*Journal for the Theory of Social Behaviour* 38:3
0021–8308

# Varieties of Social Cognition

ERIC LUIS UHLMANN, DAVID A. PIZARRO AND PAUL BLOOM

After reading words related to stereotypes of the elderly, such as "Florida" and "wrinkle," people tend to walk more slowly (Bargh, Chen, & Burrows, 1996). Being subliminally exposed to pictures of African American males makes people hostile, and thinking about professors improves their performance at Trivial Pursuit (Bargh et al., 1996; Dijksterhuis & van Knippenberg, 1998). People are more likely to mistakenly judge a male to be famous, and an African American to be a criminal (Banaji & Bhaskar, 2000; Banaji & Greenwald, 1995; Payne, 2001). Based on these and other similarly dramatic findings, we and many other psychologists have come to agree with Bargh and Chartrand (1999), who proposed that: "... most of a person's everyday life is determined not by their conscious intentions and deliberate choices but by mental processes that are put into motion by features of the environment and that operate outside of conscious awareness and guidance" (p. 462; see also Dijksterhuis & Bargh, 2001).

The present article addresses the criteria used to determine the consciousness and unconsciousness of mental processes, specifically those central to social cognition. We use the term unconscious to mean "currently inaccessible to conscious introspection." This is similar to the definition of unconsciousness offered by Baars (1997, p. 187), specifically that which is "largely inaccessible at any given time." Our definition is also close to Armstrong's (1980) notion of introspective unconsciousness, which he views as the most important form of unconsciousness and the one closest to commonsense notions of what it means to be unconscious. Operationally, this means that the person is unable to accurately report the relevant mental phenomenon when asked.

There are, however, two importantly different types of unconscious social cognition: (i) unconsciousness of the *influences* on judgment and behavior and (ii) unconsciousness of the *mental states* (i.e., attitudes and feelings) that give rise to such judgments and behaviors. An unconscious influence occurs when an individual is unaware that a stimulus in the environment has led to changes in her feelings, attitudes, judgments or actions. For instance, she may be unaware that the order in which she viewed a series of nightgowns influenced which one she decided to buy (Nisbett & Wilson, 1977), of the role played by location in her choice of

colleges (Wilson & Stone, 1985), that an interaction partner's subtle chin-rubbing has led her to rub her own chin (unconscious mimicry; Chartrand & Bargh, 1999), of the factors that influenced her to like one painting more than another (Wilson & LaFleur, 1995), or that exposure to words related to recklessness has led her to perceive a man as reckless (Higgins, Rholes, & Jones, 1977).

This is to be distinguished from an unconscious *state*. This exists when the person is unaware of the feeling or attitude that gave rise to her judgment or action. For instance, an individual may be unaware that she is suddenly feeling especially committed to the goal of behaving cooperatively or that she has negative automatic attitudes towards Black people. As these examples illustrate, the state in question could be either temporary (e.g., a fleeting commitment to cooperativeness temporarily primed by the environment) or stable (e.g., a person's longstanding attitude towards members of a social group). Notably, state unconsciousness implies influence unconsciousness because it is impossible to be aware of influences on a state if you are not aware of the state itself. As such, state unconsciousness can be considered a deeper or more profound form of unconsciousness.

The distinction between unconscious influences and unconscious states is an intuitive one, and relevant to multiple areas of psychology and related fields. Consider the effects of violent media on aggressive behavior, which are of interest not only to social psychologists (Anderson & Bushman, 2002), but also developmental psychologists (Huesmann, Eron, Lefkowitz, & Walder, 1973), clinical psychologists (Singer & Singer, 1983), and sociologists (Phillips, 1979, 1983). Unconsciousness of the effects of violent media may take different forms. An individual may be unaware that he is behaving aggressively due to watching a violent television program (influence unconsciousness) or he can be unaware that he is feeling aggressive at all (unconsciousness of the mediating state).

This distinction between unconscious influences and unconscious states is not intended to be an exhaustive categorization of unconscious cognitive phenomena. There also exists unconscious perception (Holender, 1986; Reingold & Merikle, 1988, 1990), unconscious learning (Reber, 1989), unconscious memory (Jacoby, 1991; Roediger, 1990; Schacter, 1992), and even unconscious behavior (Armstrong, 1980; Kihlstrom, Mulvaney, Tobias, & Tobis, 2000). However, unconscious influences and states have deservedly received special attention from researchers due to their profound implications for social judgment and behavior, among these unconscious aspects of consumer choices (Brunel, Tietje, & Greenwald, 2004), managerial behavior (Chugh, 2004; Kay, Wheeler, Bargh, & Ross, 2004), anxiety and shyness (Asendorpf, Banse, & Mücke, 2002; Egloff & Schmukle, 2002), cooperation (Bargh, Gollwitzer, Lee-Chai, Barndollar, & Troetschel, 2001), self-esteem (Greenwald & Banaji, 1995), health-related behaviors (Czopp, Monteith, Zimmerman, & Lynam, 2004; Sherman, Rose, Koch, Presson, & Chassin, 2003), social affiliation (Chartrand & Bargh, 1999), career choices (Rudman & Heppen, 2003), sexual

harassment (Bargh, Raymond, Pryor, & Strack, 1995), intellectual performance (Dijksterhuis & van Knippenberg, 1998), moral judgments (Haidt, 2001), phobias (Teachman & Woody, 2003), and suicide (Phillips, 1979). At the same time, there has been relatively little formal discussion of the criteria used to conclude unconsciousness of influences and states. The present article is an effort to address this significant gap in the literature.

We will argue here that while there is conclusive evidence that people are often unconscious of the influences on their judgments and behaviors, unconscious mental states are much more difficult to verify empirically. One reason for this is that in contemporary research on unconsciousness of mental states, such as the unobtrusive priming of behavior and the implicit measurement of attitudes, unconsciousness is typically the null hypothesis (i.e., that evidence of awareness will *not* emerge). Because it is difficult to confirm the null hypothesis, it is also difficult to conclusively demonstrate state unconsciousness. In an effort to further research on this topic, we propose criteria for assessing the unconsciousness of mental states, and apply them to recent empirical findings.

It is important to emphasize that our goal is to identify *conservative* criteria for verifying unconscious influences and states. All of the evidence of unconscious social cognition reviewed here—including experimental designs that equate unconsciousness with the null—provides meaningful evidence that people frequently lack introspective access into their mental processes. In many if not most cases, it is perfectly reasonable to argue that the preponderance of the evidence points to unconscious influences and states. Notably, researchers often refer to influences and states that are strong candidates for unconscious social cognition as "automatic," "nonconscious," and "implicit," perhaps in part to avoid making overly strong claims about consciousness vs. unconsciousness (e.g., Bargh et al., 1996, 2001; Greenwald & Banaji, 1995). However, more conservative criteria are required in order to conclusively demonstrate unconscious social cognition.

## EVIDENCE FOR INFLUENCE UNCONSCIOUSNESS

The evidence that people are often unconscious of the influences on their judgments and behaviors comes from five primary sources: debriefings of experimental participants, subliminal priming studies, comparisons of the judgments of actors and observers, the effects of asking individuals to analyze the reasons for their attitudes, and studies that manipulate participants' awareness of potential influences on their judgments. In part because they equate unconsciousness with the null hypothesis, debriefings and subliminal priming studies provide significant but not conclusive evidence of unconscious influences. In contrast, actor-observer comparisons, reasons analyses, and experimental manipulations of awareness provide conclusive evidence.

## 1. Participant Debriefings

A classic review by Nisbett and Wilson (1977) questioned the common assumption that we have introspective access into the influences on our judgments. For example, they noted that participants' preferences for different nightgowns are demonstrably influenced by the order in which the nightgowns are presented, that their opinions on school busing change in response to persuasive messages, that their evaluations of a teacher's attributes (his appearance, accent, and mannerisms) are affected by the warmth or coldness of his personality, and that subtle hints from the experimenter improve their performance on logic problems. Yet when subsequently asked to report the factors that influenced their judgment, participants are unable to accurately assess the influence of these experimenter-induced factors. Rather, they believe that they chose the most objectively appealing product, that their opinions were unchanged from the week before, that their evaluations of the teacher were caused by their evaluations of his attributes (when the reverse was actually the case), and that they solved the logic problem by virtue of their own inspired intellect. Nisbett and Wilson (1977) concluded that we lack genuine introspective access into our mental processes, and instead use shared cultural theories to construct *post-hoc* explanations for our judgments—and mistake them for conscious awareness of the causes of our behavior.

While they provide important evidence of unconscious influences, participant debriefings have some shortcomings. First, they can make it difficult to distinguish between unconsciousness and forgetfulness. A person who was only momentarily aware that she was being influenced by a stimulus may appear, as far as the experimenter conducting the debriefing is concerned, to have never had any such awareness.

A second issue is that debriefings rely on a null effect to demonstrate a lack of consciousness awareness. Taking the null hypothesis as evidence for unconscious processes is less than ideal because null effects can occur for any number of reasons, many of them methodological. In the case of debriefings, there is always the possibility that better questions, a different demeanor on the part of the experimenter, or some other change in the experimental situation would produce evidence of conscious awareness. Participant debriefings therefore provide significant but not conclusive evidence of unconscious influences.

## 2. Subliminal Priming Studies

Also supportive of influence unconsciousness are studies in which subliminally presented stimuli influence judgments and behaviors (e.g., Bargh et al., 1996; Bargh & Pietromonaco, 1982). For example, Bargh and Pietromonaco (1982) had participants complete an ostensive vigilance task. During the task, words related to hostility were flashed for a fraction of a second. The primes were further masked using a string of Xs that appeared immediately afterward. Participants so

primed subsequently perceived an ambiguously described target person as hostile. An additional condition demonstrated that participants were unable to guess what the masked prime words were. Such studies provide important evidence for influence unconsciousnes. After all, if participants never consciously perceived the primes, they could not possibly have been aware of their influence.

However, whether the procedures used in such studies actually rule out conscious perception of the primes is somewhat controversial (see Draine & Greenwald, 1998; Holender, 1986, and the commentaries following both articles). The debate largely boils down to whether subjective or objective thresholds should be used to assess conscious perception (Reingold & Merikle, 1988, 1990). A subjective threshold relies on participants' self-reports of the conscious perceptibility of the prime (as in Bargh & Pietromonaco, 1982). In contrast, an objective threshold uses participants' performance on another task, rather than their self-reports, as the measure of awareness. For example, to determine whether participants could consciously perceive the valence of their subliminal primes, Draine and Greenwald (1998) administered a second task requiring them to consciously categorize the briefly flashed words as positive or negative by pressing one of two computer keys. Performance at chance levels was taken as evidence that there was no conscious perception of the primes. If one accepts a subjective threshold for determining unconscious perception, studies like Bargh and Pietromonaco (1982) prove unconsciousness of influence. But if one favors a conservative objective threshold, they do not (Draine and Greenwald, 1998).

