

Principles, processes, and puzzles of social cognition: An introduction for the special issue on social cognitive neuroscience

Matthew D. Lieberman

Department of Psychology, Franz Hall, University of California, Los Angeles, CA 90095-1563, USA

Received 14 July 2005; accepted 14 July 2005

Available online 19 August 2005

This article introduces the special issue of *NeuroImage* focused on social cognitive neuroscience. Social psychology has a rich history of making sense of the often paradoxical aspects of social cognition and the social world. This article reviews the principles, processes, and puzzles of social cognition and behavior that have been examined by social psychologists for decades. Five principles of social cognition and behavior are reviewed including: (1) the power of the situation over behavior, (2) blindness for situational influences, (3) social perception and self-perception are constructive processes, (4) blindness for the constructed nature of social and self-perception, and (5) self-processes are social. Four processes of social cognition are reviewed including: (1) cognitive architecture; (2) automaticity and control; (3) motivated reasoning; and (4) accessibility, frames, and expectations. Finally, five areas of social cognition that contain enduring puzzles are described including (1) the self, (2) attitudes, (3) reflective social cognition, (4) automatic social cognition, and (5) social motives. In several of the areas of study reviewed, cognitive neuroscience is well positioned to make important contributions to these research traditions either by allowing for new tests of hypotheses or by allowing for unobtrusive measurement of social cognitive processes.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Social cognitive neuroscience; Automaticity vs. control; Self; Situational pressure; Social perception; Social cognition

Social cognitive neuroscience is a recently developing area of research that strives to answer fundamental questions about the nature of human social cognition by adding cognitive neuroscience techniques to the arsenal of experimental methods already used by social psychologists and other social scientists to study these questions (Ochsner and Lieberman, 2001). Researchers have examined the biological consequences of social factors for more than two decades (Blascovich and Mendes, 2000; Cacioppo et al., 2002; Taylor et al., 2000). Social cognitive neuroscience has added to this research by studying the neurocognitive mechanisms supporting social

cognition itself, rather than focusing primarily on the physiological after effects of social cognition.

A number of isolated studies relying on techniques such as event-related potentials (ERP) and neuropsychology conducted in the 1990s can rightfully be said to be the forerunners to current social cognitive neuroscience research. Cacioppo et al. (1996) used ERPs to localize regions of cortex that are differentially involved in evaluative and non-evaluative beliefs, suggesting that different kinds of cognition support these two kinds of attitudes. Klein et al. (1996) examined a patient with temporary amnesia, caused by traumatic brain injury, to determine whether episodic memory is critical to the integrity of self-knowledge. This patient was observed to have intact self-knowledge despite the lack of memory for the behaviors that presumably contributed to this self-knowledge in the first place.

Social cognitive neuroscience was catalyzed into a coherent area of study as researchers turned to functional magnetic resonance imaging (fMRI). Early imaging studies focused on stereotypes (Hart et al., 2000; Phelps et al., 2000), self-knowledge (Kelley et al., 2002), and theory of mind (Baron-Cohen et al., 1994; Frith and Frith, 1999); however, work has now extended into several areas of social psychological inquiry including self-serving biases (Blackwood et al., 2003), self-awareness (Gusnard et al., 2001; Keenan et al., 2001; Eisenberger et al., in press), judgment and decision-making (De Quervain et al., 2004; Sanfey et al., 2003), cooperation (Kosfeld et al., 2005; Rilling et al., 2004), self-schemas (Lieberman et al., 2004), person knowledge (Mitchell et al., 2004a), social exclusion (Eisenberger et al., 2003), attitudinal evaluation (Cunningham et al., 2003; Wood et al., 2005), regulation of stereotypes (Amodio et al., 2003; Lieberman et al., 2005; Richeson et al., 2003; Wheeler and Fiske, 2005), expectancy effects (Lieberman et al., 2004; Petrovic et al., 2002; Wager et al., 2004), relational cognition (Aron et al., 2005; Iacoboni et al., 2004), empathy (Carr et al., 2003; Singer et al., 2004), and emotional reappraisal (Beauregard et al., 2001; Ochsner et al., 2002). This special issue devoted to social cognitive neuroscience brings new light to these existing themes (Cunningham, Espinet, DeYoung, and Zelazo, this issue; Mitchell, Banaji, and Macrae, this issue; Ochsner et al., this issue; Sander et al., this issue) and tackles areas new to social cognitive neuroscience such as

E-mail address: lieber@ucla.edu.

