people's guesses.

In the Diallo case, the question is whether prejudice resulted in a lessened ability to discriminate between the visual features of a gun versus a wallet or, alternatively, had its effect through an automatic influence on guessing. By the latter alternative, prejudice did not change "true" perception but, rather, because of an inability to distinguish between a gun and a wallet, police relied on their "guess," which may have been, perhaps unconsciously, influenced by Diallo's race. In this chapter, we separately measure the contributions of controlled and automatic processes within a task (e.g., judging whether a gun was present) and treat a measure of accessibility bias as reflecting an automatic, implicit attitude. Note that the contrast we draw between controlled and automatic uses of information is not between explicit versus implicit attitudes. Rather, we identify cognitive control with the ability to respond in a manner consistent with a goal (e.g., task instructions) based on appropriate information in the task at hand (e.g., distinguishing between a gun and a wallet). We end by showing the generality of our approach as a means of analyzing a wide range of accessibility effects of the sort that have been prominent in social psychology.

Automatic and Controlled Processes in Social Cognition

One of the important findings to emerge from the contemporary social cognition literature is that the use of category-based knowledge can be guided by both automatic and controlled processes (e.g., Devine, 1989). Noticing a person's race, for example, has the potential to trigger both a spontaneous stereotype and efforts to control that stereotype (e.g., Dunton & Fazio, 1997). In their efforts to study how automatic and controlled processes guide behavior, researchers have developed several innovative techniques to isolate the two types of processes. One popular approach in recent social cognition research identifies automatic processes with performance on indirect tests, and controlled processes with performance on direct tests.

Implicit attitude studies build on implicit memory studies, in which indirect tests have been used to measure the effects of past experience in the absence of conscious memory for an event (e.g., Jacoby & Dallas, 1981; Roediger & McDermott, 1993). Memory researchers have used tasks such as word fragment and stem completion, lexical decision, and other indirect tests as measures of automatic memory influences. Direct tests, such as recall and recognition, are used to assess explicit memory.

Evidence to support the distinction between implicit (automatic) and controlled (explicit) uses of memory has been gained by showing dissociations

between performances on the two types of test. As an example, amnesiacs show striking memory dissociations. Because of neurological damage, these patients show severe deficits on direct memory tests and may claim no conscious memory for material they have studied. However, on indirect tests such as word stem completions, amnesiac patients perform very similar to the neurologically healthy (see Shimamura, 1986). Similar dissociations between direct and indirect test performance have been shown by people with normally functioning memory (for a review, see Roediger & McDermott, 1993). Finding situations in which direct and indirect test performance can be dissociated provides evidence that processes underlying the two types of test are distinct.

Social psychologists have used indirect methods including word completions (Gilbert & Hixon, 1991), priming tasks (Devine, 1989; Fazio, Jackson, Dunton, & Williams, 1995), and implicit association tasks (Greenwald, McGhee, & Schwartz, 1998) to measure automatic influences of stereotypes and attitudes (see Bargh & Chartrand, 2000). Self-report scales have been used to measure explicit attitudes. Correlational studies have sometimes shown that direct test performance is dissociated from (uncorrelated with) indirect test performance (e.g., Devine, 1989; Fazio et al., 1995). However, other studies have shown that direct and indirect measures of stereotypes or attitudes covary with one another (Lepore & Brown, 1997; Wittenbrink, Judd, & Park, 1997). In the next section, we describe an experiment in which we used separate tasks to measure explicit and implicit attitudes. We show that even when they are uncorrelated with each other, implicit and explicit attitudes can have independent roles in predicting prejudiced responses to stereotyped people.

Identifying Processes With Tasks

The contrasting of direct and indirect measures has led to important theoretical advances in models of stereotyping and prejudice. Social cognition researchers have focused in part on developing indirect tests as measures of individual differences in automatic processing. In this way, automatic and controlled processes are identified with different tasks, which relate to different behaviors. The question addressed by this approach is whether, and under what conditions, each kind of process is active in guiding people's overt responses. A pioneering study by Fazio and colleagues (1995) showed that explicit racial attitudes—as measured by a traditional attitude scale—predicted blatantly race-related judgments such as satisfaction with the verdict in the case of Rodney King. This judgment was considered easily controllable and clearly related to race. Implicit racial attitudes, as measured by an indirect priming task, did not correlate with opinions about the Rodney King

verdict. However, indirect test performance predicted subtler behavior, such as participants' friendliness during an interaction with an African American experimenter (for a similar conceptual and methodological approach, see Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997).

However, it would be a mistake to generalize from these unambiguous examples to all social behaviors. Both automatic and controlled processes simultaneously contribute to most social behaviors, although it may be easier to detect one in some situations than others (see also Wilson, Lindsey, & Schooler, 2000). A study in our laboratory provides a case in point (Lambert, Payne, Ramsey, & Shaffer, in press). In this research, we investigated the joint contribution of implicit and explicit racial attitudes to impressions of a single individual. Participants' explicit racial attitudes were assessed using a number of self-report measures (e.g., the Modern Racism Scale [MRS]; McConahay, Hardee, & Batts, 1981). Implicit attitudes were measured using a lexical decision task as an indirect measure. The direct and indirect measures were used to predict participants' subsequent judgments of the target person.

In the indirect measure, participants decided whether various letter strings were words after being primed (200 ms duration) with either the words *Whites* or *Blacks* or a row of Xs used as a control prime. The words in this study included personality adjectives that varied in whether they were favorable and whether they were related to the Black or White stereotype. The interval between primes and target words (stimulus onset asynchrony, SOA) in this task was 200 ms, well within the range typically used to prevent participants from intentional control of their responses (cf. Fazio et al., 1995; Neely, 1977). As the dependent variable, participants read a short story about a character whose behaviors were ambiguous. They could be interpreted as high or low in hostility and in intelligence. Information participants received was identical except for the race of the target person (Black or White). Although participants' evaluations were subject to strategic control, the race of the target was manipulated subtly, in the context of other demographic information. As a result, there was no clear norm for the appropriate or socially desirable response.

Results showed that, while the indirect (lexical decision performance) and direct (MRS) attitude measures were uncorrelated with one another, each measure independently predicted impressions of the Black target. In other words, participants' overall impressions of the Black target appeared to be driven by two distinct types of attitudes: one measured by the indirect test and the other by the direct test. Neither measure predicted judgments of the White target. These data are provocative in that they suggest a somewhat different view of explicit versus implicit measures compared to that offered by Fazio et al. (1995) and Dovidio et al. (1997). That research suggests that explicit and implicit tasks should be predictive of different classes of behaviors. In contrast, our findings suggest that any given judgment (e.g., the extent to

which a Black person is perceived to be intelligent) can be influenced by both explicit and implicit processes.

However, the contrasting of direct and indirect measures has limitations. When a dissociation between direct and indirect tests is found, questions can arise about whether the lack of a relationship reflects differences in the processes they are intended to measure, or low reliability in the indirect measure (Cunningham, Preacher, & Banaji, 2001; Kawakami & Dovidio, 2001). Second, neither indirect nor direct tests are likely to represent a process-pure measure. On one hand, controlled processes may "contaminate" performance on indirect tests, so that people's strategies can distort the measurement of the memories, attitudes, or stereotypes that researchers are trying to capture. For example, a word completion task may require participants to complete a stem with whatever first comes to mind. If nothing springs to mind, a participant may intentionally search for a word to use from memory. On the other hand, automatic processes may bias responses to self-report measures (Jacoby, 1991). If participants do not know how to respond to a question, they may answer with whatever comes to mind most readily, which can reflect their implicit attitude.