Proponents of objective thresholds (e.g., Draine & Greenwald, 1998) argue that subjective thresholds are too lax since participants could be somewhat aware of the primes, but not enough to provide perfectly accurate reports. They further note that researchers sometimes ask participants the wrong question, such as reporting the specific word rather than inquiring about the critical feature of the prime (such as whether it was a positive word or a negative word). Proponents of a subjective threshold (ourselves included) argue that above-chance performance on objective tasks need not reflect conscious awareness. Quick motor responses on an awareness task as to whether a briefly flashed word is good or bad might be correct due to unconscious processing of its meaning. That is, when participants are obliged to respond quickly, their key presses might be influenced by unconscious perceptions of the stimuli rather than conscious choices. Due to this unresolved controversy over whether objective (i.e., performance-based) or subjective (i.e., self-report based) thresholds are more appropriate, a conclusive case for influence unconsciousness cannot be made solely based on subliminal priming studies using a subjective threshold.

## 3. Comparisons of Actors and Observers

A particularly convincing source of evidence for influence unconsciousness is the comparison of actors' assessments of the influences on their judgments with the assessments of observers (Nisbett & Bellows, 1977; Nisbett & Wilson, 1977;

Wilson & Brekke, 1994). In these studies, one set of participants (actors) formulates a judgment (for example, of the deceptiveness of interviewees; Kraut & Lewis, 1982) and assesses the role of various potential influences on their judgment (e.g., the role of eye contact in their judgments of deceptiveness). Other participants (observers) are merely given a description of the judgmental situation and asked to predict either how a specific actor (Wright & Rip, 1981), the average actor (Wilson, Laser, & Stone, 1982), or they themselves (Kraut & Lewis, 1982) would formulate their judgments. Within-subject correlations are calculated between actual and perceived influences for both actors and observers, and their relative accuracy is compared. Remarkably, actors are not more accurate at estimating the influences on their own judgments than are observers (see Wilson & Stone, 1985, for a review). This suggests that rather than introspecting, they rely on cultural theories of influence that they share with observers (Nisbett & Wilson, 1977; Wilson & Stone, 1985).

Actor-observer comparisons technically rely on null effects, in the form of no difference between the accuracy of actors and observers. However, since both actors and observers show significant correspondence between estimated and actual influences (e.g., highly significant within-subject correlations; Wilson & Stone, 1985), such null findings are attributable to the accuracy of observers rather than measurement problems.

## 4. The Effects of Asking Individuals to Analyze the Reasons for Their Attitudes

Additional evidence for influence unconsciousness that does not rely on equating unconsciousness with the null can be found in studies that ask people to reflect on the reasons for their evaluations of attitude objects. The attitude objects can range from beverages and vacation pictures to academic courses, political candidates and even significant others. Individuals who are asked to analyze the reasons for their attitudes tend to change their original attitudes, are less satisfied with the choices they make based on those attitudes, show weaker correspondence between their attitudes and their behaviors, and are less able to predict their future behaviors—all of which suggests that they are unaware of the true causes of their attitudes (Wilson & LaFleur, 1995; Wilson et al., 1993; Wilson, Dunn, Bybee, Hyman, & Rotondo, 1984; Wilson, Dunn, Kraft, & Lisle, 1989). In other words, if attitudes are significantly changed when individuals are asked to consciously report the reasons for their attitudes, then these original attitudes were most likely not the product of those conscious reasons.

## 5. Experimental Manipulations of Conscious Awareness

Perhaps the strongest evidence for unconscious influences on judgment is that participants who are made more aware of such influences correct for them. These

effects have been demonstrated in research on the effects of primed trait words on the impressions formed of target persons. For example, presenting words related to the trait "reckless" in the context of an unrelated task (e.g., unscrambling sentences) leads participants to assimilate to the primes in the impressions they form of others (e.g., individuals evaluate the target person's actions as more reckless; Higgins et al., 1977; see also Bargh & Pietromonaco, 1982). However, priming effects of this sort are eliminated and even reversed when participants are made more aware of the potentially biasing influence of the primes. For example, individuals who cannot remember the words with which they had been previously primed fall prey to the typical assimilation effect. Conversely, individuals who can consciously recall the primes demonstrate contrast effects—rendering judgments *opposite* to those the primes would ordinarily lead them to make (for example, judging a target person to be *less* stubborn after being primed with words related to stubbornness; Lombardi, Higgins, & Bargh, 1987; see also Newman & Uleman, 1990).

Similar effects are found when awareness is experimentally increased. For example, individuals subtly reminded of the priming episode just prior to reading about the target person (for example, by asking them whether the priming task was easy or difficult) also demonstrate a contrast effect, correcting away from the influence of the prime. Participants who are not given such a reminder demonstrate assimilation (Erb, Bioy, & Hilton, 2002; Strack, Schwarz, Bless, Kübler, & Wänke, 1993). Blatantly priming participants with trait-related words prior to the impression formation task (as opposed to more subtle manipulations, such as embedding the words in a supposedly unrelated task), likewise produces contrast effects (Newman & Uleman, 1990; see also Moskowitz & Roman, 1992).

Moreover, such contrast effects appear to depend on participants' capacity for deliberative processing. Moskowitz and Skurnick (1999) primed participants with words related to the trait "hostile" using word puzzles. Again, being reminded of the primes led to a contrast effect in judgments of the target person, such that they were actually deemed less hostile after hostility had been primed. However, this contrast effect was eliminated in a condition in which participants were placed under cognitive load via the distracting task of keeping the titles of academic articles in memory (see Martin, Seta, & Crelia, 1990, for a similar finding).

This indicates that assimilation to primes is unconscious by showing that increased awareness of being primed makes the assimilation effect disappear (and is replaced by a contrast effect). That even minimal increases in participants' degree of conscious awareness of being primed decreases the influence of primes on social judgments strongly argues that that influence is unconscious—even if the degree of awareness participants require in order to make their corrections is not perfect.

Again, one of the strengths of this methodology is that it does not require accepting the null to conclude a lack of conscious awareness. Rather, it utilizes the presence of an effect (a contrast effect with increased awareness) to show that

another effect (assimilation to unobtrusively presented prime words) is unconscious. Another notable strength of this methodology is that, because it does not depend on participants' recollections of having been influenced, it does not risk confusing forgetfulness with unconsciousness. One is therefore on safe ground making the strong claim that unobtrusively primed concepts unconsciously influence judgments and behaviors.

## Summary

People are clearly often unconscious of the influences on their judgments. Participant debriefings and subliminal priming studies, which tend to equate unconsciousness with the null hypothesis, together provide significant, but not conclusive, evidence for unconscious influences. Conclusive evidence of influence unconsciousness comes from studies that compare the judgments of actors and observers, ask participants to analyze the reasons for their attitudes, and manipulate conscious awareness of primed concepts. It appears a fundamental aspect of social cognition that we lack introspective access into many influences on our feelings, judgments, and behaviors.

## STATE UNCONSCIOUSNESS I: ARE PEOPLE AWARE OF THE MENTAL STATES THAT MEDIATE INFLUENCES ON THEIR BEHAVIOR?

There are two different versions of the proposal that people are unconscious of the mental states that give rise to their judgments and behaviors. We will discuss here work suggesting that environments prime behavior without any awareness of intervening mental states, then turn to research suggesting that implicit measures reveal attitudes that are unconscious.

The possibility that the environment can prime behavior with no awareness of intervening mental states is raised by studies in which unobtrusively primed concepts led to dramatic changes in social judgments and behaviors (Bargh & Ferguson, 2000; Bargh & Chartrand, 1999; Bargh, Raymond, Pryor, & Strack, 1995; Chen, Lee-Chai, & Bargh, 2001; Lakin & Chartrand, 2003; Spencer, Fein, Wolfe, Fong, & Dunn, 1998; Van Baaren, Maddux, Chartrand, De Bouter, & Van Knippenberg, 2003). For example, Bargh et al. (1996) found that priming words related to politeness (using an ostensibly unrelated sentence-unscrambling task) led participants to wait longer before interrupting the experimenter. Taken together, debriefings and the effects of blatant versus subtle priming conclusively demonstrate that participants are not aware of the influence of the primes. But at the same time, testing to see whether a person is unconscious of the influence of the primes is different from testing to see whether they are unconscious of the internal states that mediate the effects of the priming. For instance, a person might be unaware

that exposure to words related to competitiveness has caused her to behave competitively, but still be aware that she is feeling particularly competitive at the moment.

## Unobtrusive Priming can Influence Conscious States

Indeed, recent studies indicate that unobtrusive priming can influence conscious mental states (e.g., Baldwin, Carrell, & Lopez, 1990; DeMarree, Wheeler, & Petty, 2005; Kawakami, Dovidio, & Dijksterhuis, 2003; Kay, Wheeler, Bargh, & Ross, 2004; Kay & Ross, 2003; Levesque & Pelletier, 2003; McCoy and Major, in press; Mussweiler, Rüter, & Epstude, 2004; Nelson & Norton, 2005; Shah & Kruglanski, 2002, 2003; Shah, 2003a/b; Stapel & Blanton, 2004; Tamir, Robinson, Clore, Martin, & Whitaker, 2004; Winkielman, Berridge, & Wilbarger, 2005). In relevant investigations, participants primed with words related to autonomy reported greater intrinsic motivation (Levesque & Pelletier, 2003), subliminal exposure to happy and sad faces influenced the consciously perceived value of a fruit flavored drink as well as the amount of the drink participants intended to consume (Winkielman et al., 2005), being subliminally primed with the scowling face of department head Bob Zajonc led graduate students to explicitly evaluate their own research ideas more negatively (Baldwin et al., 1990), Catholic women who read a sexually explicit passage and were subliminally primed with the face of the Pope consciously perceived themselves as worse Catholics (Baldwin et al., 1990), priming the concept "superhero" increased college students' explicit willingness to help others in hypothetical situations (Nelson & Norton, 2005), individuals engaged in a competitive game consciously perceived their performance more positively when subliminally primed with smiling faces (Tamir et al., 2004), primes related to business environments led participants to perceive competitiveness as more situationally appropriate (Kay et al., 2004), participants primed with words related to cooperativeness were more likely to report the intention to cooperate in a Prisoner's Dilemma situation (Kay & Ross, 2003), and unobtrusively priming words related to merit increased the extent to which participants consciously endorsed meritocracy (McCoy & Major, in press).