Available online on ScienceDirect (www.sciencedirect.com).

attribution (Harris, Todorov, and Fiske, this issue; Heberlein and Saxe, this issue), attachment (Gillath, Bunge, Shaver, Wendelken, and Mikulincer, this issue), self-esteem (Pruessner et al., this issue), and intention planning (Ouden, Frith, Frith, and Blakemore, this issue).

One complaint often launched against social cognitive neuroscience research by social psychologists is that cognitive neuroscience takes more from social cognition than it gives in return. fMRI studies of social cognition have been used to identify the social cognitive function of different brain regions such as medial prefrontal cortex (MPFC), superior temporal sulcus (STS), and the amygdala; however, many question whether our understanding of social cognition has been enhanced by knowing their neural correlates.

A second complaint from social psychology is that the work does not always respect the complexities of social psychology. Because the language of social psychology is the language of everyday life, there is sometimes an implicit assumption that the principles of social psychological phenomena can be derived from our lay theories of everyday life. However, as is evident from any introductory social psychology course, social psychology is in many ways the study of how wrong our lay theories of everyday life are (Wilson, 2002; cf. Krueger and Funder, 2004). People think they know much more about the rules and patterns of the social world than they really do. The contributors to this special issue answer this challenge by giving sophisticated treatment to the intricacies of social cognition. To help contextualize this work and introduce social cognition to a cognitive neuroscience audience, the remainder of this article has thus been written as a primer on some of the principles, processes, and puzzles of social cognition identified by social psychologists over the past five decades. These findings are important for social cognitive neuroscience researchers to appreciate, but more importantly, may provide many new avenues of research as these complexities are embraced in future research.

Principles of social cognition

Power of the situation over behavior

If a social psychologist was going to be marooned on a deserted island and could only take one principle of social psychology with him it would undoubtedly be the ‘power of the situation’. All of the most classic studies in the early days of social psychology demonstrated that situations can exert a powerful force over the actions of individuals. Observers often mistake situationally induced behavior as implying that others are evil, stupid, or weak willed, when in fact these individuals are good, decent individuals, just like us, who unwittingly succumb to the power of the situation. The most famous of these studies is Milgram’s (1963) work on obedience to authority in which he demonstrated that individuals are willing to shock a complete stranger at high voltages when pressed by a scientific experimenter who insists that “the experiment must continue”. Asch (1956) conducted conformity experiments in which participants gave conspicuously wrong answers to simple problems because other confederates, who were strangers to the participant, consistently gave the same wrong answer before the participant had an opportunity to respond. Finally, Latane and Darley (1970) found that a person in need is more likely to get help from a single individual than from a group of individuals because

group members tend to remain passive while they look to each other to determine what to do and end up doing nothing because those around them are doing nothing (Miller and Ratner, 1998). In the years since these classic studies, countless others have demonstrated the consequences of groups of all sizes (culture, family, significant others, and strangers) on the behaviors of others.

Blindness for situational influences

If the power of the situation is the first principle of social psychology, a second is that people are largely unaware of the influence of situations on behavior, whether it is their own or someone else’s behavior. Milgram consulted with psychiatrists prior to conducting his obedience studies to ensure that his paradigm would not cause long-term psychological harm to his participants. The psychiatrists indicated that they were not worried because they expected all of the participants to refuse to do anything that could be remotely harmful to another individual. In other words, the psychiatrists did not appreciate the power of the situation and neither did another sample of adults who had the experiment described to them and were asked to predict the point at which they would stop administering the shocks. As seen in Fig. 1, there was virtually no overlap between the predictions made by either psychiatrists or healthy controls and the responses made by actual subjects in the study.