Finally, using different tests to measure separate processes leaves the underlying process confounded with the properties of the test used to measure it. For example, if the results of a self-report measure are sensitive to a manipulation while the results of a word completion task are not, does that dissociation reflect a difference in explicit versus implicit processes, or does it reflect a difference in the sensitivity of the measures?

Attitudes as Accessibility Bias

A complementary approach to teasing apart automatic and controlled processes is to arrange experimental conditions such that the contributions of the two types of processes can be estimated within the same task. Our approach does this by measuring automaticity as a systematic bias in the way people respond. We treat bias or guessing as reflecting attitudes. To illustrate how guessing patterns can be informative about implicit processes, consider the following example of an early indirect attitude test.

Hammond (1948): Guessing as a Measure of Implicit Attitudes

Guessing, by definition, does not fully reflect judgments that are based on objective knowledge about the world. Nevertheless, systematic biases in guessing can be a rich source of information about the ways that people think, as well as the ways they respond. In a classic work, Sherif (1935)

noted that there are cases in which "objective determination is lacking, thus allowing internal factors such as attitudes, subjective norms, and values to play the dominant role in organization of the perceptual field" (p. 60).

An early and clever study by Hammond (1948) showed how a systematic bias in guesses can reveal those internal factors. This article anticipated so closely issues currently being debated in social cognition research that it is worth considering in some detail. Hammond was concerned with measuring unintentional effects of attitudes, even labeling the technique an indirect test. Much like our focus, his method aimed at dissociating the unintentional influence of attitudes from other bases for responding to questions. Finally, Hammond was concerned with the reliability of indirect tests compared to direct tests of attitudes, foreshadowing present efforts to measure and correct for the reliability of indirect measures (Cunningham et al., 2001; Kawakami & Dovidio, 2001).

Fascinated with projective tests such as the Rorschach (1942) inkblot being used to diagnose personality disorders, social psychologists had begun searching for indirect methods for measuring attitudes that could penetrate people's concerns with self-presentation. But being disenchanted with the interpretability of projective tests, Hammond designed a bogus "information test" to reveal unexpressed attitudes. "Much of the difficulty with present methods of attitude measurement," wrote Hammond (1948, p. 38), "lies in the trouble authors have in deciding just what it is they are trying to measure."

In one test, questions about Russia and organized labor were used to measure attitudes toward Communism. In these questions, there was either no factual answer (e.g., "Russia's removal of heavy industry from Austria was (a) legal (b) illegal") or the alternatives were both incorrect (e.g., "Man-days lost because of strikes from January to June, 1946, were (a) 35 million (b) 99 million"). In the latter example, the true answer is 67 million days, midway between the two alternatives provided. The response alternatives provided one favorable choice and one choice unfavorable toward socialist policies (e.g., making union activity appear very costly).

By making no answer factually correct, Hammond attempted to eliminate objective knowledge as a basis for responding, so participants had to guess. Those guesses were not random, but were informed by participants' attitudes toward socialist policies. Hammond compared the responses of American business owners with the responses of union workers. During that historical period, businesspeople were strongly opposed to organized labor and collectivist policies because they feared that such policies would undermine the free market system on which they depended. Union workers were, understandably, expected to be more sympathetic toward collectivist policies. The businesspeople tended to choose antilabor test alternatives to a greater extent than did union members (55% versus 9%). Not only were responses systemat-

ically related to the two groups' political leanings, but the responses showed high consistency, producing reliability coefficients from .78 to .87.

Hammond's article is striking in the extent to which it anticipated many of the theoretical and methodological questions that psychologists are grappling with today. He was interested in the contrast between direct and indirect measures of attitudes and the different processes they revealed. Although Hammond's work is an excellent example of how guesses can reflect implicit processes, we are now in a position to go far beyond the possibilities raised by this early indirect method. We use Hammond's test as an example to illustrate how distinct bases for responding can be measured within the same task.

Separating Knowledge and Attitude

A weakness of the approach introduced by Hammond is that it does not allow for the possibility that some subjects in his experiments knew the correct answer to his questions. As an example, suppose that someone knew how many man-days were lost because of strikes. How would such knowledge influence responding? Differences in knowledge could be measured by adding a condition in which one of the alternatives was the correct one. The importance of separating these two sources of information is evidenced by the different kinds of reactions people have to decisions based on each. When people are concerned with fairness and social justice, responses stemming from prejudice rather than from knowledge trigger outrage. That outrage is based on the assumption that prejudice (attitudes) and knowledge serve as alternative bases for responding. How can the contribution of the two bases for responding be separated?

To do this, suppose we constructed two types of test items. Each test item includes the correct response paired with an incorrect response that is either larger or smaller than the correct response. First, we have an item in which the alternatives are 67 million (correct) versus 35 million (incorrect). From the perspective of the businessperson, the correct alternative can be chosen either because of knowledge or, in the absence of knowledge, because of prejudice (it is the alternative least favorable toward socialism). In the other form of the question, the alternatives might be 67 million (correct) versus 99 million. For the businessperson, the larger, incorrect answer will be chosen because of his or her attitude in the absence of knowledge. In this case, knowledge opposes prejudice.

To illustrate how knowledge and prejudice can be disentangled, we present some idealized data from a fictional test in which businesspeople and union members from 1948 have answered questions of the sort we just described. Anti-Communist items are those on which the correct answer is the anti-

Communist answer. Pro-Communist items are those on which the correct answer is pro-Communist. Table 15.1 shows the hypothetical data.

Look first at the column of pro-Communist items (second column). Notice that the businesspeople made more errors than the union members, showing a tendency to err on the side against Communism, just as Hammond (1948) found. However, these overall performance measures do not allow us to separate knowledge from bias. To do so, consider the processes by which these errors could be made. An anti-Communist error will be made if a person does not know the answer $(1 - K)$, and relies on anti-Communist attitudes (A). The probability of an error in the pro-Communist condition can be written as $p(\text{error} \mid \text{pro-Communist item}) = A(1 - K)$.

Next look at the column of anti-Communist items. Both knowledge and anti-Communist attitudes would lead to a correct answer on these items. For these items, the probability of scoring a correct answer can be written as the probability of using knowledge (K) plus the probability of using attitudes (A) in the absence of knowledge $(1 - K)$. So, $p(\text{correct} \mid \text{anti-Communist item}) = K + A(1 - K)$.

With these two equations, we can estimate knowledge and attitudes by solving for K and A algebraically. Knowledge can be solved as $p(\text{correct} \mid \text{anti-Communist item}) - p(\text{error} \mid \text{pro-Communist item})$. That amounts to this simple "proof" equation: $K = K + A(1 - K) - A(1 - K)$. Consulting our table, we see that for the businesspeople, $K = .90 - .40 = .50$. For the union workers, $K = .60 - .10 = .50$. This shows that the businesspeople and union members know the same amount about the facts in question.