Other work has examined the effects of being primed with members of stereotyped groups on personal attitudes. In three experiments, Kawakami et al. (2003) found that subliminally or otherwise unobtrusively priming college students with elderly persons led them to report more positive attitudes towards publicly funded health care and more negative views of nudity on television. In a fourth experiment, students primed with skinheads reported more negative attitudes towards immigrants and racial minorities.

At the same time, however, there are also studies in which primed concepts did not influence a particular self-reported state (e.g., Aarts, Gollwitzer, & Hassin, 2004; Bargh et al., 2001; Chartrand & Bargh, 1996; Fitzsimons & Bargh, 2003;

Shah & Kruglanski, 2002, 2003; Winkielman et al., 2005). For instance, priming words related to cooperativeness did not influence self-reported commitment to the goal of behaving cooperatively (Bargh et al., 2001), priming words related to memorizing had no effect on self-reported memorization goals (Chartrand & Bargh, 1996), subliminally priming people with the name of their best friend had no effect on their motivation to be understanding about another person's behavior (Fitzsimons & Bargh, 2003), subliminally priming a goal relevant to an alternative task had no effect on the perceived difficulty of a current task (although it did influence self-reports of thinking about and feeling distracted by the alternative goal; Shah & Kruglanski, 2002, 2003), and subliminally presenting happy and sad faces did not influence participants' conscious mood (although effects were found on conscious desire to consume a beverage as well as the consciously perceived value of the beverage; Winkielman et al., 2005). However, that primes have significantly influenced conscious states in a sizeable number of investigations at least raises the possibility that conscious states can mediate priming effects on behavior.

## Changes in Conscious States can Mediate Priming Effects

Additional work finds evidence that changes in conscious states can mediate prime-to-behavior effects (Chartrand, van Baaren, & Bargh, 2006; Hertel & Kerr, 2001; Shah & Kruglanski, 2002, 2003; Shah, 2003a/b; Strahan, Spencer, & Zanna, 2002). For example, Shah (2003b) found that subliminally priming participants with the names of significant people in their life influenced the expected difficulty and perceived importance of an anagram task, and that these conscious states statistically mediated (Baron & Kenny, 1986) the effects of the primes on behavior. Chartrand et al. (2006) found that consciously reported mood significantly mediated the effects of subliminal affect primes on stereotypic judgments. In other relevant work, priming an alternative goal increased conscious thoughts about the alternative goal, negatively impacted performance at a current task, and self-reports of feeling distracted by the alternative goal mediated the effect of the primes on performance (Shah & Kruglanski, 2002, 2003).

Also, Strahan et al. (2002) found that subliminally priming the word "thirst" increased the influence of an advertisement on both self-reported attitudes and taking coupons for the advertised sports drink. Moreover, it seems likely that changes in conscious attitudes towards the sports drinks at least partly mediated the effects of the priming manipulation on behavior, given that the self-report attitude measure and behavioral measure of taking coupons were highly correlated ($r = .71$). Hertel and Kerr (2001) report that priming the concept "loyalty" using an ostensible verbal memory task led to the conscious perception that loyalty was expected of group members, greater self-reported ingroup identification, and behavior favoring the ingroup (allocating more resources to ingroup members

than to outgroup members). In addition, the consciously perceived normativeness of group loyalty mediated the effects of the primes on behavior—controlling for the effects of the priming manipulation on perceived norms reduced the behavioral effect to nonsignificance.

Again, such studies are simply existence proofs that conscious states *can* mediate prime-to-behavior effects, and cannot demonstrate that such effects are *usually* mediated by conscious states. As noted earlier, other studies find little to no evidence that particular self-reported states mediate priming effects (e.g., Aarts et al., 2004; Bargh et al., 2001; Chartrand & Bargh, 1996; Fitzsimons & Bargh, 2003). However, the results of studies like Shah (2003b), Chartrand et al. (2006), and Hertel and Kerr (2001) do suggest that at least some priming effects are mediated by states of which the person is conscious.

## Was the Correct State Assessed?

In our view, the most critical issue with demonstrating that unconscious states mediate priming effects is that it is difficult to figure out what the right state to test conscious awareness of is. For example, Shah (2003b) observed effects of priming significant others on conscious expectations of success on an anagram task, but no effects on a general sense of self-efficacy. Similarly, Hertel and Kerr (2001) measured both self-reported group identification and perceived norms, but only the latter mediated the effects of the priming manipulation.

In fact, relatively little is known about what sorts of mental states typically mediate the effects of priming manipulations (Cesario, Plaks, & Higgins, 2006; DeMarree et al., 2005; Wheeler & Petty, 2001). Wheeler and Petty (2001) provide an extensive list of potential mediators, among them changes in mood, approach-avoidance states, goals, behavioral tags, motivations, perceived norms, and self-concept. DeMarree et al. (2005) argue that changes in the self-concept mediate priming effects, and present some supporting evidence (see also Wheeler, DeMarree, & Petty, 2005). Bargh et al. (2001) and Cesario et al. (2006) propose that stereotype priming effects might be mediated by goals associated with stereotyped groups (see Cesario et al., 2006, for some empirical evidence). Other potential mediating states readily come to mind, for instance expectations, beliefs, abstract attitudes, and attitudes towards the behavior (Ajzen & Fishbein, 1977).

Judging the relative accuracy of each of these theories concerning potential mediators is critical to determining whether prime-to-behavior effects are mediated by conscious or unconscious states. Consider for a moment the possibility that DeMarree et al. (2005) are right, and changes in self-concept mediate many priming effects. If so, it is difficult to resolve the issue of whether people are state conscious or not because very few studies assessed conscious self-concept. And if Bargh et al. (2001) and Cesario et al. (2006) are correct that stereotype primes activate goals, it is hard to tell whether such primes are mediated by a conscious

state because measures of self-reported goals were typically not assessed in stereotype priming studies. This is not to argue that DeMarree et al. (2005), Bargh et al. (2001), or Cesario et al. (2006) are necessarily correct, only that we need to know what sorts of state mediates a priming manipulation before we are in a position to test for conscious awareness of that state.

## Criteria for the Verification of an Unconscious State

We propose that two steps are necessary to show that the effects of a priming manipulation are mediated by changes in an unconscious state. First, we must have at least a general idea of what sort of state (e.g., self-concept, mood) mediates the particular priming manipulation in question. Otherwise, the possibility remains open that some other conscious state mediated the effects of the priming (e.g., it was perceived norms rather than self-concept, or attitudes rather than mood).

For the most part, this criterion has not been met. One important exception is work showing that some priming manipulations are mediated by changes in goal states (Bargh et al., 2001; Cesario et al., 2006). Goals are known to increase in strength until acted on (Atkinson & Birch, 1970), and priming words like "succeed" produces effects on behavior that are even stronger after a brief delay (Bargh et al., 2001). People persist in pursuing their goals even in the face of obstacles (Gollwitzer & Moskowitz, 1996), and individuals primed to succeed continue to work on a task even when instructed to stop (Bargh et al., 2001). Such results strongly support the hypothesis that words related to success prime a goal to succeed on achievement-related tasks.

Second, evidence is needed that while self-reports of the relevant state do not mediate the effects of unobtrusively presented primes, the same self-report measure does mediate a more explicit manipulation. To give a hypothetical example, priming people for success leads them to solve more anagrams with no noticeable change in conscious states, but a conscious goal to be successful does mediate the effects of explicitly instructing them to try to succeed. Since in this hypothetical example the self-report measure performed well under theoretically expected circumstances, it was clearly both valid and relevant for that particular situation. Because participants were randomly assigned to either the unobtrusive priming condition or the explicit instruction condition, it would further be clear that the consciousness with which the goal was adopted caused self-reported goals to mediate more effectively in one condition than in the other.

This second criterion was first proposed by Bargh et al. (2001), who carried out a relevant empirical investigation. Participants took part in a commons game that gave them the opportunity to either share resources with others or use up communal resources for their own personal gain. After the game, participants completed self-report measures of their commitment to behaving cooperatively during the game. Both participants with a conscious and unobtrusively primed

goal to cooperate behaved more cooperatively, soaking up less communal resources in return for personal profit. While individuals with a conscious goal to be cooperative evidenced a correlation between their self-reported commitment to cooperativeness and their behavior in the commons game, individuals primed with cooperation did not. This provides conclusive evidence that the effects of the primes were mediated by an unconscious goal.

### Summary

The available data indicates that priming effects are sometimes mediated by conscious states, while null effects suggest state unconsciousness in other cases. Verifying unconscious states is challenging due to persistent uncertainty regarding what sorts of states mediate prime-to-behavior effects. Two criteria should be met before concluding that a priming manipulation was mediated by an unconscious state. First, independent evidence is needed regarding what sort of state (e.g., self-concept, attitude, mood, goals, perceived norms) mediates the priming effect. Second, a self-report measure of that state should be unresponsive to the priming manipulation, yet mediate the effects of a more explicit manipulation.

## STATE UNCONSCIOUSNESS II: ARE PEOPLE AWARE OF THEIR SOCIAL ATTITUDES?

One of the most thought provoking and frequently discussed hypotheses regarding social cognition is that we are often unaware of the attitudes that give rise to our judgments and behaviors. While some leading researchers have hypothesized that social attitudes are often unconscious (Banaji, 2001; Greenwald & Banaji, 1995), others have argued that they are generally conscious (Fazio & Olson, 2003; Gawronski & Bodenhausen, 2006; Gawronski, Hofmann, & Wilbur, 2006; Gawronski, LeBel, & Peters, 2007).

Gawronski, Bodenhausen, and their colleagues (Gawronski & Bodenhausen, 2006; Gawronski et al., 2006, 2007) propose a useful distinction between the automatic activation of evaluative associations in memory (e.g., automatic evaluations) and propositional processes concerned with whether the association is valid (e.g., explicit evaluations). For instance, seeing an African American person activates concepts related to violence, after which the social perceiver consciously decides whether or not the association is accurate. Gawronski (2007) further argue that while evaluative associations are automatically activated, they are nonetheless available to conscious introspection.