Imagine the consequences that this blindness has for our understanding of the participants in this study. If we start with the assumption that we would have stopped administering shocks in the low end of the shock range and that any normal person would do the same, then it is only natural to conclude that the actual subjects in this study were fundamentally cruel. Of course, given the distribution of actual responses it is likely that we too would have been susceptible to the effects of authority. One thing that most are unlikely to consider in imagining themselves in the Milgram experiment is the slippery slope of justification built into the study. Every time the confederate answers incorrectly, the participant is expected to give a shock that is 15 V higher than the previous shock given. The slippery slope emerges because on each trial subjects are only being asked to give slightly more shock than they were willing to a moment earlier. In order for the subject to decide that “now things have gotten out of hand and I will not continue” the subject has to condemn their own previous behavior. One cannot easily decide that giving 345 V is inappropriate after deciding 330 V was perfectly fine. Thus, there is pressure on the subject to construe each new voltage level as appropriate in order to be consistent with one’s prior behavior (Festinger, 1957). However, this pressure is likely to go unnoticed by observers.

Consequently, although we want to label the subjects willing to administer shocks up to 450 V as cruel people, we are not justified in using this behavior to label them at all because situational pressure, rather than individual dispositions, was the driving force behind these behaviors. Most observers will also miss the actual opportunity to label some of the subjects in terms of their character. The few subjects who stopped administering shocks at 255 V are likely to be the Good Samaritans of the world. It may be hard to imagine, but given the baseline effect of the situation in this study, these individuals can be seen as fighting against the power of the situation to help another person. This research is a powerful illustration of both the power of the situation and the dangerous consequences of our blindness to this power.

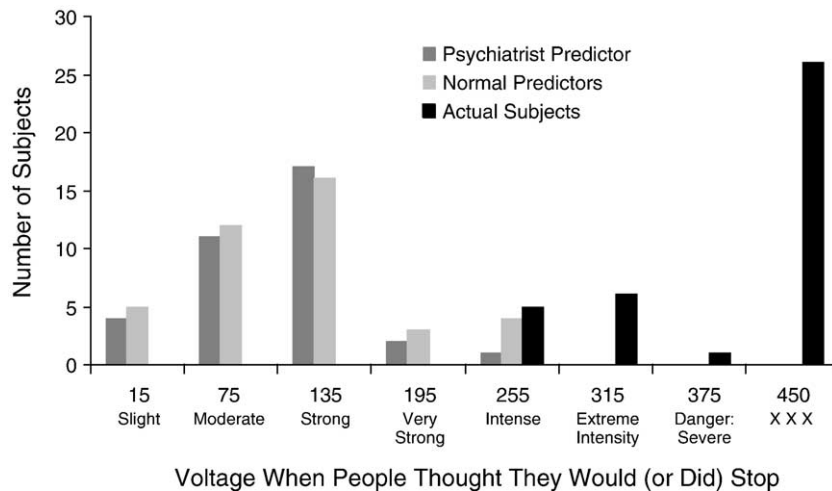


Fig. 1. Distribution of predicted and actual stopping voltages in the Milgram (1963) experiment. Psychiatrists (dark grey bars) and other normal individuals (light grey bars) made predictions about the voltage at which they thought they would stop if they were run through the experimental protocol. None of these individuals thought they would go further than 255 V. Actual subjects (black bars) all shocked to a level of 255 V or more. Several of the individuals that went all the way to 450 V (the highest voltage available) did not stop on their own but continued to shock at 450 V until stopped by the experimenter. The descriptive labels below the voltages (e.g., “Extreme Intensity”, “XXX”) were written on the shock device that the participants used in the experiment.

Social perception and self-perception are constructive processes

It has long been accepted that our experience of color and other perceptual characteristics of the world has as much to do with the organization of our mind as it does with the world itself (Kant, 1781/1965). Perception is thus acknowledged to be a constructive process. Social perception and self-perception are similarly constructive processes. We see the social world filtered through the lenses of all the social learning, thinking, and practicing that we have done throughout our lives as well as by our current and recent mental activity. Our interpretation of others and ourselves is guided by social schemas (Markus, 1977), chronically accessible constructs (Higgins, 1996), stereotypes (Allport, 1954), and self-evaluative motives (Dunning and Cohen, 1992; Tesser, 1988) among others.

In a classic study of social perception, Higgins primed participants with either “adventurousness” or “recklessness” and then had them read a paragraph about Donald that could either be interpreted as adventurous or reckless (Higgins et al., 1977). Participants’ interpretations of Donald were significantly influenced by the priming episode such that they tended to rate Donald in ways consistent with the particular trait construct that was primed. Higgins (1996) subsequently suggested that repeated thinking about a particular construct could promote greater accessibility of that construct at all times (‘chronic accessibility’) and lead that construct to influence the interpretation of behavior even when it had not been primed recently. In other words, if someone has chronic accessibility for recklessness, they would tend to see ambiguous behaviors as more reckless than would other individuals on a regular basis. Thus, it is clear that our social perception, although constrained by reality, is constructed on the basis of currently activated mental representations, motives, and processes.