To measure the respondents' attitude-based biases, we begin with the errors in the pro-Communist condition. Recall that the probability of producing an error here is $p(\text{error} \mid \text{pro-Communist item}) = A(1 - K)$. To solve for A, we need to divide the term $A(1 - K)$ by $(1 - K)$. Because we have solved for K already, we know that $(1 - K)$ for both the businesspeople and the union members is $1 - .50 = .50$. Consulting table 15.1, we see that the attitude bias for businesspeople is $.40 \div .50 = .80$. The attitude bias for union members is $.10 \div .50 = .20$. A bias value of .50 represents unbiased responding, whereas scores greater than .50 represent an anti-Communist bias, and scores below

Table 15.1 Hypothetical Responses to an Adapted Version of Hammond's Indirect Attitude Test

	Correct on Anti-Communist Items	Errors on Pro-Communist Items	Estimates	
Social Group	$K + A(1 - K)$	$A(1 - K)$	Knowledge	Attitude
Businesspeople	.90	.40	.50	.80
Union	.60	.10	.50	.20

that point represent pro-Communist bias. The businesspeople have a much stronger bias against Communism than the union members, who actually have a pro-Communist bias.

Use of our equations shows that the businesspeople and union workers differed not in their knowledge but in their biases. Some readers will have recognized the above equation for measuring knowledge as being the same equations commonly used to "correct for guessing" on a multiple-choice test. Interest has typically been in "true" scores, with guessing treated as being largely random and of little interest. In contrast, our approach treats guesses as revealing unintended influences of memory, attitudes, needs, and so on. Our example thus far has shown how guessing bias can reflect attitudes, and how bias can be separated from knowledge. However, there is nothing in what we have done to guarantee that the attitude bias is automatic or that the use of knowledge is controlled. Next, we consider how our approach defines automaticity and control, and how these assumptions are tested.

Defining Automaticity: Processes versus Tasks

Automatic processing has typically been defined as unintentional, unconscious, uncontrollable, and highly efficient, in the sense that an automatic process operates rapidly and does not demand attentional effort (Shiffrin & Schneider, 1977). Several theorists have argued that no task meets all the criteria to be considered automatic in an unqualified way (Bargh, 1989; Logan & Cowan, 1984). We agree that tasks do not meet the criteria of automaticity. However, our interest is in processes, not tasks. We hold that some processes or components of task performance are automatic.

What about the automatic versus strategic qualities in our measure of attitude toward Communism defined as bias? In our approach, we make the qualities traditionally associated with automatic processing independent variables and predict dissociations between the two processes. For instance, by placing intentional and unintentional processes both in opposition to one another and in concert with one another, we demonstrate one process that is sensitive to intentions, and one that is invariant with respect to intentions. By manipulations such as divided attention and speeded responding, we show that controlled processes require attention, and automatic processes operate with little attention.

In the case of Hammond's indirect attitude test, we could test whether the attitude bias was automatic by manipulating factors traditionally associated with automaticity and predicting dissociations. If responses based on conscious knowledge were controlled but prejudice were automatic, then manipulations of fast responding or divided attention would influence knowledge but not bias.

Discriminability and Accessibility Bias in Perception

The above thought experiment serves to introduce the process dissociation procedure (Jacoby, 1991). In this section, we begin by further illustrating that procedure by describing its use to examine effects of expectancy on word identification. Next, we return to the Diallo example and use the process dissociation procedure to analyze how prejudice has its effects.

The words of Sherif (1935) quoted earlier reflect the Gestalt emphasis on internal factors that help construct people's perceptions. Later those ideas were incorporated into the New Look movement (Bruner, 1957), which emphasized the role of expectations in perceptions. Much of the research done to support the New Look movement showed that expectations serve to resolve ambiguity in ways relied upon when projective tests are used. The notion of perceptual defense (McGinnes, 1949) serves as an example. In that research, it was claimed was that the perceptual system protects against noxious stimuli such as obscene words. Ambiguity was created by flashing words for a duration that was so short as to not reliably allow their full identification. Results showed that words that were obscene had to be flashed for a longer duration to be identified than did words that were not obscene, and this disadvantage of obscene words was said to reflect perceptual defense. However, later research sought to explain such "defense" in more mundane ways by appealing to effects of factors such as frequency in the language. It was noted that obscene words occur less frequently in the written language than do words that are similar but not obscene, and the poor perception of obscene words was explained as produced by their low frequency (for a review, see Erdelyi, 1974).

How do expectations that reflect frequency in a person's prior experience have their effect? One possibility is that expectations based on experience serve to truly influence what is perceived so that reality is constructed. Alternatively, such expectations might serve as a source of accessibility bias. For example, suppose that when a four-letter word was flashed, only three of its four letters were identified (e.g., sh_t). The partial information does not allow the word to be identified, so the viewer must guess. The ambiguous stimulus might be interpreted as the word that fits the fragment and is most readily accessible due to prior experience. An experiment by Jacoby (reported by Jacoby, McElree, & Trainham, 1999) examined effects of expectations in identifying ambiguous words.

In Jacoby's experiment, expectations were created by means of training. During training, participants were presented with context cues paired with word fragments that could be plausibly completed in two ways (e.g., knee-b_n_). Participants guessed the completion, and then were given feedback from the computer as to the "correct" completion. In one condition, a typical completion (e.g., bend) was created by presenting that completion on two

thirds of the trials, and an atypical response (e.g., bone) on the remaining one third of the trials. This training phase created expectancies of different strength for how the ambiguous stimulus, the word fragment, would be completed. Frequent and recent pairings between the cue and certain targets made those typical targets highly accessible completions for the fragment.

Following training, participants were again asked to complete fragments. However, this time they were asked to complete the fragments with a masked word that was flashed briefly just before the fragment was shown. The duration of the flash was either 20 or 40 ms. Those trials in which the typical word from training was also the flashed word can be considered congruent trials, because both the typical, accessible response from training and perception of the flashed word would lead to the same correct response. On incongruent trials, the atypical word from training was flashed. On those trials, perception of the flashed word would lead to a correct response, but relying on the accessible response from training would cause an error.

On congruent trials, the correct response might be given either by perceptually identifying the flashed word (P) or by relying on accessibility bias (A), created during training, in the absence of perceptual identification ($1 - P$). Thus, the probability of a hit in the congruent condition is $P + A(1 - P)$. On incongruent trials, accessibility bias is pitted against perception, so that participants are expected to make errors when they rely on accessible information in the absence of perception. If a participant successfully perceives the word completion based on the objective information flashed, then the accessibility of a habitual response is not expected to influence responses. The probability of a false alarm in the incongruent condition is $A(1 - P)$, reflecting that a false alarm is likely to the extent that accessibility bias is active (A) but perception is not ($1 - P$).

Perceptual identification can be estimated by subtracting hits in the congruent condition from false alarms in the incongruent condition: $P = p(\text{Hit} \mid \text{congruent}) - p(FA \mid \text{incongruent})$. Mathematically, this can be seen by the fact that subtracting the term $A(1 - P)$ from the term $P + A(1 - P)$ yields P. Given an estimate of perception, accessibility bias can be estimated as $A = p(FA \mid \text{incongruent})/(1 - P)$. It should be noted that this estimation procedure is the same as illustrated in our thought experiment based on Hammond's (1948) study.