Several considerations make it difficult to conclusively demonstrate that an attitude is unconscious. It is important to again distinguish between being unconscious of the factors that influence a mental state versus being unconscious of the state itself (see also Fazio & Olson, 2003; Gawronski et al., 2006, 2007). For

example, mere exposure studies demonstrate that subliminal exposure to a stimulus increases subsequent liking of the stimulus (Zajonc, 1980). However, the resulting attitude is an unconscious effect of the subliminal exposure, not an unconscious attitude—indeed, since the dependent variable in mere exposure research is self-reported liking, the attitude is by definition consciously accessible. Similarly, one can be consciously aware of one's attitude, yet be completely unaware that it was created by an unobtrusive classical conditioning procedure (e.g., Olson & Fazio, 2001).

It is also important to distinguish between unconsciousness of an attitude versus unconsciousness that the attitude is being assessed (see also Fazio & Olson, 2003; Gawronski et al., 2007). Reaction time measures like evaluative priming (Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Fazio, Jackson, Dunton, & Williams, 1995) and the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) use response latencies when representatives of a social group (e.g., "Republicans") are paired with evaluative words (e.g., "Bad") to assess social attitudes. But neither the IAT nor evaluative priming demonstrates that the association being assessed is unconscious in nature—only that the participant cannot consciously control her responses to the stimuli.

## Correlations between Different Measures of Attitude

Considerable research compares responses on implicit and explicit measures of attitude. In this work, attitudes are measured either directly using explicit self-report questionnaires, or indirectly (i.e., implicitly) through reaction time. In many cases, researchers find that implicit and explicit measures of attitude are either weakly correlated or uncorrelated with one another (Dovidio et al., 1997; Fazio et al., 1995; Gaertner & Dovidio, 1986; Greenwald et al., 1998).

However, dissociations between implicit and explicit measures cannot be used to conclude independent unconscious and conscious systems (Roediger, 1990; Schacter, 1992). The reason is that dissociations are not only common between implicit and explicit measures, but also between different implicit measures (for examples from the attitude literature, see Boniecki & Jacks, 2002; Bosson, Swann, & Pennebaker, 2000; Fazio & Olson, 2003; Nosek & Hansen, 2006; Rudman, Ashmore, & Gary, 2001).

These dissociations often occur due to meaningful differences between implicit measures. For example, priming measures assess the average evaluative response to a series of *individual* African American faces, whereas IAT measures assess evaluations of the *category* "African Americans" (De Houwer, 2001; Olson & Fazio, 2003). But dissociations between priming and IAT measures are observed even when both assess attitudes towards categories (Boniecki & Jacks, 2002; Bosson et al., 2000; Nosek & Hansen, 2006; Rudman et al., 2001). Thus, low or null correlations between different implicit measures probably also occur for methodological reasons. Indeed, as research using multi-method, multi-trait matrices has shown,

measures often cluster together based on method variance rather than the construct they are ostensibly tapping (Campbell & Fiske, 1959; Cook & Selltiz, 1964; Cote & Buckley, 1987). This raises the possibility that dissociations between different implicit measures, as well as between implicit and explicit measures, result to a considerable extent from random and systematic measurement error.

With regard to random error, implicit measures are less statistically reliable than explicit measures (Bosson et al., 2000; Cunningham, Preacher, & Banaji, 2001; Gawronski et al., 2007; Nosek & Hansen, 2006; Olson & Fazio, 2003). For example, a recent meta-analysis indicates that the average correlation between priming and questionnaire measures of racial attitudes is only .16 (average $r = .24$ collapsing across all implicit measures; Dovidio, Kawakami, & Beach, 2001). But consider that the reliability of standard priming measures averages about .30 across studies (see Bosson et al., 2000; Cunningham et al., 2001; Kawakami & Dovidio, 2001; Nosek & Hansen, 2006; Olson & Fazio, 2003). Such high degrees of measurement error constitute a leveling bias that may mask the true relationship between implicit and explicit measures.

Consistent with this idea, use of latent variable analysis to correct for measurement error increases the correlation between evaluative priming and IAT measures and self-reported racial attitudes from below .20 to .45 (the correlation between the IAT and evaluative priming measures also rises dramatically; Cunningham et al., 2001). Similarly, correcting for measurement error raises the correlations between IAT and questionnaire measures of attitudes towards African Americans vs. European Americans, the rich vs. the poor, gay vs. straight men, Americans vs. foreigners, and Christians vs. Jews to .47 (Cunningham, Nezlek, & Banaji, 2004). Notably, the average implicit-explicit correspondence for racial attitudes revealed by meta-analysis ($r = .24$; Dovidio et al., 2001) is close to the uncorrected correlations found by Cunningham et al. (2001, 2004).

Recent work by Payne and his colleagues illustrates the role of systematic measurement error in the correspondence between implicit and explicit measures. Their work employed the Affect Misattribution Procedure (AMP; Payne, Cheng, Govorun, & Stewart, 2005), a task in which Black and White faces are briefly flashed just before participants evaluate ambiguous pictographs. White Americans tend to evaluate the ambiguous pictographs more negatively when a Black face has just been flashed, a pattern of responses reflective of automatic prejudice. The correlations between AMP scores and self-reported racial attitudes are systematically higher when the implicit and explicit tasks are equated on irrelevant methodological variables (Payne, Burkley, & Stokes, 2008). When carefully matched, correlations upwards of .60 are observed, even without correcting for random error.

Social desirability pressures further attenuate the relationship between implicit and explicit measures of attitude. Many attitude domains judged by independent raters to be low in social desirability pressures (e.g., political, academic, and consumer attitudes) are marked by correlations between IAT and explicit measures that approach or exceed .50 even without correcting for measurement error

(Brunel, Tietje, & Greenwald, 2004; Karpinski, Steinman, & Hilton, 2005; Maison, Greenwald, & Bruin, 2001; Nosek, 2005; Nosek & Banaji, 2002; Nosek, Banaji, & Greenwald, 2002; Swanson, Rudman, & Greenwald, 2001; although see Hofman, Gawronski, Gschwendner, Le, & Schmitt, 2005). (Implicit and self-report measures do exhibit low correlations in many non-sensitive domains, such as attitudes towards flowers vs. insects; Greenwald et al., 1998; Karpinski & Hilton, 2001). Assessed at the level of individual differences, a general concern about providing socially desirable responses similarly attenuates the relationship between priming and explicit measures of prejudice against African Americans (Dunton & Fazio, 1997; Fazio et al., 1995; Payne, 2001; see also Payne et al., 2005) and IAT and self-report measures of attitudes towards gay men (Banse, Seise, & Zerbes, 2001).

The moderating role of social desirability pressures is somewhat difficult to interpret because the attitudes people are motivated to hide from others they may even be loathe to admit to themselves. In other words, implicit and explicit measures of prejudice may correlate less highly than similar measures of political attitudes because the implicit measure is tapping into an attitude which the individual consciously denies to herself. There is, however, some evidence that the more specific motivation to hide one's prejudices from *other people* reduces implicit-explicit correspondence. Whereas individuals who report motivation to respond without prejudice specifically to avoid social censure show no relationship between IAT and self-report measures of attitudes towards gay men, individuals who report feeling little pressure to lie show substantial implicit-explicit correspondence (Lemm, 2000). In an experimental study, when participants were asked to be completely honest in their explicit responses, the correlation between a self-report and IAT measure of self-esteem increased significantly (Olson, Fazio, & Hermann, in press). These findings indicate that 1) motivations to provide socially desirable responses on questionnaires attenuate the relationship between implicit and explicit measures and 2) this at least partly results from an effort to deceive others about one's attitude.

## Awareness of Automatic Attitudes

In some cases, an individual likely does have two distinct attitudes, one tapped by explicit measures, the other by an implicit measure such as the IAT (see Kihlstrom, 2004; Wilson, 2002; Wilson, Lindsey, & Schooler, 2000). While we have noted the difficulty of concluding a dissociation based on low correlations, scores on implicit and explicit measures can also show mean differences, such as when White Americans show a more consistent preference for Whites over Blacks on the IAT than on self-report measures (Greenwald et al., 1998). Such *mean dissociations* (Nosek, 2002, 2005) can be so dramatic that they are difficult to attribute entirely to social desirability concerns. However, even when a person does hold dual attitudes (Wilson, 2002; Wilson et al., 2000), it does not necessarily

follow that her second attitude is unconscious. She may be aware of her second attitude but not report it because she does not perceive it as reflecting her true self. For example, she may be consciously aware of her unwanted automatic prejudices towards racial minorities, but express her more personally endorsed, egalitarian views when asked for her attitude.

In fact, there is some evidence that White Americans are aware of their automatic prejudices. While White students tend to explicitly reject prejudiced ideas and beliefs about minorities, about 75% agree to questionnaire items like: "Although I don't necessarily want to, I sometimes have prejudiced reactions (like spontaneous thoughts and gut feelings) towards racial minorities that I'm not sure I can prevent" (Uhlmann & Nosek, 2005). In addition, respondents correctly estimate that their first, gut response to an African American person will be more prejudiced than their later responses (Voils & Monteith, 2004). Moreover, individuals who report discrepancies between how they should and would act towards African Americans (e.g., I *should* vs. *would* think stereotypical thoughts) demonstrate more prejudiced behaviors, but only under cognitive load—suggesting that people can also be aware of their automatic prejudiced behaviors (Monteith & Voils, 1998).

Additional suggestive evidence for awareness of automatic attitudes comes from work showing that implicit and explicit measures interact to predict judgments and behaviors. These interactions suggest that people not only compensate, but in some cases even *overcompensate* for their automatic attitudes. For example, individuals who are automatically prejudiced but who are consciously motivated to respond without prejudice respond even more favorably towards Black targets in terms of their trait judgments (Olson & Fazio, 2004) and willingness to interact with the person (Towles-Schwen & Fazio, 2003) than individuals who are not automatically prejudiced (see also Dasgupta, 2004). As noted earlier, increased awareness of an automatic process can lead to correction effects (Newman & Uleman, 1990; Moskowitz & Roman, 1992).