Blindness for the constructed nature of social and self-perception

Although social and self-perception are constructive processes, they are typically experienced as direct perceptions of reality. When we see an individual that we interpret as being reckless, we

experience the recklessness as ‘out there in the world’ rather than as being constructed on the basis of primes and chronically accessible constructs in our heads. We see reckless people, not people that we have decided are reckless (Hastorf and Cantril, 1954). This can lead to interpersonal misunderstandings as people have different mental processes and active representations influencing their own social perceptions and thus reach different conclusions about social phenomena. Because we are typically unaware of our own constructive contributions to our social perceptions, it is quite difficult to appreciate why others would ever see the social world differently than we do. When others report seeing the social world differently from us, our first reaction is often that the other person must be crazy, stupid, or trying to take advantage of us. Ross and his colleagues (Griffin and Ross, 1991) have described this phenomenon of not recognizing the constructive nature of social perception and thinking badly of those who come to conclusions different from our own as *naïve realism* and have suggested it helps explain conflicts as small as debates over who did more of the household chores (Ross and Sicoly, 1979) and as large as national ideological disputes (Robinson et al., 1995).

Little is known about why naïve realism is pervasive in our social and self-perception. One suggestion has been that automatic social inference processes occur outside of awareness and so the products of these inferences are taken to be part of reality rather than creations of the mind (Lieberman et al., 2002; Neisser, 1976). In support of this idea, when individuals are made aware of the possible link between a priming episode and the subsequent measure on which the prime typically has an effect, the priming effects tend to disappear or even reverse (Moscowitz and Roman, 1992). This suggests that priming effects tend to exert their influence in part because we are unaware of them.

Self-processes are social

It may seem odd that social psychologists, who specialize in analyzing the social world, take the study of the self to be a major area of study for them. The self appears to be an encapsulated,

hermetically sealed entity that is primarily a representation of our most distinctive features that differentiate us from others and to which we have special introspective access that others could never have. Who could know whether and how much I like chocolate ice cream more than me? I have the experience of enjoying or not enjoying it and thus it seems that only I have access to that information. Although there are undoubtedly aspects of self-knowledge that are private and unknown to others, many aspects of the self are highly social in nature.

When individuals are asked to answer the question “Who am I?” several times, almost all of the answers given are enduring relationships to different social groups (“I am male”, “I am Latino”, “I am a college student”, “I have two sisters”, “I am a pre-med student”). Rarely do students include private non-observable information (“I am hungry”, “I am frustrated by this task”). A private investigator could ascertain all of these social facts about a person just by observing the person and without ever speaking to the individual. Some have suggested that we each engage in this kind of detective work in coming to understand ourselves (Bem, 1972). From this perspective, it might be more reasonable to say we often learn about ourselves through *extrospection* rather than introspection.

Mead (1934) and Cooley (1902) have both suggested that the self comes into existence as we come to see ourselves through the eyes of others such as our parents. The idea is that we recognize our caregivers as sentient beings before we recognize ourselves as such (Dixon, 1957; Pipp et al., 1987; Pipp-Siegel et al., 1995). As our caregivers treat us as sentient beings we make this inference that we too have a self. From this perspective, knowing and understanding others would thus become a prominent source of self-knowledge. If we come to know ourselves through the feedback we get from others, this gives others tremendous power to influence the structure and content of our selves. Moreover, this social genesis of the self would promote self-evaluations that rely heavily on the perceived evaluations of others. If Mead and others are correct, then it suggests that what we take to be our most intimate personal perception, our self-concept, is really the product of social forces, both in its inception and its upkeep. Although a great deal of anecdotal evidence supports Mead’s (1934) claims, there has been little compelling work on the developmental aspects of this process (Harter, 1999), and the work that has been done has depended entirely on self-reports of individuals and those who know them. In many ways, the links between self and social processes are actually easier to examine with fMRI because the common neural basis of self and social perception can be identified (Chaminade and Decety, 2002; Lieberman and Pfeifer, 2005; Ochsner et al., this issue).