Results from Jacoby's experiment (table 15.2) show that flash duration influenced perception but had no influence on estimated accessibility bias. That is, accessibility bias was nearly identical in the short and long flash-duration conditions. This is important in showing that differences in perception did not influence accessibility bias ("guessing") but, rather, only influenced the opportunity for guessing to drive responses. Again, consider the equation for correct responses in the congruent condition: $P + A(1 - P)$. By that equation, A has an effect only when perception fails, which happens

Table 15.2 Probability of Fragment Completions on Congruent and Incongruent Trials With Estimates of Perception and Habit

Duration (ms)	Correct on Congruent Trials $P + A(1 - P)$	Errors on Incongruent Trials $A(1 - P)$	Estimates	
			Perception	Habit
20	.60	.60	.01	.60
40	.78	.35	.44	.62

Note: Data from Jacoby, McElree, and Trainham (1999).

more often in the short-duration condition. We emphasize this point because we later argue that it is an important one for theories about how prejudice has its effects.

The dissociation displayed in table 15.2 shows an influence on perception with accessibility bias left unchanged, whereas the dissociation in table 15.1 shows an influence on accessibility bias with knowledge left unchanged. We have found dissociations of both sorts in memory and perception experiments. Estimates of accessibility bias can be influenced by manipulating expectancies or typicality in training. Doing so produced differences in estimated accessibility bias that approximated the difference in training probabilities (i.e., probability matching), but left the estimated contribution of controlled processing (recollection in their experiments) unchanged (e.g., Hay & Jacoby, 1996). Note that in the experiment described above, the estimate of accessibility bias is near the training probability (.61 vs. .67). Jacoby and colleagues have found probability matching of this sort in several experiments (e.g., Hay & Jacoby, 1996). Manipulating the amount of time allowed for responding reduced controlled processing but left estimated accessibility bias unchanged, just as would be expected if accessibility bias was a more automatic basis for responding. Results from these and other experiments (e.g., Jacoby, Debner, & Hay, 2001) justify treating accessibility bias as an automatic basis for responding.

One of the assumptions underlying the process dissociation procedure is that the processes of discrimination and bias are independent of one another. This assumption has generated some controversy (see Hintzman & Curran, 1997; Jacoby, Begg, & Toth, 1997, for a discussion of these assumptions). Finding variables such as those reported here that selectively affect the estimated contribution of the two types of processes provides support for the validity of the independence assumption. For example, findings that presentation duration and amount of time required for responding selectively influence perception while prior expectations selectively influence accessibility bias suggest that the two bases for responding operate independently.

Discriminability and Accessibility Bias in Stereotypical Inferences

The experiments described above showed that expectations can bias perceptual judgments and that the influence of expectations can be separated from perception itself. We have been concerned with stereotypes as one source of expectations that may create a similar accessibility bias. Consider the case alluded to earlier, of a police officer in a confrontation with a suspect who is holding an object. The judgment of whether that object is a weapon is urgent, and it must be made quickly. Unfortunately, such a fast decision is the type that is most likely to be influenced by the social category of the suspect (Macrae, Milne, & Bodenhausen, 1994).

Specifically, the fact that African Americans are stereotypically expected to be more violent and criminal than White Americans creates the worry that people might let their stereotypical expectations influence their perceptual identification of weapons. Can race affect people to the point that they claim to see guns where there are none? Under some conditions, it can. Payne (2001) conducted a study with the aim of dissociating automatic bias and controlled responding when racial prejudice influences perceptual judgments. Participants saw pictures of handguns and hand tools on a computer monitor. Their task was to classify each item as either a gun or a tool. Immediately before each target item, faces of White and Black men were flashed briefly, but visibly (200 ms). Some participants were required to respond quickly (in under 500 ms), while others were allowed to respond at their own pace.

Participants under time pressure showed a stereotypical pattern of errors. In addition to making more errors overall, participants mistakenly classified a tool as a gun more often when it was paired with a Black face than when paired with a White face. There was no actual correlation between the color of the face and the identity of the target within the experiment. However, because of people's preexisting stereotypes, a Black face paired randomly with an object was sufficient to cause that object to be misclassified as a gun.

As table 15.3 shows, participants' performance was separated into estimates of accessibility bias and discriminability using the process dissociation equations. Congruent conditions were those in which a Black face preceded

Table 15.3 Process Estimates in Each Prime and Deadline Condition

	Accessibility Bias		Discriminability	
	Black Prime	White Prime	Black Prime	White Prime
Deadline	.57	.49	.40	.44
No deadline	.61	.48	.86	.88

Source: From Payne (2001).

a gun, or a White face preceded a tool. The Black-tool and White-gun pairs provided the incongruent conditions. Accessibility bias was coded so that higher scores indicate a tendency to respond "gun." Process estimates showed that requiring participants to respond quickly cut their discriminability by more than half, but left their bias estimate unchanged. In contrast, the race of the face influenced the accessibility bias estimate, but did not change discriminability. Black faces led to a greater bias toward classifying an object as a gun, compared to the White faces.

We regard the accessibility bias as an automatic influence for several reasons. First, the bias occurred whether or not it was consistent with the goals required for the task. The instructions were to respond "gun" if and only if a gun is present; respond "tool" if and only if a tool is present. The face primes influenced performance regardless of which response was appropriate to the task goal. Second, the bias took place within a very short time period (200 ms SOA), and the response deadline did not affect it. The magnitude of participants' accessibility bias was the same whether or not their processing time was restricted. Together, this evidence suggests that the racial bias caused by the faces of different races may, under some circumstances, be an unintentional and efficient use of information akin to implicit memory.

In addition to showing a double dissociation between accessibility bias and discriminability within the priming task, we tested the relationship between this automatic bias and participants' racial attitudes as measured by a direct test. We measured participants' racial attitudes with the MRS (McConahay et al., 1981) and the Motivation to Control Prejudice Scale (MCP; Dunton & Fazio, 1997). The MCP measures the extent to which people are willing to express any negative attitudes toward other racial groups. We found a positive correlation between the automatic bias estimate and attitudes expressed via the MRS ($\beta = +.51$), but only for participants who were unmotivated to avoid racial prejudice, as measured by the MCP scale. For those who found it inappropriate to express unfavorable attitudes about Blacks, the bias estimate did not correlate with MRS scores ($\beta = -.23$, nonsignificant). These findings converge with other research using different procedures that shows participants' automatic biases correlated positively with their directly expressed attitudes only when they were not motivated to appear unprejudiced (Fazio et al., 1995).

To summarize, in the studies reviewed here, stereotypes and habits created an accessibility bias. Several lines of research suggest that accessibility biases and discriminability often operate simultaneously and independently to jointly determine people's actions. The experiments described above show why process dissociation is useful for breaking down complex effects into separate, quantifiable processes. Those processes can each then be studied individually. When factors traditionally identified with automatic processing are either

built into an experimental design (e.g., using the logic of opposition) or manipulated, it is possible to draw conclusions about underlying processes.