In addition—and consistent with theories positing that narcissism and defensiveness stem from effortful attempts to compensate for negative automatic views of the self—individuals high in self-reported self-esteem but low in automatic self-esteem are more narcissistic and defensive (for example, rationalizing their behaviors more; Jordan, Spencer, Zanna, Hoshino-Browne, & Correll, 2003; McGregor & Marigold, 2003). Because in order to make such corrections, people likely possess at least some fleeting awareness of the attitude (Wilson & Brekke, 1994), this research further supports the hypothesis that people can be aware of automatic attitudes that conflict with their intentionally endorsed views.[1]

## Alternative Criteria for Concluding Unconsciousness vs. Consciousness of Attitudes

What alternative criteria might be used to verify that an implicit measure is tapping into an unconscious attitude? We suggest that a good criterion would

draw upon multiple correlated implicit measures. As noted earlier, implicit measures generally do not correlate with each other, making it difficult to interpret low implicit-explicit correlations. Multiple correlated implicit measures are necessary to reduce the possibility that low implicit-explicit correlations are due to methodological variables like random and systematic measurement error. When different implicit measures show convergent validity (e.g., Cunningham et al., 2001; Olson & Fazio, 2003), we can have more confidence that their failure to correspond with explicit measures reflects dissociated attitudes.

Also, self-perceived automatic attitudes should be measured in addition to intentionally endorsed attitudes. Some evidence suggests that people can be aware of their unwanted automatic attitudes. It is therefore necessary to compare responses on implicit measures to people's beliefs about their automatic attitudes, not just evaluations that they intentionally endorse.

We suggest that there are two types of dissociations between implicit and explicit measures that can be used to verify an unconscious state. First, automatic and self-reported attitudes may be *uncorrelated at an individual differences level*. In other words, those individuals who report positive evaluations of the attitude object are not the same individuals who have positive automatic associations with the attitude object. When multiple correlated implicit measures fail to correspond with self-report measures of the same construct, both intentionally endorsed and self-perceived automatic attitudes are assessed, and a correction for measurement error is made, null implicit-explicit correlations become much more diagnostic of unconscious attitudes.

Second, participants may exhibit *self-reported and automatic evaluations that are opposed in valence*. In other words, people prefer attitude object A over attitude object B on self-report measures, yet exhibit more positive automatic associations with attitude object B than with attitude object A. Even in cases in which implicit and explicit measures of attitude are highly correlated at an individual differences level, very different mean preferences suggest a conscious-unconscious dissociation.

Conversely, if implicit measures are tapping into an attitude of which the person is conscious, implicit and self-report measures should correlate highly at an individual differences level and further reveal about the same mean preference. The self-report measure could be of either intentionally endorsed attitudes or self-perceived automatic attitudes. Below we illustrate how these criteria apply to recent empirical findings.

## Applying the Criteria

On average, people exhibit a preference for flowers over insects on both self-report measures of attitude and on the Implicit Association Test (Greenwald et al., 1998; Karpinski & Hilton, 2001). But at the same time, self-reported attitudes and automatic associations are virtually uncorrelated at an individual differences

level. This appears to be a case of implicit-explicit correspondence in terms of mean preferences, along with an implicit-explicit dissociation at an individual differences level. This pattern of results suggests that people are not aware of their automatic associations. If they were aware, the people who self-reported the strongest preferences for flowers over insects should have exhibited the most positive associations with flowers relative to insects. Note, however, that this conclusion would be considerably stronger if more than one implicit measure had been used and self-perceived automatic attitudes assessed.

Also, while self-report and IAT measures of attitudes towards evolutionary theory vs. creationism are highly correlated at an individual-differences level ($r = .60$), people actually exhibit *opposite* preferences on self-report and IAT measures (Nosek, 2002, 2005). Specifically, while participants on average self-report a preference for evolutionary theory, on the IAT they exhibit a preference for creationism. While this pattern of results would again be more conclusive if more implicit measures were used and self-perceived automatic attitudes assessed, it does suggest a lack of awareness of the automatic attitude. Even when implicit and explicit measures are highly correlated across individuals, if people exhibit opposite preferences on implicit vs. explicit measures there are probably unconscious attitudes at work.

We now turn to a few findings that suggest awareness of the evaluative associations assessed by implicit measures. The first notable study compared participant's self-perceived automatic attitudes towards gay people with their scores on implicit measures of automatic associations (Ranganath, Smith, & Nosek, in press; see also Smith & Nosek, 2005). Participants reported both their "gut feelings" and deliberatively endorsed feelings towards gay people. Participants' self-reported gut feelings (i.e., self-perceived automatic attitudes) toward gay people were significantly more negative than their deliberatively endorsed feelings. Moreover, while their self-reported gut feelings correlated significantly with Implicit Association Test and Go/No-Go Association Task (Nosek & Banaji, 2001) measures of automatic associations with gay people, their deliberatively endorsed feelings did not. These findings indicate that people can be aware of automatic feelings that are discrepant from their intentionally endorsed attitudes.

A final illustrative example is that of evaluations of major political candidates (e.g., Kerry vs. Bush). Scores on implicit and explicit measures of political attitudes are highly correlated at an individual-difference level ($.50 < rs < .80$). At the same time, people express similar mean preferences on both types of measure (i.e., there is no overall preference in either case, reflecting the even split between Democrats and Republicans in U.S. society). This pattern of results has been replicated using both the IAT (Nosek, 2005; Nosek et al., 2002) and AMP (Affective Misattribution Paradigm) measures (Payne et al., 2005). Because implicit-explicit concordance was present at both an individual-differences level and when examining mean preferences, these results suggest that people are consciously aware of their automatic political attitudes.

Summary

Conclusively identifying unconscious social attitudes is methodologically challenging because implicit measurement does not guarantee that the attitude is inaccessible to conscious awareness. Correlations between implicit and explicit measures improve considerably once social desirability concerns and measurement error are accounted for. Even when dissociations between implicit and explicit measures do result from distinct attitudes, people can still be consciously aware of their second, more automatic attitude. A more rigorous criterion for concluding state unconsciousness is to demonstrate that multiple, correlated implicit measures fail to correspond with self-report measures of both intentionally endorsed attitudes and self-perceived automatic attitudes.

## CONCLUSION

There are many conclusive demonstrations that we are influenced by factors that we are unaware of, and that these have a profound effect on our feelings, judgments, and actions. In contrast, while recent empirical findings provide significant evidence in favor of unconscious mental states, such states are difficult to verify. Because it is difficult to confirm the null hypothesis, operationalizing state unconsciousness as the null makes it difficult to conclusively demonstrate a lack of awareness of such states.

In an effort to further this line of research, we have proposed empirical criteria for verifying unconscious states. We recommend that two criteria be met before concluding that a prime-to-behavior effect was mediated by an unconscious state. First, independent evidence is needed regarding what sort of state (e.g., self-concept, attitude, mood, goals, perceived norms) mediates the priming effect. Second, a self-report measure of that state should be unresponsive to the priming manipulation, yet mediate the effects of a more explicit manipulation (Bargh et al., 2001).

With regard to social attitudes, we have proposed that unconsciousness can be concluded if multiple, correlated implicit measures failed to correspond with self-report measures of both intentionally endorsed attitudes and self-perceived automatic attitudes. Such dissociations could either take the form of low individual differences correlations (e.g., no correlation between implicit and explicit measures of attitudes towards flowers) or very different mean preferences (e.g., an on average explicit preference for evolution over creationism, but on average an implicit preference for creationism over evolution).

At this point, it seems important to consider the broader implications of the conclusive evidence currently available for unconscious social cognition—that for unconscious influences. What implications do unconscious influences hold for theories in which behavior is the product of conscious mental processes?

One possibility is that conscious deliberation and unconscious influences operate more or less independently of one another. Sometimes behavior is driven by consciously endorsed attitudes and assessments of situational norms, each of which in turn contributes to our conscious intention to carry out the behavior (Ajzen, 1985; Fishbein & Ajzen, 1975). At other times, our behavior is the consequence of factors that exert an unconscious influence. For example, sometimes we act competitively because we consciously desire to compete, and at other times because we have been nonconsciously primed to compete by stimuli in the environment.

It seems likely, however, that conscious and unconscious cognition are related in interesting ways (Bargh et al., 2001; Ouellette & Wood, 1998; Wilson, 2002). Our environments may unconsciously influence us by activating dispositions that we consciously endorse. For example, a student may have competition goals automatically activated in school situations because she has habitually competed with other students in the past (Bargh et al., 2001). This automatic activation may reflect the gradual habitualization of her consciously chosen goal to compete in academic situations. As a result, she may be aware of currently feeling competitive (i.e., she is state conscious), without knowing the reason why (i.e., she is unconscious of the influence of the immediate situation).

There is another way in which both unconscious influences and conscious processes may be in effect much of the time. That is, consistent with the theories of reasoned action and planned behavior (Ajzen, 1985; Fishbein & Ajzen, 1975), our behaviors are proximally caused by conscious attitudes, perceived norms, and intentions—but these conscious states are themselves determined by distal factors we are not always conscious of. For instance, a man might be aware that he is voting for a candidate because he personally likes the candidate and his friends support the candidate, yet remain unaware that the candidate's height is exerting an influence on his preference. This idea finds support in research on the limited introspective access people have into the causes of their conscious attitudes (Wilson & LaFleur, 1995; Wilson et al., 1984; see also Haidt, 2001).

In the end, there may be surprisingly little discrepancy between unconscious influences and theories in which conscious states play a central role in human action. Self-reported intentions, attitudes toward the behavior and perceived social norms are powerful predictors of behaviors ranging from driving violations and smoking marijuana to using birth control pills, having an abortion, organ donation, occupational choices, and voting (for meta-analytic reviews, see Armitage & Conner, 2001; Randall & Wolff, 1994; Sheppard, Hartwick, & Warshaw, 1988). But all of this is fully consistent with the idea that conscious attitudes, beliefs and intentions 1) can become automatized (and once automatic can become difficult to control) and 2) are themselves subject to unconscious influences.

To summarize, we have distinguished between two important types of unconscious social cognition: unconsciousness of the influences on judgments and actions, and unconscious of the mental states that give rise to judgments and actions.

Unconscious influences are corroborated by conclusive evidence from a variety of experimental paradigms. While there is some significant evidence of unconscious states, such states are more difficult to verify. We have proposed new criteria aimed at providing conclusive evidence of state unconsciousness. We have further argued that unconscious influences on feelings, judgments and actions are in many respects compatible with theories in which conscious states are important causes of behavior.

*Eric Luis Uhlmann*
*Kellogg School of Management*
*Northwestern University*
*Evanston, IL 60208-2001*
*e-uhlmann@northwestern.edu*

*David A. Pizarro*
*Cornell University*
*dap54@cornell.edu*

*Paul Bloom*
*Yale University*
*Paul.Bloom@yale.edu*

## NOTE

[1] As yet unanswered questions include potential individual and cultural differences in the ability to accurately introspect about one's mental states. While speculative, it seems possible that individuals who score low on self-deception (Paulhus, 1984) and high in need for cognition (Cacioppo & Petty, 1982) are more effective at introspecting.