Relevance to social cognitive neuroscience

These principles together help to organize and explain many aspects of social life and social cognition. Yet, how and why these principles function is rarely considered perhaps because behavioral data alone are ill-suited for answering these questions. Knowing that situations are powerful and that we are blind to them explains countless findings in social psychology, but why are we so sensitive to situations in our behavior but not in our perception? A common refrain from social psychology is that situations are invisible, but so are intentions, beliefs, desires, and dispositions. Why do we ‘see’ these invisible mental attributes everywhere and yet do not see invisible situational factors? What is it about our

neural wiring that makes this so? A social cognitive neuroscience approach can contribute to this line of questioning, and more importantly the results of this work could re-shape our understanding of these principles. Identifying which brain regions are especially (in)sensitive to situational factors could allow us to interrogate exactly what counts (or not) as a situation for these brain regions. Rather than defining situations and situational influence theoretically, it may be possible to define them in terms of the responses of the brain and the mental and behavioral consequences of those brain activations. It is possible that our lay definition of situations and situational influence will not line up with the brain’s response. Indeed, it would not be the first time that our naive theories about the social world turned out to be wrong. Similar kinds of investigations are possible for each of the social psychological principles which are taken as foundational.

Processes of social cognition

Above, I have described five of the broad principles derived from countless social psychology findings. When social psychologists want to understand the mechanisms by which these and other phenomena occur, they typically invoke one or more of the processing dynamics or structural features described below. All have become central areas of study in their own right within the study of social cognition and the ability to unobtrusively measure these processes with fMRI would be enormously useful.

Cognitive architecture

Social psychologists often invoke arguments from cognitive architecture to account for their findings (Nisbett and Ross, 1980). For instance, the correspondence bias in attribution (Gilbert and Malone, 1995; Jones and Harris, 1967) can be partially explained as a social instance of the anchoring and adjustment heuristic (Quattrone, 1982; Tversky and Kahneman, 1974). Many stereotyping effects can be thought of in terms of the representativeness heuristic and other basic categorization processes (Allport, 1954). Finally, some self-enhancing biases can be explained by the availability of diagnostic information in memory. For instance, both partners in a couple may believe they do more of the household chores than the other (Ross and Sicoly, 1979) because each can retrieve instances of their own contributions but may not have actually seen or paid as much attention to the partner doing chores. In each of these cases, the argument is similar to the one made regarding the relationship of visual illusions to visual processing in general: visual illusions point to features of the visual system that help to efficiently make sense of the visual world. These processes make some assumptions about the visual world and when these assumptions are violated, errors in the form of visual illusions occur. Similarly, one can argue that the cognitive architecture is highly adapted to allow for efficient perception of and participation in a social world and consequently makes certain assumptions that can produce systematically biased social judgments when those assumptions are violated.

Automaticity and control

Automaticity and control can be thought of as an extension of the cognitive architecture argument and yet the processing distinction has made so many contributions to the understanding

of social cognition that it is often appealed to on its own (Wegner and Bargh, 1998). Automatic processes are those that highly efficient, feel effortless, require no intention to operate, and/or occur outside of conscious awareness (Bargh, 1989; see Sander et al., this issue). Controlled processes are those that can be interrupted, feel effortful, require an intention to operate, and/or occur with conscious awareness. Automaticity and control are often referred to jointly in dual-process models (Chaiken and Trope, 1999) that suggest there are some kinds of information that are processed automatically and others that are only processed if one is motivated to think carefully and is not distracted by other thought processes at the same time. For instance, the correspondence bias (also known as the fundamental attribution error, Ross, 1977) involves inferring that a person has an enduring disposition (e.g., an anxious personality) to engage in behavior x after seeing a single instance of behavior x (e.g., a person behavior anxiously), even when the observer is aware that there is a strong situational explanation for why the target engaged in behavior x (e.g., the person was asked to talk to a stranger about her sexual fantasies). Gilbert (1989) has argued that people automatically infer from behaviors the corresponding personality traits. Gilbert also suggested that the initial judgment can be corrected in light of situational information; however, this correction process is a controlled process and thus only occurs when the observer is both motivated to be accurate and has controlled processing resources free. Thus, people may not engage this correctional process in most cases, resulting in a substantial correspondence bias. This model can be linked back to the anchor and adjustment heuristic mentioned under process 1. Combining the two explanations, it appears that dispositional attributions are automatically anchored onto the observed behavior and only adjusted to take account of situational factors if controlled processing resources are brought to bear on the attribution. Dual-process explanations have been generative in the study of persuasion (Petty and Cacioppo, 1986; Chaiken et al., 1989), stereotyping (Devine, 1989), attribution (Gilbert et al., 1988; Trope, 1986), and attitude and behavior consistency (Fazio and Towles-Schwen, 1999).