Consciousness and Control

In many of the domains we have described, cognitive control as modeled by discriminability estimates is closely related to conscious subjective experiences. In an explicit test of the relationships between objectively estimated memory processes and subjective experience, Jacoby et al. (2001) had participants describe the phenomenology of each response on a memory test as something they consciously "remembered," just "knew," or "guessed." Those experiments showed a very strong association between estimates of memory discriminability (recollection) and the conscious experience of remembering. Some provocative work has been interested in the possibility of nonconscious control (chapters 7 and 8, this volume; Moskowitz, Gollwitzer, Wasal, & Shaal, 1999). As our approach shows, we see the relationship between consciousness and cognitive control as an empirical question. Having a means for estimating cognitive control independent of participants' reports allows the two to be empirically compared.

Just as the consciousness of control mechanisms may be put to direct test, so can the relationships between consciousness and accessibility biases of different sorts. In some situations, for example, people may be able to become aware of a process that is initially unconscious. As suggested by the psychodynamic tradition, achieving awareness of a mental process can provide a basis for controlling it. An experiment has shown that people can harness their biases strategically if they are made aware of them.

Dolan and Jacoby (2000) conducted a memory study in which the colors of words carried biasing information. During a recognition test, words were tested in either red or green. Two thirds of the old words were tested in green, and one third of the old words were tested in red. Thus, if a word was tested in green, it was more likely to be old than new. Results showed that participants had learned this pattern, in that they made more hits and false alarms for green words than for red words. That is, they were more likely to respond "old" to green words whether they were actually studied or not. However, postexperiment questioning revealed that participants were not aware that there was a relationship between color and the status of the word.

A second experiment avoided relying on participants' self-reports of awareness. In this experiment, participants were either allowed to remain unaware of the color relationship or they were forced to notice it. The unaware condition was the same as described above. In the aware condition, however, participants had to use different response keys, depending on the color of the

word. This arrangement forced participants to attend to the word color. In this experiment, participants were allowed to respond both at their own pace and at a speeded pace in a separate phase of the experiment.

Results showed that neither the color of the words nor awareness of the color relationship affected discriminability. The speeded response condition, in contrast, did reduce discriminability. Expectancies based on word colors affected the bias estimate—participants were biased toward calling green words "old." When participants responded at their own pace, the color bias was the same for aware and unaware conditions. Critically, when they were required to respond rapidly, participants who were aware of the color relationship relied on it heavily, creating a color bias much larger than that of the unaware participants. The point to note is that the bias was in place in all conditions, but awareness of the bias allowed participants to use it strategically. Under difficult speeded conditions, aware participants informed their judgments with the bias, using it to accomplish the task goal. Unaware participants had no such flexible control over their biases. In the next section, we describe additional subtle and even ironic effects that may arise from biases created by accessible attitudes.

A New Look at Some Old Constructs: Habits and Dominant Responses

The construct of habit is one of the oldest in experimental psychology, as it lies at the core of many classic theories of motivation and learning (e.g., Hull, 1951; Spence, 1956). More recently, the importance of habit emerged in the context of the social facilitation literature. Habitual behaviors (or "dominant responses" as they are called in this area) may become more likely when organisms perform a task with, or while being observed by, members of the same species (Zajonc, 1965). The notions of "habit" and "dominant response" have historically been construed in behavioral terms. The well-learned response typically includes some behavior that the organism (be it a rat or a college sophomore) has acquired through repetition, such as successfully negotiating a maze or solving an anagram. However, there is no reason why such constructs could not be usefully extended to domains more familiar to the "cognitive" domain (see James, 1890).

Indeed, a theme running tacitly through many of the experiments reported here is that certain stimuli (or situations) may stimulate well-learned, habitual bases for responding. These habits may be fairly mundane, such as the tendency to automatically think of *bend* when presented with the word *knee*. Other habits carry more immediate consequences, such as the tendency for mere presentation of a Black face to automatically activate images of guns or other stereotypical associations. This raises the more general point that the

construct of attitudes (perhaps the most ubiquitous construct in social psychology) fits quite well within the habit framework, in that attitudes are often defined as a well-learned association between an object and one's positive or negative appraisal of it (Fazio, 1986).

On one hand, framing attitudes as habits or dominant responses does not represent a startling theoretical advance per se. Nevertheless, if this conceptualization is used to make new connections between recent social-cognitive work on attitudes and a long tradition of investigating dominant or habitual responses, then the notion may prove quite useful. In particular, research in the social facilitation literature has shown that the presence of other people facilitates dominant responses, while interfering with subordinate responses (Zajonc, 1965; see also Spence, 1956).

Somewhat surprisingly, our review of this literature shows that researchers have focused almost exclusively on the effects of audiences on performance. Although we are not aware of any previous efforts to make this connection, it seems reasonable that attitudes, like other well-learned responses, may be facilitated by the presence of an audience. If attitudes are construed as dominant evaluative responses, then one would expect the relationship between people's attitudes on one hand and their judgments and behaviors on the other to be strengthened when they are in the presence of others, or are generally aroused.

When applied to racial attitudes, this rationale makes an extremely counterintuitive prediction. Both research and intuition suggest that people are motivated to appear unprejudiced in public settings (e.g., Gaertner & Dovidio, 1986; Plant & Devine, 1998). Thus, it would seem natural that evaluations of stereotyped individuals would be more positive in public than in private settings, and that such evaluations would be more consistent with the perceiver's attitudes in private settings. However, if attitudes are understood as dominant or habitual responses, then one would predict precisely the opposite results: assuming that most people have some level of prejudice, this framework suggests that evaluations of a stereotyped person should be more negative and more closely related to racial attitudes in public settings. In fact, this is exactly what was found in a series of experiments reported by Lambert, Cronen, Chasteen, and Lickel (1996). These researchers found that there was greater "attitude-behavior consistency" in public compared to a private setting, insofar as participants' stereotypical attitudes toward Blacks (as measured by an explicit measure prior to the main study) were more strongly related to judgments of a single Black individual in the former compared to the latter condition.

One ambiguity of these findings is that they can be interpreted in one of two ways. On one hand, the "public expression effect" could reflect heightened accessibility bias. This interpretation is more or less consistent with a Hullian/Zajonc model of social facilitation, which emphasizes the energizing

or drivelike effects of audiences on habitual responses. Another interpretation, however, is that the findings by Lambert et al. (1996) reflect reduced cognitive control, an interpretation that might be predicted by attentional conflict models of social facilitation (Groff, Baron, & Moore, 1983). One way of distinguishing between these accounts is that they presume either that something is being added (increased stereotypical bias) or that something is being taken away (control). It is worth emphasizing that this ambiguity applies not only to the Lambert et al. (1996) investigation but to the social facilitation area as a whole, as theorists in this area have not yet resolved whether drive-based versus control-based accounts provide viable accounts of social facilitation. Nor has a methodology been proposed that can successfully tease these models apart.