## REFERENCES

Aarts, H., Gollwitzer, P.M., & Hassin, R.R. 2004. Goal Contagion: Perceiving is for Pursuing. *Journal of Personality and Social Psychology*, 87, 23–37.

Ajzen, I. 1985. From Intentions to Actions: A Theory of Planned Behavior. In J. Kuhl & J. Beckmann (Eds.), *Action Control: From Cognition to Behavior* (pp. 11–38). Berlin: Springer-Verlag.

Ajzen, I., & Fishbein, M. 1977. Attitude-Behavior Relations: A Theoretical Analysis and Review of Empirical Research. *Psychological Bulletin*, 84, 888–918.

Anderson, C.A., & Bushman, B.J. 2002. The Effects of Media Violence on Society. *Science*, 295, 2377–2378.

Armitage, C.J., & Conner, M. 2001. Efficacy of the Theory of Planned Behavior: A Meta-Analytic Review. *British Journal of Social Psychology*, 40, 471–499.

Armstrong, D.M. 1980. *The Nature of Mind and Other Essays*. Ithaca, New York: Cornell University Press.

Asendorpf, J.B., Banse, R., & Mücke, D. 2002. Double Dissociation between Implicit and Explicit Personality Self-Concept: The Case of Shy Behavior. *Journal of Personality and Social Psychology*, 83, 380–393.

Atkinson, J.W., & Birch, D. 1970. *A Dynamic Theory of Action*. New York: Wiley.

Baars, B.J. 1997. Contrastive Phenomenology: A Thoroughly Empirical Approach to Consciousness. In N. Block, O. Flanagan, & G. Guzeldere (Eds.) *The Nature of Consciousness: Philosophical Debates* (pp. 187–202). Cambridge: MIT Press.

Baldwin, M.W., Carrell, S.E., & Lopez, D.F. 1990. Priming Relationship Schemas: My Advisor and the Pope are Watching Me from the Back of My Mind. *Journal of Experimental Social Psychology*, 26, 435–454.

Banaji, M.R. 2001. Implicit Attitudes can be Measured. In H.L. Roedeger, III, J.S. Nairne, I. Neath, & A. Surprenant (Eds.), *The Nature of Remembering: Essays in Honor of Robert G. Crowder* (pp. 117–150). Washington, DC: American Psychological Association.

Banaji, M.R., & Bhaskar, R. 2000. Implicit Stereotypes and Memory: The Bounded Rationality of Social Beliefs. In D.L. Schacter, & E. Scarry (Eds.). *Memory, Brain, and Belief* (pp. 139–175). Cambridge, MA: Harvard University Press.

Banaji, M.R., & Greenwald, A.G. 1995. Implicit Gender Stereotyping in Judgments of Fame. *Journal of Personality and Social Psychology*, 68, 181–198.

Banse, R., Seise, J., & Zerbes, N. 2001. Implicit Attitudes Towards Homosexuality: Reliability, Validity, and Controllability of the IAT. *Zeitschrift für Experimentelle Psychologie*, 48, 145–160.

Bargh, J.A., & Chartrand, T.L. 1999. The Unbearable Automaticity of Being. *American Psychologist*, 54, 462–479.

Bargh, J.A., Chen, M., & Burrows, L. 1996. Automaticity of Social Behavior: Direct Effects of Trait Construct and Stereotype Activation on Action. *Journal of Personality and Social Psychology*, 71, 230–244.

Bargh, J.A., & Ferguson, M.J. 2000. Beyond Behaviorism: On the Automaticity of Higher Mental Processes. *Psychological Bulletin*, 126, 925–945.

Bargh, J.A., Gollwitzer, P.M., Lee-Chai, A., Barndollar, K., & Troetschel, R. 2001. The Automated Will: Nonconscious Activation and Pursuit of Behavioral Goals. *Journal of Personality and Social Psychology*, 81, 1014–1027.

Bargh, J.A., & Pietromonaco, P. 1982. Automatic Information Processing and Social Perception: The Influence of Trait Information Presented Outside of Conscious Awareness on Impression Formation. *Journal of Personality and Social Psychology*, 43, 437–449.

Bargh, J.A., Raymond, P., Pryor, J.B., & Strack, F. 1995. Attractiveness of the Underling: An Automatic Power→ Sex Association and Its Consequences for Sexual Harassment and Aggression. *Journal of Personality and Social Psychology*, 68, 768–781.

Baron, R.M., & Kenny, D.A. 1986. The Moderator-mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. *Journal of Personality and Social Psychology*, 51, 1173–1182.

Boniecki, K.A., & Jacks, J.Z. 2002. The Elusive Relationship between Measures of Implicit and Explicit Prejudice. *Representative Research in Social Psychology*, 26, 1–14.

Bosson, J.K., Swann, W.B., Jr., & Pennebaker, J.W. 2000. Stalking the Perfect Measure of Implicit Self-Esteem: The Blind Man and the Elephant Revisited? *Journal of Personality and Social Psychology*, 79, 631–643.

Brunel, F.F., Tietje, B.C., & Greenwald, A.G. 2004. Is the Implicit Association Test a Valid and Valuable Measure of Implicit Consumer Social Cognition? *Journal of Consumer Psychology*, 14, 385–404.

Cacioppo, J.T., & Petty, R.E. 1982. The Need for Cognition. *Journal of Personality and Social Psychology*, 42, 116–131.

Campbell, D.T., & Fiske, D.W. 1959. Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix. *Psychological Bulletin*, 56, 81–105.

Cesario, J., Plaks, J.E., & Higgins, E.T. 2006. Automatic Social Behavior as Motivated Preparation to Interact. *Journal of Personality and Social Psychology*, 90, 893–910.

Chartrand, T.L., & Bargh, J.A. 1996. Automatic Activation of Impression Formation and Memorization Goals: Nonconscious Goal Priming Reproduces Effects of Explicit Task Instructions. *Journal of Personality and Social Psychology*, 71, 464–478.

Chartrand, T.L., & Bargh, J.A. 1999. The Chameleon Effect: The Perception-Behavior Link and Social Interaction. *Journal of Personality and Social Psychology*, 76, 893–910.

Chartrand, T.L., van Baaren, R.B., Bargh, J.A. 2006. Linking Automatic Evaluation to Mood and Information Processing Style: Consequences for Experienced Affect, Impression Formation, and Stereotyping. *Journal of Experimental Psychology: General*, 135, 70–77.

Chen, S., Lee-Chai, A.Y., & Bargh, J.A. 2001. Relationship Orientation as a Moderator of the Effects of Social Power. *Journal of Personality and Social Psychology*, 80, 173–187.

Chugh, D. 2004. Societal and Managerial Implications of Implicit Social Cognition: Why Milliseconds Matter. *Social Justice Research*, 17, 203–222.

Cook, S.W., & Selltiz, C. 1964. A Multiple-Indicator Approach to Attitude Measurement. *Psychological Bulletin*, 62, 36–55.

Cote, J.A., & Buckley, M.R. 1987. Estimating Trait, Method, and Error Variance: Generalizing Across 70 Construct Validation Studies. *Journal of Marketing Research*, 24, 315–318.

Cunningham, W.A., Nezlek, J.B., & Banaji, M.R. 2004. Conscious and Unconscious Ethnocentrism: Revisiting the Ideologies of Prejudice. *Personality and Social Psychology Bulletin*, 30, 1332–1346.

Cunningham, W.A., Preacher, K.J., & Banaji, M.R. 2001. Implicit Attitude Measures: Consistency, Stability, and Convergent Validity. *Psychological Science*, 12, 163–170.

Czopp, A.M., Monteith, M.J., Zimmerman, R.S., & Lynam, D.R. 2004. Implicit Attitudes as Potential Protection from Risky Sex: Predicting Condom Use with the IAT. *Basic and Applied Social Psychology*, 26, 227–237.

Dasgupta, N. 2004. Implicit Ingroup Favoritism, Outgroup Favoritism, and their Behavioral Manifestations. *Social Justice Research*, 17, 143–169.

De Houwer, J. 2001. A Structural and Process Analysis of the Implicit Association Test. *Journal of Experimental Social Psychology*, 37, 443–534.

DeMarree, K.G., Wheeler, S.C., & Petty, R.E. 2005. Priming a New Identity: Self-Monitoring Moderates the Effects of Non-Self Primes on Self-Judgments and Behavior. *Journal of Personality and Social Psychology*, 89, 657–671.

Dijksterhuis, A., & Bargh, J.A. 2001. The Perception-Behavior Expressway: The Automatic Effects of Social Perception on Social Behavior. In M.P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 33, pp. 1–40). San Diego, CA: Academic Press.

Dijksterhuis, A., & van Knippenberg, A. 1998. The Relation between Perception and Behavior, or How to Win a Game of Trivial Pursuit. *Journal of Personality and Social Psychology*, 74, 865–877.

Dovidio, J.F., Kawakami, K., & Beach, K.R. 2001. Implicit and Explicit Attitudes: Examination of the Relationship between Measures of Intergroup Bias. In R. Brown & S.L. Gaertner (Eds.) *Blackwell Handbook of Social Psychology: Intergroup Processes* (pp. 175–197). Malden, MA: Blackwell.

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.

Draine, S.C., & Greenwald, A.G. 1998. Replicable Unconscious Semantic Priming. *Journal of Experimental Psychology: General*, 127, 286–303.

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.

Egloff, B., & Schmukle, S.C. 2002. Predictive Validity of an Implicit Association Test for Assessing Anxiety. *Journal of Personality and Social Psychology*, 83, 1441–1455.

Erb, H., Bioy, A., & Hilton, D.J. 2002. Choice Preferences without Inferences: Subconscious Priming of Risk Attitudes. *Journal of Behavioral Decision Making*, 15, 251–262.

Fazio, R.H., Jackson, J., Dunton, B., & Williams, C. 1995. Variability in Automatic Activation as an Unobtrusive Measure of Racial Attitudes: A Bona Fide Pipeline? *Journal of Personality and Social Psychology*, 69, 1013–1027.

Fazio, R.H., & Olson, M.A. 2003. Implicit Measures in Social Cognition Research: Their Meaning and Uses. *Annual Review of Psychology*, 54, 297–327.

Fishbein, M., & Ajzen, I. 1975. *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*. Reading, MA: Addison-Wesley.