Motivated reasoning

Although some social phenomena can be explained by invoking known cognitive processes (see processes 1 and 2 above), others are not so easily reducible to these known processing categories. The fact that individuals consistently provide self-enhancing answers to a wide assortment of questions (“How smart are you?”, “How good of a friend are you?”) attests to this assertion. These effects even occur when it is not immediately obvious how someone should respond in order to be most self-enhancing (Dunning and Cohen, 1992), suggesting that the impact of motivation on reasoning may be outside of awareness. Additionally, these effects are often found to be stronger when an individual’s self-image is threatened (Beauregard and Dunning, 1998; Brown and Gallagher, 1992) and weaker when an individual’s self-image is affirmed (Steele, 1988; Tesser and Cornell, 1991), suggesting that a motivation to maintain one’s self-image is a driving force in these effects. These *positive illusions* (Greenwald, 1980; Taylor and Brown, 1988) and *motivated reasoning* processes (Kunda, 1990) seem to conflict with the general view that cognitive processes are designed to represent the world accurately; when it comes to ourselves, we appear instead to want to view ourselves positively, although not

always (Swann et al., 1989). The extent to which these motivated reasoning processes overlap with the effect of mood and affect on judgment (Ashby et al., 1999; Forgas, 1995; Gray, 1999) is still unclear but this may be a starting point for cognitive neuroscience investigations.

Accessibility, frames, and expectations

Social information is more readily perceived and interpreted in a particular way based on various mental constraints, expectations, and pre-activated representations as described in principle 3. In its simplest cognitive form, priming of a concept (“doctor”) will increase the activation of related concepts (“nurse”) such that the related concepts are more accessible to consciousness (Collins and Loftus, 1975). The study (Higgins et al., 1977) in which “adventurous” or “reckless” was primed is a classic example of using priming to promote a particular interpretation of ambiguous social information over another reasonable but unprimed interpretation. In real life, people are always in some situation or other that may serve to prime various representations and frames for understanding ongoing behavior—behavior that might well be construed as having another meaning if seen in a different context. Primes can be activated outside of awareness, but they need not be as long as individuals are unaware of how the primes are serving as primes (Bargh, 1992). Whole frameworks for understanding and motivating behavior can be primed by describing scenarios in slightly different terms (Higgins, 1997; Tversky and Kahneman, 1981). Finally, fully explicit expectations can serve as primes. If a person expects that a new acquaintance is extraverted, then he might inadvertently ask questions and behave in ways likely to elicit extraverted behavior from the target, regardless of whether this individual is actually extraverted or not (Rosenthal and Jacobson, 1968; Snyder and Swann, 1978). Thus, the explanatory mechanisms in this category cut across the automaticity–control distinction and the cognitive–motivation distinction and form a separate class of processing effects that are central to social cognition.

Relevance to social cognitive neuroscience

Each of these processes has been used fruitfully in the study of social psychological phenomena. One limitation of assessing the role of these processes in a phenomenon of interest is measurement. At least two different measurement issues arise that social cognitive neuroscience may be able to address. First, there is the basic issue of whether self-report and other cognitive assessment tools are accurately measuring the processes and constructs of interest. Neuroimaging is by no means a panacea for this issue; however, psychology has long rallied behind the converging measures approach. Additional measures, such as neural activation, can only improve our measurement capacities. Second, there is the related issue of whether self-report and cognitive measurements alter the very processes they are trying to measure. Does asking someone about their mood, self-esteem, frequency of self-focused attention, or sense of being socially rejected contaminate the individual’s mental activity such that the intended mental phenomena of study are irrevocably altered? This Heisenbergian dilemma may be addressable *in part* by using fMRI to assess the presence or absence of different mental processes. That is, if the neural signature of self-reflection can be pinpointed such that we can confidently infer that the presence of altered activation in a

particular neural circuit indicates the presence or absence of self-reflection, then it would not be necessary to interrupt subjects to probe their current level of self-reflection.