We have gained some important leverage in these matters (Lambert et al., 2003). In one study, we followed up on the findings reported by Lambert et al. (1996), showing that the public expression effect was moderated by dispositional levels of anxiety. Highly anxious participants showed strong facilitation of racial attitudes in public compared to private, whereas less anxious participants did not. (Low-anxiety participants showed a nonsignificant reversal of this pattern.) These data are useful insofar as they further demonstrate the important parallelism between our line of work and the social facilitation area, which typically shows these sorts of moderation effects as well.

Even more important, a second study was able to provide a direct test of the two theoretical accounts described above. In this study, participants were randomly assigned to perform the gun versus tool identification task used by Payne (2001) in either a private or anticipated public context. Results revealed significantly more stereotypical errors in public compared to private—especially among participants experiencing high levels of anxiety—conceptually replicating and extending our earlier work in the impression-formation domain. Moreover, use of the process dissociation procedure offered strong leverage in testing the viability of a drive-based account (which would predict heightened bias) versus a control-based account (which would predict lower discriminability). Results showed much stronger support for the latter, as the increase in stereotypical errors was due entirely to lower discriminability in the public condition. These and other current efforts in our laboratory may thus hold potential for clarifying issues in both the stereotyping and the social facilitation areas.

Toward Choosing Among Models: The Tyranny of Automaticity?

Social psychologists have explained a broad range of judgmental biases by the accessibility of mental categories. For example, Bruner (1957) described

perceptual distortions as a result of categorization, which is influenced by the relative accessibility of different categories. Recent use of a trait category can make it more accessible and therefore more likely to be used when forming an impression of a new person (Higgins, Rholes, & Jones, 1977). Even without a recent priming, chronically used attitudes and trait categories can influence the way that social information is processed (Fazio, 1986; Higgins, King, & Mavin, 1982; Wyer & Srull, 1989). Such accessibility effects have typically been revealed by using ambiguous stimuli or by showing errors that result from reliance on accessibility.

Studies of social perception have placed a heavy emphasis on ambiguous situations (e.g., Higgins et al., 1977). The conclusions of research using ambiguous situations are applicable to the extent that everyday judgments are made under ambiguity. No doubt, situations can be ambiguous. However, unambiguous information is frequently available to use as a basis for controlled responding. A problem with using truly ambiguous stimuli, in which there is no "correct" response, is that ambiguous stimuli tell us little about the basis for accessibility effects. In those cases, accessibility effects could arise either from an influence on perception or from accessibility bias. With ambiguous materials, it is impossible to choose between mechanisms. In contrast, we use situations in which there is a "correct" answer and arrange conditions such that bias is either congruent or incongruent with the objective basis for obtaining that correct answer. By doing so, we can separate effects of accessibility on perception and bias.

In some ways, our approach is similar to that of Banaji and Greenwald (1995). As did we, they used a measure of bias to index implicit attitudes. For example, they used signal-detection theory (SDT) to separate effects of discriminability and bias in people's ability to distinguish between famous and nonfamous names. Finding a bias against judging female names as famous was used as a measure of implicit sexism. However, an important difference between their approach and ours is the meaning of the term *bias*. By SDT, a single-process model, bias refers to a quantitative difference in the amount of information needed to make a decision. In contrast, by our dual-process model, accessibility bias reflects a basis for judgments that is qualitatively different from that used for a consciously controlled judgment (e.g., discriminating between a tool and a gun) with regard to the type of information used. Jacoby, McElree, and Trainham (1999) discuss the relation between single- and dual-process models of bias effects (see also Jacoby, Kelley, & McElree, 1999).

Several models of stereotyping take the form of dual-process models (e.g., Brewer, 1988; Devine, 1989; Fiske & Neuberg, 1990). However, research testing these models does not typically measure both processes within the same task. As a result, the operation of each process and the relationships between processes cannot be directly examined. Separate estimates of control

and accessibility bias provide a means to test the relationships between processes, and so provide a basis for choosing between competing models.

This ability represents an important strength, because as Gilbert (1999) noted, dual-process models can take many forms, with separate processes combining in any number of ways to determine behavior. A productive way to proceed is to identify plausible competing theories and test them against one another (Popper, 1994). In the next section, we describe an approach we have taken that exploits the estimates derived from our approach to test which of two plausible dual-process models best accounts for the data we gathered in the weapon identification experiment described earlier.

A Test of Two Models: Accessibility Bias or Attitude Inhibition?

The basic process dissociation model spells out how controlled and automatic processes interact, with automatic processing, accessibility bias, serving as a basis for responses only in the event that control fails. Controlled processes have clear priority, constraining the likelihood that a bias will have the opportunity to affect behavior. This arrangement contrasts with "inhibition" models that are common in current social psychology research. These models assume that when people enter into a social situation, automatic processing constitutes a first stage, which must then be suppressed or inhibited for a person to make a controlled response (Bodenhausen & Macrae, 1998; Devine, 1989; Wegner, 1994).

Lindsay and Jacoby (1994) found that an inhibition model of this sort provides a good description of performance in the Stroop task. In the Stroop task, participants are asked to name the color of the pigment in which words are printed. The critical words are the names of colors themselves. The typical finding is that people have more difficulty naming the color in which a word is printed if the word is the name of some other color. It is believed that in this task, word reading is an automatic process that operates so readily that it must be suppressed if one is to perform well by naming the color of the word rather than the word itself.

Are stereotypes the sort of things that leap out at us with such force that we must struggle to inhibit them if we are to relate objectively to another person? Is our hypothetical officer likely to "read" an object in the hand of a Black person as a gun as spontaneously as one reads words for meaning? Or do stereotypes operate more subtly, so that they have an opportunity to shape judgments primarily when more controlled processes fail?

We have used a multinomial processing tree (MPT) approach to empirically test which kind of model best fits the data generated in our weapon identification paradigm (see Batchelder & Riefer, 1999, for a discussion of

MPT modeling). In an MPT model, mental processes are represented as a series of branches. Figure 15.1 shows the tree diagrams for an accessibility bias (top panel) and an inhibition model (bottom panel). As with a traditional flowchart, a process can operate or not operate at each step. What distinguishes this approach from a traditional flowchart is that the likelihood that each process is active is represented by a probability. On the right hand side of the figure are the responses expected as a result of each processing path. The number of correct and incorrect responses in each condition from an experiment can be used to test whether each model fits the pattern of actual data.