Fitzsimons, G.M., & Bargh, J.A. 2003. Thinking of You: Nonconscious Pursuit of Interpersonal Goals Associated with Relationship Partners. *Journal of Personality and Social Psychology*, 84, 148–163.

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). San Diego, CA: Academic Press.

Gawronski, B., & Bodenhausen, G.V. 2006. Associative and Propositional Processes in Evaluation: An Integrative Review of Implicit and Explicit Attitude Change. *Psychological Bulletin*, 132, 692–731.

Gawronski, B., LeBel, E.P., & Peters, K.R. 2007. What do Implicit Measures Tell Us? Scrutinizing the Validity of Three Common Assumptions. *Perspectives on Psychological Science*, 2, 181–193.

Gawronski, B., Hofmann, W., & Wilbur, C.J. 2006. Are "Implicit" Attitudes Unconscious? *Consciousness and Cognition*, 15, 485–499.

Gollwitzer, P.M., & Moskowitz, G.B. 1996. Goal Effects on Action and Cognition. In E.T. Higgins & A.W. Kruglanski (Eds.), *Social Psychology: Handbook of Basic Principles* (pp. 361–399). New York: Guilford Press.

Greenwald, A.G., & Banaji, M.R. 1995. Implicit Social Cognition: Attitudes, Self-Esteem, and Stereotypes. *Psychological Review*, 102, 4–27.

Greenwald, A.G., McGhee, D.E., & Schwartz, J.L.K. 1998. Measuring Individual Differences in Implicit Cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, 74, 1464–1480.

Haidt, J. 2001. The Emotional Dog and its Rational Tail: A Social Intuitionist Approach to Moral Judgment. *Psychological Review*, 108, 814–834.

Heath, A. 1976. *Rational Choice and Social Exchange*. Cambridge: Cambridge University Press.

Hertel, G., & Kerr, N.L. 2001. Priming In-Group Favoritism: The Impact of Normative Scripts in the Minimal Group Paradigm. *Journal of Experimental Social Psychology*, 37, 316–324.

Higgins, E.T., Rholes, W.S., & Jones, C.R. 1977. Category Accessibility and Impression Formation. *Journal of Experimental Social Psychology*, 13, 141–154.

Hofman, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. 2005. A Meta-Analysis on the Association between the Implicit Association Test and Explicit Self-Report Measure. *Personality and Social Psychology Bulletin*, 31, 1369–1385.

Holender, D. 1986. Semantic Activation without Conscious Identification in Dichotic Listening, Parafoveal Vision, and Visual Masking: A Survey and Appraisal. *Behavioral and Brain Sciences*, 9, 1–23.

Huesmann, L.R., Eron, L.D., Lefkowitz, M.M., & Walder, L.O. 1973. Television Violence and Aggression: The Causal Effect Remains. *American Psychologist*, 28, 617–620.

Jacoby, L.L. 1991. A Process Dissociation Framework: Separating Automatic from Intentional Uses of Memory. *Journal of Memory and Language*, 30, 513–541.

Jordan, C.H., Spencer, S.J., Zanna, M.P., Hoshino-Browne, E., & Correll, J. 2003. Secure and Defensive High Self-Esteem. *Journal of Personality and Social Psychology*, 85, 969–978.

Karpinski, A. & Hilton, J.L. 2001. Attitudes and the Implicit Association Test. *Journal of Personality and Social Psychology*, 81, 774–788.

Kawakami, K., & Dovidio, J.F. 2001. The Reliability of Implicit Stereotyping. *Personality and Social Psychology Bulletin*, 27, 212–225.

Kawakami, K., Dovidio, J.F., & Dijksterhuis, A. 2003. Effect of Social Category Priming on Personal Attitudes. *Psychological Science*, 14, 315–319.

Kay, A.C., & Ross, L. 2003. The Perceptual Push: The Interplay of Implicit Cues and Explicit Situational Construals on Behavioral Intentions in the Prisoner's Dilemma. *Journal of Experimental Social Psychology*, 39, 634–643.

Kay, A.C., Wheeler, S.C., Bargh, J.A., & Ross, L. 2004. Material Priming: The Influence of Mundane Physical Objects on Situational Construal and Competitive Behavior Choice. *Organizational Behavior and Human Decision Processes*, 95, 83–96.

Kihlstrom, J.F. 2004. Implicit Methods in Social Psychology. In C. Sansone, C. Morf, & A. Panter (Eds.). *Handbook of Methods in Social Psychology*. Thousand Oaks, CA: Sage.

Kihlstrom, J.F., Mulvaney, S., Tobias, B.A., & Tobis, I.P. 2000. The Emotional Unconscious. In E. Eich (Ed.), *Cognition and Emotion*. New York: Oxford University Press.

Kraut, R.E., & Lewis, S.H. 1982. Person Perception and Self-Awareness: Knowledge of Influences on One's Own Judgments. *Journal of Personality and Social Psychology*, 42, 448–460.

Krueger, J.I., & Funder, D.C. 2004. Towards a Balanced Social Psychology: Causes, Consequences, and Cures for the Problem-Seeking Approach to Social Behavior and Cognition. *Behavioral and Brain Sciences*, 27, 313–327.

Krueger, J. 1998 The Bet on Bias: A Foregone Conclusion?. Psycoloquy, 9(46).

Lakin, J., & Chartrand, T.L. 2003. Using Nonconscious Behavioral Mimickry to Create Affiliation and Rapport. *Psychological Science*, 14, 334–339.

Lemm, K.M. 2000. *Personal and Social Motivation to Respond Without Prejudice: Implications for Implicit and Explicit Attitudes and Behavior*. Doctoral dissertation, Yale University.

Levesque, C., & Pelletier, L.G. 2003. On the Investigation of Primed and Chronic Autonomous and Heteronomous Motivational Orientations. *Personality and Social Psychology Bulletin*, 29, 1570–1584.

Libet, B. 1985. Unconscious Cerebral Initiative and the Role of Conscious Will in Voluntary Action. *Behavioral and Brain Sciences*, 8, 529–566.

Lombardi, W.J., Higgins, E.T., & Bargh, J.A. 1987. The Role of Consciousness in Priming Effects on Categorization: Assimilation Versus Contrast as a Function of Awareness of the Priming Task. *Personality and Social Psychology Bulletin*, 13, 411–429.

Maison, D., Greenwald, A.G., & Bruin, R. 2001. The Implicit Association Test as a Measure of Consumer Attitudes. *Polish Psychological Bulletin*, 2, 79.

Martin, L.L., Seta, J.J., & Crelia, R.A. 1990. Assimilation and Contrast as a Function of People's Willingness and Ability to Expend Effort in Forming an Impression. *Journal of Personality and Social Psychology*, 59, 27–37.

McCoy, S.K., & Major, B. (in press). Priming Meritocracy and the Psychological Justification of Inequality. *Journal of Experimental Social Psychology*.

McGregor, I., & Marigold, D.C. 2003. Defensive Zeal and the Uncertain Self: What Makes You So Sure? *Journal of Personality and Social Psychology*, 85, 838–852.

Monteith, M.J., & Voils, C.I. 1998. Proneness to Prejudiced Responses: Toward Understanding the Authenticity of Self-Reported Discrepancies. *Journal of Personality and Social Psychology*, 75, 901–916.

Moskowitz, G.B., Gollwitzer, P.M., Wasel, W., & Schaal, B. 1999. Preconscious Control of Stereotype Activation through Chronic Egalitarian Goals. *Journal of Personality and Social Psychology*, 7, 167–184.

Moskowitz, G.B., & Roman, R.J. 1992. Spontaneous Trait Inferences as Self-Generated Primes: Implications for Conscious Social Judgment. *Journal of Personality and Social Psychology*, 62, 728–738.

Moskowitz, G.B., & Skurnik, I.W. 1999. Contrast Effects as Determined by the Type of Prime: Trait Versus Exemplar Primes Initiate Processing Strategies that Differ in How Accessible Constructs are Used. *Journal of Personality and Social Psychology*, 6, 911–927.

Mussweiler, T., Rüter, K., & Epstude, K. 2004. The Man Who wasn't There—Subliminal Social Standards Influence Self-Evaluation. *Journal of Experimental Social Psychology*, 40, 689–696.

Nelson, L.D., & Norton, M. 2005. From Student to Superhero: Situational Primes Shape Future Helping. *Journal of Experimental Social Psychology*, 41, 423–430.

Newman, L.S., & Uleman, J.S. 1990. Assimilation and Contrast Effects in Spontaneous Trait Inference. *Personality and Social Psychology Bulletin*, 16, 224–240.

Nisbett, R.E., & Bellows, N. 1977. Verbal Reports about Causal Influences on Social Judgments: Private Access vs. Public Theories. *Journal of Personality and Social Psychology*, 35, 613–624.

Nisbett, R.E., & Wilson, T.D. 1977. Telling more than We Can Know: Verbal Reports on Mental Processes. *Psychological Review*, 84, 231–259.

Nosek, B.A. 2002. *Moderators of the Relationship between Implicit and Explicit Attitudes*. Doctoral dissertation, Yale University.

Nosek, B.A. 2005. Moderators of the Relationship between Implicit and Explicit Evaluation. *Journal of Experimental Psychology: General*, 134, 565–584.

Nosek, B.A., & Banaji, M.R. 2001. The Go/No-go Association Task. *Social Cognition*, 19, 625–666.

Nosek, B.A., & Banaji, M.R. 2002. [Polish Language] (At Least) Two Factors Moderate the Relationship between Implicit and Explicit Attitudes. In R.K. Ohme & M. Jarymowicz (Eds.), *Natura Automatyzmow* (pp. 49–56). Warszawa: WIP PAN & SWPS.

Nosek, B.A., Banaji, M.R., & Greenwald, A.G. 2002. Harvesting Implicit Group Attitudes and Beliefs from a Demonstration Website. *Group Dynamics*, 6, 101–115.

Nosek, B.A., & Hansen, J. 2006. *The Associations in Our Heads Belong to Us: Measuring the Multifaceted Attitude Construct in Implicit Social Cognition*. Manuscript submitted for publication.

Olson, M.A., & Fazio, R.H. 2001. Implicit Attitude Formation through Classical Conditioning. *Psychological Science*, 12, 413–417.

Olson, M.A., & Fazio, R.H. 2003. Relations between Implicit Measures of Prejudice: What are We Measuring? *Psychological Science*, 14, 636–639.