Of course, finding the neural signature of different social cognitive processes is no easy task. Typically, the same brain region is activated by a number of different tasks and therefore it is a painstaking empirical process to determine the precise computations performed by a brain region or neural circuit of multiple regions (Botvinick et al., 1999). Still, this could be one possible goal of social cognitive neuroscience research. Currently, cognitive neuroscientists can determine from brain activity if an individual is looking at faces or not (Hasson et al., 2004; Hasson et al., 2001) and whether an individual is familiar with the song they are listening to (Kraemer et al., 2005). It is possible that at some point in the future, social cognitive neuroscience could accurately determine when an individual is engaged in attribution processes, thinking about possible futures, and so on.

Puzzles of social cognition

The previous sections have reviewed the bedrock of social psychology in terms of its principles and processes. Perhaps the most productive area in social psychology has been in identifying complex and often paradoxical findings about everyday experience (Wegner and Gilbert, 2000). Recall that two of the major principles of social psychology focus on our blindness for important social psychological phenomena. The truth of these principles is thus partially responsible for the counterintuitive nature of social psychology (Griffin and Ross, 1991). This section on the enduring puzzles of social psychology may be of greatest interest to cognitive neuroscientists because it describes many of the surprising social psychological phenomena that have not yet been fully explained and may well benefit from cognitive neuroscience techniques.

The self

The self is as strange as psychological phenomena get. No one has ever seen their own self (Hume, 1739/1890) and yet many would argue that there is nothing we are more intimately connected to (James, 1890/1950). Still others would argue that the self is largely, if not completely, an illusion (Suzuki, 1964), while some would argue that although the self exists, it is a tool that operates for the society's benefit rather than in one's own best interest (Nietzsche, 1872/1999).

Social psychologists typically divide the study of the self up into several topics such as self-awareness, self-control, self-esteem, and self-knowledge (for a review, see Baumeister, 1998). Self-awareness and self-control are puzzling because there seem to be two selves involved in each of these acts. Consider the statements "I thought about what I wanted to eat" and "I made myself keep studying". In the first there appears to be both a knower and a known self (James, 1890/1950) and in the second there appears to be both a controller and a controlled self (Lakoff and Johnson, 1999). Are there really multiple selves that interact or are these descriptions nothing more than linguistic accidents of western culture?

Recall that Mead (1934) argued that the active components of self-awareness and self-control are like a mental simulation of what one's parents and other authority figures have provided as

expectations of desirable and appropriate behavior within a particular culture. Consistent with this idea, a large literature has shown that increasing self-focused attention increases the likelihood that an individual will follow societal or situational norms (Carver and Scheier, 1981; Wicklund, 1975). The fact that focusing on the self increases conformity rather than defiance against conformity as one affirms one's personal beliefs and values is a surprising finding and one that has never been entirely explained. Also, one must wonder, if self-awareness and self-control processes are really simulations of external admonishments, why do these processes feel so personal rather than feeling like social assessments? Social cognitive neuroscience is in a strong position to access the overlap in self and social cognition (Lieberman and Pfeifer, 2005; Ochsner et al., this issue). Finally, although engaging in self-focused attention and self-control may help guide behavior in socially acceptable ways, why do these processes also lead to poorer performances under some conditions (i.e., choking under pressure; Baumeister, 1984). Many expert performances occur when the self, in some sense, seems to disappear such that one feels at one with the performance (Csikszentmihalyi, 2000; Herrigel, 1953). And yet, when one gains this sort of performative expertise, these self-less performances can become a large part of one's self-concept (e.g., the person for whom skiing is a major part of their self-concept despite their sense of self perhaps being most absent when skiing).

Less esoteric but no less important is the phenomenon of self-esteem. Particularly in Western societies (Heine et al., 1999), people go to great lengths to boost their own self-esteem and that of their loved ones. But what exactly is self-esteem and why should aisles and aisles of self-help books be devoted to improving it? Self-esteem refers to one's assessment of one's own self-worth (suggesting one self that possesses some amount of worth and a second self that is aware of this worth). Why do we make this assessment and why should there be consequences of how much we think we are worthy? Is self-esteem a basic aspect of how the mind functions and regulates itself? Is self-esteem primarily an assessment of how valued we are by the groups we identify with (Leary and Baumeister, 2000), and if so why does it feel like a self-assessment rather than a social judgment? Why does high self-esteem sometimes lead to helpful behavior and at other times lead to aggressive behavior (Brown and Smart, 1991; Bushman and Baumeister, 1998).