Look first at the accessibility bias model. This model represents the same equations used in the process dissociation procedure described throughout. In this model, controlled discriminability of the target either succeeds with probability C, or it fails with probability $(1 - C)$. If control fails, then an accessibility bias may occur with probability A, or fail to occur with probability

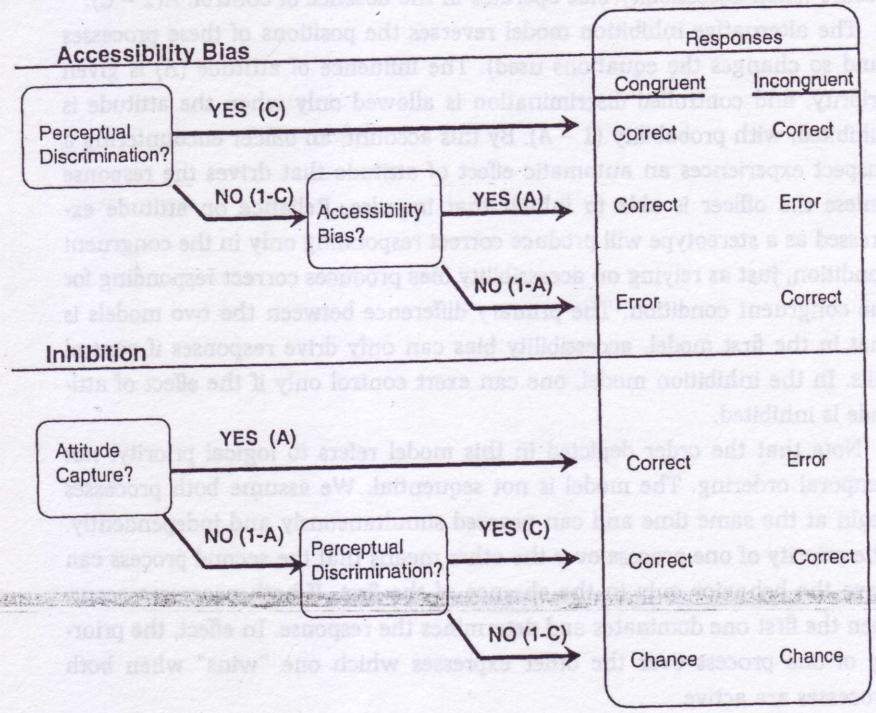

Figure 15.1 A comparison of two models of the relationships between controlled and automatic processes. Note that the order of processes depicted does not refer to temporal stages, but to the relative dominance of each process. In the event that both processes occur, the process in the first position determines the response.

$(1 - A)$. Rather than operating only when able to inhibit an automatic response, controlled processing begins very early. However, it may take longer to complete than automatic processing. Psychologically, this sequence would describe a process in which an officer encountering a suspect attempts to discern whether the object is a gun or not, based on the physical characteristics of the object. If controlled processing is successful, then racial stereotypes do not influence the weapon identification.

In contrast, when control is impaired (e.g., by rushed responding due to the intense pressure of the situation), the automatic processing of the suspect's race may cause a stereotypical judgment to be made. This is a case of an accessibility bias left unopposed by controlled processing. To see the relationship between the equations used in the process dissociation procedure and the probabilities represented in the diagram, notice the paths denoted by the arrows. As shown in figure 15.1, a respondent can give a correct answer in the congruent condition either by control of the response (C) or by an accessibility bias in the absence of control: $A(1 - C)$. The fact that accessibility bias can only drive responses when control fails amounts to the equation used in process dissociation, where an error in the incongruent condition occurs when accessibility bias operates in the absence of control: $A(1 - C)$.

The alternative inhibition model reverses the positions of these processes (and so changes the equations used). The influence of attitude (A) is given priority, and controlled discrimination is allowed only when the attitude is inhibited, with probability $(1 - A)$. By this account, an officer encountering a suspect experiences an automatic effect of attitude that drives the response unless the officer is able to inhibit that impulse. Reliance on attitude expressed as a stereotype will produce correct responding only in the congruent condition, just as relying on accessibility bias produces correct responding for the congruent condition. The primary difference between the two models is that in the first model, accessibility bias can only drive responses if control fails. In the inhibition model, one can exert control only if the effect of attitude is inhibited.

Note that the order depicted in this model refers to logical priority, not temporal ordering. The model is not sequential. We assume both processes begin at the same time and can proceed simultaneously and independently. The priority of one process over the other means that the second process can drive the behavior only in the absence of the first. If both processes occur, then the first one dominates and determines the response. In effect, the priority of one process over the other expresses which one "wins" when both processes are active.

Once the branches of the multinomial tree model are filled out using equations based on the probabilities just described, the model can be statistically tested against experimental data. Rather than solving for the estimates algebraically, the control and accessibility parameters are estimated using a

solver algorithm. The data from Payne (2001) were used to test the models. The pattern of errors and correct responses predicted by the different models was compared to experimental data, and a statistical goodness-of-fit test was used. In testing the fit of multinomial models, a G^2 statistic is used, which is similar to a chi-square statistic. For this particular model, with a significance level of .01, the critical value for G^2 is 13.28. Values below that limit are considered an acceptable fit, and values above that limit indicate that the model is rejected.

The accessibility bias model fit the data well ($G^2 = 3.64$), but the inhibition model did not ($G^2 = 27.23$). The estimates yielded by the accessibility bias model converged nicely with the process dissociation estimates calculated originally on this data set. The controlled parameter was held constant across both prime races, and was different for short ($C = .42$) and long ($C = .87$) response deadlines. The accessibility bias parameter was held constant across deadlines, and varied between the Black ($A = .59$) and White ($A = .53$) prime conditions. Our results suggest that participants were not "blinded" by race so that they could only discriminate between weapons and tools when they inhibited the race bias. Instead, the actual objects and racial bias served as separate bases for responding, with decisions based on perceptual discrimination requiring cognitive control. When that control failed, the automatic race bias had its effect.

Taken in combination, these findings emerging from our laboratory challenge contemporary models of stereotyping in several ways. First, most models assume a two-stage mechanism in which the automatic activation of the stereotype can be subsequently overridden by controlled processes, but only if perceivers have the motivation and ability to do so (e.g., Devine, 1989; see also Bodenhausen & Macrae, 1998; Gilbert & Hixon, 1991). Second, the fact that stereotypes have greater impact on social judgment under cognitive load—or other manipulations that compromise careful processing—is typically interpreted as a consequence of "knocking out" perceivers' ability to implement the secondary correction process. The research on weapon misidentifications discussed above suggests that these rarely tested assumptions may not always be correct. Particularly important is the idea that constraints on controlled processing affect discriminability but not accessibility bias. Thus, the notion that prejudice plays a greater role in social perception under situational duress—such as might have occurred in the Diallo case—may reflect compromised ability to respond based on objective properties rather than a failure to successfully override the stereotype.

Our argument is not that inhibition models are never applicable. As alluded to earlier, Lindsay and Jacoby (1994) found that a similar model was a good description of the way that interference occurs in the Stroop task. However, different models are useful to describe different psychological effects. It is important to have a means of choosing between different models of the

processes underlying different kinds of behaviors. The choice of models used to explain the effects of prejudice has important consequences for attempts to control prejudice.

An inhibition model suggests that if objective information is to guide judgments, automatic impulses must be inhibited. If they are not, then the more controlled basis for responding does not matter. Inhibition-related assumptions lead to an emphasis on suppressing prejudiced reactions. Research has focused on attempts to suppress stereotypical thoughts from initially entering consciousness (Wegner, 1994), to replace stereotypical thoughts with egalitarian ones (Monteith, 1993), and to adjust judgments once stereotypical thoughts have already come to mind (Wilson & Brekke, 1994).

Attempts to reduce prejudice by these routes have left some theorists pessimistic. It is argued that in order to correct stereotypical thinking by "thinking twice," people must be aware of their bias as well as being motivated and able to override it (Bargh, 1999; Wilson & Brekke, 1994). They may not be aware of the bias because it has operated automatically, and thus it may be unconscious. This pessimism is supported by studies showing that attempts to suppress stereotypes sometimes fail, causing an ironic increase in stereotype use (Macrae, Bodenhausen, Milne, & Jetten, 1994; Monteith, Spicer, & Toomen, 1998; Payne, Lambert, & Jacoby, 2002). In addition, the demands of constantly second-guessing oneself may be overwhelming.