Olson, M.A., & Fazio, R.H. 2004. Trait Inferences as a Function of Automatically-Activated Racial Attitudes and Motivation to Control Prejudiced Reactions. *Basic and Applied Social Psychology*, 26, 1–11.

Olson, M.A., Fazio, R.H., & Hermann, T.D. 2007. Reporting Tendencies Underlie Discrepancies between Implicit and Explicit Measures of Self-Esteem. *Psychological Science*, 18, 287–291.

Ouellette, J.A., & Wood, W. 1998. Habit and Intention in Everyday Life: The Multiple Processes by Which Past Behavior Predicts Future Behavior. *Psychological Bulletin*, 124, 54–74.

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., 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., Burkley, M., & Stokes, M.B. 2008. Why do Implicit and Explicit Attitude Tests Diverge? The Role of Structural Fit. *Journal of Personality and Social Psychology*, 94, 16–31.

Paulhus, D.L. 1984. Two-Component Models of Socially Desirable Responding. *Journal of Personality and Social Psychology*, 46, 598–609.

Phillips, D.P. 1979. Suicide, Motor Vehicle Fatalities, and the Mass Media: Evidence Toward a Theory of Suggestion. *American Journal of Sociology*, 84, 1150–1174.

Phillips, D.P. 1983. The Impact of Mass Media Violence on U.S. Homicides. *American Sociological Review*, 48, 560–568.

Randall, D.M., & Wolff, J.A. 1994. The Time Interval in the Intention-Behaviour Relationship: Meta-Analysis. *British Journal of Social Psychology*, 33, 405–418.

Ranganath, K.A., Smith, C.T., & Nosek, B.A. (in press). Distinguishing Automatic and Controlled Components of Attitudes from Direct and Indirect Measurement. *Journal of Experimental Social Psychology*.

Reber, A.S. 1989. More Thoughts on the Unconscious: Reply to Brody and to Lewicki and Hill. *Journal of Experimental Psychology: General*, 118, 242–244.

Reingold, E.M., & Merikle, P.M. 1988. Using Direct and Indirect Measures to Study Perception without Awareness. *Perception and Psychophysics*, 44, 563–575.

Reingold, E.M., & Merikle, P.M. 1990. On the Inter-Relatedness of Theory and Measurement in the Study of Unconscious Processes. *Mind and Language*, 5, 9–28.

Roediger, H.L. 1990. Implicit Memory: Retention without Remembering. *American Psychologist*, 45, 1043–1056.

Rudman, L.A., Ashmore, R.D., & Gary, M.L. 2001. "Unlearning" Automatic Biases: The Malleability of Implicit Prejudice and Stereotypes. *Journal of Personality and Social Psychology*, 81, 856–868.

Rudman, L.E., & Heppen, J. 2003. Implicit Romantic Fantasies and Women's Interest in Personal Power: A Glass Slipper Effect? *Personality and Social Psychology Bulletin*, 29, 1357–1370.

Schacter, D.L. 1992. Understanding Implicit Memory: A Cognitive Neuroscience Approach. *American Psychologist*, 47, 559–569.

Shah, J. 2003a. Automatic for the People: How Representations of Significant Others Implicitly Affect Goal Pursuit. *Journal of Personality and Social Psychology*, 84, 661–681.

Shah, J. 2003b. The Motivated Looking Glass: How Significant Others Implicitly Affect Goal Appraisals. *Journal of Personality and Social Psychology*, 85, 424–439.

Shah, J.Y., & Kruglanski, A.W. 2002. Priming Against Your Will: How Goal Pursuit is Affected by Accessible Alternatives. *Journal of Experimental Social Psychology*, 38, 368–383.

Shah, J.Y., & Kruglanski, A.W. 2003. When Opportunity Knocks: Bottom-Up Priming of Goals by Means and Its Effects on Self-Regulation. *Journal of Personality and Social Psychology*, 84, 1109–1122.

Sheppard, B.H., Hartwick, J., & Warshaw, P.R. 1988. The Theory of Reasoned Action: A Meta-Analysis of Past Research with Recommendations for Modifications and Future Research. *Journal of Consumer Research*, 15, 325–343.

Sherman, S.J., Rose, J.S., Koch, K., Presson, C.C., & Chassin, L. 2003. Implicit and Explicit Attitudes toward Cigarette Smoking: The Effects of Context and Motivation. *Journal of Social and Clinical Psychology*, 22, 13–39.

Singer, J.L., & Singer, D.G. 1983. Psychologists Look at Television: Cognitive, Developmental, Personality, and Social Policy Implications. *American Psychologist*, 38, 826–834.

Smith, C.T., & Nosek, B.A. 2005. *Gut Reactions and Considered Feelings: Relations between Self-Report and Implicit Attitudes*. Poster presented at the 2005 conference for the Society of Personality and Social Psychology.

Spencer, S.J., Fein, S., Wolfe, C.T., Fong, C., & Dunn, M.A. 1998. Automatic Activation of Stereotypes: The Role of Self-Image Threat. *Personality and Social Psychology Bulletin*, 24, 1139–1152.

Strack, F., Schwarz, N., Bless, H., Kübler, A., & Wänke, M. 1993. Awareness of the Influence as a Determinant of Assimilation Versus Contrast. *European Journal of Social Psychology*, 23, 53–62.

Stapel, D.A., & Blanton, H. 2004. From Seeing to Believing: Subliminal Social Comparisons Affect Implicit and Explicit Self-Evaluations. *Journal of Personality and Social Psychology*, 87, 468–481.

Strahan, E.J., Spencer, S.J., & Zanna, M.P. 2002. Subliminal Priming and Persuasion: Striking While the Iron is Hot. *Journal of Experimental Social Psychology*, 38, 556–568.

Swanson, J.E., Rudman, L.A., & Greenwald, A.G. 2001. Using the Implicit Association test to Investigate Attitude-Behavior Consistency for Stigmatized Behavior. *Cognition and Emotion*, 15, 207–230.

Tamir, M., Robinson, M.D., Clore, G.L., Martin, L.L., Whitaker, D.J. 2004. Are We Puppets on a String?: The Contextual Meaning of Unconscious Expressive Cues. *Personality and Social Psychology Bulletin*, 30, 237–249.

Teachman, B., & Woody, S. 2003. Automatic Processing in Spider Phobia: Implicit Fear Associations over the Course of Treatment. *Journal of Abnormal Psychology*, 112, 100–109.

Towles-Schwen, T., & Fazio, R.H. 2003. Choosing Social Situations: The Relation between Automatically-Activated Racial Attitudes and Anticipated Comfort Interacting with African Americans. *Personality and Social Psychology Bulletin*, 29, 170–182.

Uhlmann, E., & Nosek, B.A. 2005. *My Culture Made Me Do It: Lay Theories of Socialization and Responsibility*. Unpublished manuscript, Yale University.

Van Baaren, R.B., Maddux, W.W., Chartrand, T.L., De Bouter, C., & Van Knippenberg, A. 2003. It Takes Two to Mimic: Behavioral Consequences of Self-Construals. *Journal of Personality and Social Psychology*, 84, 1093–1102.

Voils, C.I., & Monteith, M.J. 2004. *Naive Theories of Racial Bias: Measuring and Influencing Beliefs about the Extent to Which Race Colors Perceptions*. Unpublished manuscript, University of Kentucky.

Wegner, D.M. 2002. *The Illusion of Conscious Will*. Cambridge, MA: The MIT Press.

Wheeler, S.C., DeMarree, K.G., & Petty, R.E. 2005. The Roles of the Self in Priming-To-Behavior Effects. In A. Tesser, J.V. Wood, & D.A. Stapel (Eds.), *On Building, Defending, and Regulating the Self: A Psychological Perspective* (pp. 245–271). New York: Psychology Press.

Wheeler, C.S., & Petty, R.E. 2001. The Effects of Stereotype Activation on Behavior: A Review of Possible Mechanisms. *Psychological Bulletin*, 127, 797–826.

Wilson, T.D. 2002. *Strangers to Ourselves: Discovering the Adaptive Unconscious*. Cambridge, MA: Belknap Press/Harvard University Press.

Wilson, T.D., & Brekke, N. 1994. Mental Contamination and Mental Correction: Unwanted Influences on Judgments and Evaluations. *Psychological Bulletin*, 116, 117–142.

Wilson, T.D., Dunn, D.S., Bybee, J.A., Hyman, D.B., & Rotondo, J.A. 1984. Effects of Analyzing Reasons on Attitude—Behavior Consistency. *Journal of Personality and Social Psychology*, 47, 5–16.

Wilson, T.D., Dunn, D.S., Kraft, D., & Lisle, D.J. 1989. Introspection, Attitude Change, and Attitude-Behavior Consistency: The Disruptive Effects of Explaining Why We Feel the Way We Do. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 22, pp. 287–343). Orlando, FL: Academic Press.

Wilson, T.D., & LaFleur, S.J. 1995. Knowing What You'll Do: Effects of Analyzing Reasons on Self-Prediction. *Journal of Personality and Social Psychology*, 68, 21–35.

Wilson, T.D., Laser, P., & Stone, J. 1982. Judging the Predictors of One's Own Mood: Accuracy and the Use of Shared Theories. *Journal of Experimental Social Psychology*, 18, 537–556.

Wilson, T.D., Lindsey, S., & Schooler, T.Y. 2000. A Model of Dual Attitudes. *Psychological Review*, 107, 101–126.

Wilson, T.D., Lisle, D.J., Schooler, J.W., Hodges, S.D., Klaaren, K.J., & LaFleur, S.J. 1993. Introspecting about Reasons can Reduce Post-Choice Satisfaction. *Personality and Social Psychology Bulletin*, 19, 331–339.

Wilson, T.D., & Stone, J.I. 1985. Limitations of Self-Knowledge: More on Telling more than We Can Know. In P. Shaver (Ed.), *Review of Personality and Social Psychology* (Vol. 6, pp. 167–183). Beverly Hills, CA: Sage.

Winkielman, P., Berridge, K.C., & Wilbarger, J. 2005. Unconscious Affective Reactions to Masked Happy Versus Angry Faces Influence Consumption Behavior and Judgments of Value. *Personality and Social Psychology Bulletin*, 31, 121–135.

Wright, P., & Rip, P.D. 1981. Retrospective Reports on the Causes of Decisions. *Journal of Personality and Social Psychology*, 40, 601–614.

Zajonc, R.B. 1980. Feeling and Thinking: Preferences Need No Inferences. *American Psychologist*, 35, 151–175.