An alternative approach suggested by the accessibility bias model emphasizes maximizing cognitive control rather than overriding the effect of accessible information. By this approach, it is not necessary to be aware of any bias, and there is no need to suppress automatic thoughts so long as objective grounds provide a basis for responding. According to this model, accessibility bias only influences responses when controlled processes fail. As an example, consider the word-perception experiments described earlier. The habitual response only influenced behavior when perception failed. If the situation had been structured to maximize perception, then even the strongest accessibility bias would have no impact on responses.

In fact, this is precisely what happened in Payne's (2001) weapon identification study. When participants were allowed unlimited time to respond, their accessibility bias was just as strong as in the speeded condition. However, this accessibility bias only translated into significantly more stereotypical errors in the speeded condition. When participants were allowed cognitive control by responding slowly, the accessibility bias was unable to cause actual errors. Situations structured so that people can use objective information as a basis for control reduce the consequences of accessible information. And they do so without (often futile) attempts to suppress automatic thoughts. The choice between theoretical models can inform choices about interventions. Thus, we believe the approach described here is as relevant to pragmatics as it is to processes.

References

Banaji, M. R., & Greenwald, A. G. (1995). Implicit gender stereotyping in judgments of fame. *Journal of Personality and Social Psychology, 68*, 181–198.

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

Bargh, J. A. (1999). The cognitive monster: The case against the controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 361–382). New York: Guilford.

Bargh, J. A., & Chartrand, T. L. (2000). The mind in the middle: A practical guide to priming and automaticity research. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 253–285). New York: Cambridge University Press.

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

Bodenhausen, G. V., & Macrae, C. N. (1998). Stereotype activation and inhibition. In R. S. Wyer, Jr. (Ed.), *Advances in social cognition* (Vol. 11, pp. 1–52). Mahwah, NJ: Erlbaum.

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

Bruner, J. S. (1957). On perceptual readiness. *Psychological Review, 64*, 123–152.

Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science, 121*, 163–170.

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

Dolan, P. O., & Jacoby, L. L. (2000). *Response biases, implicit learning, and the effects of age.* Unpublished manuscript, New York University.

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.

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

Erdelyi, M. H. (1974). A new look at the new look: Perceptual defense and vigilance. *Psychological Review, 81*, 1–25.

Fazio, R. H. (1986). How do attitudes guide behavior? In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (pp. 204–243). New York: Guilford.

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

Fiske, S. T., & Neuberg, S. L. (1990). A continuum model of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology, 23*, 1–74.

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

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

Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology, 60*, 509–517.

Greenwald, A. G. (1992). New Look 3: Unconscious cognition reclaimed. *American Psychologist, 47*, 766–779.

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.

Groff, B. D., Baron, R. S., & Moore, D. L. (1983). Distraction, attentional conflict, and drivelike behavior. *Journal of Experimental Social Psychology, 19*, 359–380.

Hammond, K. R. (1948). Measuring attitudes by error choice: An indirect method. *Journal of Abnormal and Social Psychology, 43*, 38–48.

Hay, J. F., & Jacoby, L. L. (1999). Separating habit and recollection in young and elderly adults: Effects of elaborative processing and distinctiveness. *Psychology and Aging, 14*, 122–134.

Higgins, E. T., King, G. A., & Mavin, G. H. (1982). Individual construct accessibility and subjective impressions and memory. *Journal of Personality and Social Psychology, 43*, 35–47.

Higgins, E. T., Rholes, W. S., & Jones, C. R. (1977). Category accessibility and impression formation. *Journal of Experimental Social Psychology, 13*, 141–154.

Hintzman, D. L., & Curran, T. (1997). More than one way to violate independence: Reply to Jacoby and Shrout (1997). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 511–513.

Hull, C. L. (1951). *Essentials of behavior*. New Haven, CT: Yale University Press.

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., 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., Debner, J. A., & Hay, J. F. (2001). Proactive interference, accessibility bias, and process dissociations: Valid subjective reports of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 686–700.

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

Jacoby, L. L., McElree, B., & Trainham, T. N. (1999). Automatic influences as accessibility bias in memory and Stroop tasks: Toward a formal model. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII* (pp. 461–486). Cambridge, MA: Bradford, MIT Press.

James, W. (1890). *Principles of psychology* (Vol. 1). New York: Henry Holt.

Kawakami, K., & Dovidio, J. F. (2001). The reliability of implicit stereotyping. *Personality and Social Psychology Bulletin, 27*, 212–225.

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

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

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

Lepore, L., & Brown, R. (1997). Category and stereotype activation: Is prejudice inevitable? *Journal of Personality and Social Psychology, 72*, 275–287.

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

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review, 91*, 295–327.

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Jetten, J. (1994). Out of mind but back in sight: Stereotypes on the rebound. *Journal of Personality and Social Psychology, 67*, 808–817.

Macrae, C. N., Milne, A. B., & Bodenhausen, G. V. (1994). Stereotypes as energy-saving devices: A peek inside the cognitive toolbox. *Journal of Personality and Social Psychology, 66*, 37–47.

McConahay, J. B., Hardee, B. B., & Batts, V. (1981). Has racism declined in America? It depends on who is asking and what is asked. *Journal of Conflict Resolution, 25*, 563–579.

McGinnes, E. (1949). Emotionality and perceptual defense. *Psychological Review, 56*, 244–251.

Monteith, M. J. (1993). Self-regulation of prejudiced responses: Implications for progress in prejudice reduction efforts. *Journal of Personality and Social Psychology, 65*, 469–485.

Monteith, M. J., Spicer, C. V., & Toomen, G. D. (1998). Consequences of stereotype suppression: Stereotypes on AND not on the rebound. *Journal of Experimental Social Psychology, 34*, 355–377.

Moskowitz, G. B., Gollwitzer, P. M., Wasel, W., & Schaal, B. (1999). Preconscious control over stereotype activation through chronic egalitarian goals. *Journal of Personality and Social Psychology, 77*, 167–184.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited capacity attention. *Journal of Experimental Psychology: General, 106*, 226–254.

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., 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.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology, 75*, 811–832.

Popper, K. R. (1994). *The myth of the framework*. London: Routledge.

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.

Rorschach, H. (1942). *Psychodiagnostics: A diagnostic test based on perceptions.* Bern, Switzerland: Hans Huber.

Sherif, M. (1935). A study of some social factors in perception. *Archives of Psychology, 187,* 60.

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.

Spence, K. W. (1956). *Behavior theory and conditioning.* New Haven, CT: Yale University Press.

Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review, 101,* 34–52.

Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Review, 116,* 117–142.

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

Wittenbrink, B., Judd, C., & 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., Jr., & Srull, T. K. (1989). *Memory and cognition in its social context.* Hillsdale, NJ: Erlbaum.

Zajonc, R. B. (1965). Social facilitation. *Science, 149,* 269–274.