Tying all of these paradoxical aspects of the self together, Baumeister and colleagues (Baumeister, 1990; Heatherton and Baumeister, 1991) have suggested that self-esteem, self-control, and self-awareness processes may collude to create a desire to escape from self-awareness which may perpetuate a cycle of increasing self-regulatory failures. In this phenomenon, unattainable self-standards (e.g., body image) lead to attempts at self-control (e.g., dieting) that are ultimately unsuccessful and experienced as highly distressing in the light of self-focused attention. In an attempt to escape one's own self-focused attention, one may engage in behaviors that diminish self-awareness (e.g., drinking, eating, and various risky behaviors). In the absence of self-awareness, social and self-imposed rules may be ignored leading to impulsive behavior (e.g., more eating) that may be regretted later when self-focused attention reemerges. The cycle thus perpetuates itself by increasing the discrepancy between the kind of self one thinks one should have and the kind of self one does have (Higgins, 1987), making future episodes of self-awareness increasingly painful. It is surprising that self-phenom-

ena which distinguish us from most, if not all, other species could have such potentially maladaptive consequences. This work portrays a very different view of self-processes than are typically examined in fMRI studies. Nevertheless, this behavioral work indicates these other aspects of the self are in need of investigation as well.

Attitudes

Attitudes were one of the earliest topics of study in social psychology (Thurstone, 1931) and they have been an early area of study in social cognitive neuroscience as well (see Cunningham et al., this issue). Attitudes were defined as mental dispositions towards particular kinds of behaviors and thus the assumption was that the best way to predict behavior prior to observing it was to assess a person's attitudes. One of the unexpected findings of the attitude literature is that attitudes and behavior are not all that highly correlated. Over the years, many factors have been found that mediate and moderate the relationship between attitudes and behavior such as the accessibility of the attitude (Bargh et al., 1992; Fazio et al., 1986) and having a specific intention linking the attitude to a planned behavior (Ajzen, 1991; Gollwitzer et al., 1990; see also Ouden et al., this issue), but it is still unclear why there is such a disconnect. Examining the neurocognitive bases of attitudes and behavior and how these interact may shed new light on this old problem.

One of the most famous findings in the attitude literature is that when one is made aware of the inconsistency between one's attitudes or between one attitude and a relevant behavior, there is a tendency to alter one's attitude to make it consistent with the other attitude or behavior. This process of cognitive dissonance reduction (Cooper and Fazio, 1984; Festinger, 1957; Harmon-Jones and Mills, 1999) looks like rationalization from the outside, but despite thousands of studies little is known about how it takes place from the inside. Does the person know they are rationalizing (Lieberman et al., 2001)? If they do know, how can they do it in good conscience, and if they do not know, how might such a process actually take place? An examination of constraint satisfaction processes in the brain may be a good place to begin looking (Shultz and Lepper, 1995; Simon et al., 2004).

Reflective social cognition

The ability to reflect on and think about one's experience and mental activity is by no means the province of social psychologists alone (Lieberman, *in press*). Cognitive psychologists and cognitive neuroscientists have examined reflective thought from many perspectives including working memory, executive processing, goal maintenance, conflict monitoring, source monitoring, and analogical thought (Anderson et al., 2003; Botvinick et al., 1999; Christoff et al., 2001; Mitchell et al., 2004b; Smith and Jonides, 1999). Nevertheless, social psychologists have uncovered some particularly quirky aspects of reflective thought that tend to highlight some of the less adaptive consequences of reflective thought.

In a landmark paper, Nisbett and Wilson (1977) provided substantial evidence regarding the systematic fallibility of introspective processes. In particular, they focused on the reasons people gave for recent behaviors and preferences. Again and again, they found that although people had an explanation they could readily give for their behavior, these explanations systematically

deviated from the actual factors influencing behavior. For instance, people who had been primed with words like 'moon' and 'ocean' were more likely to select 'Tide' as their favorite detergent, but they never referred to the priming episode when explaining why it was their favorite. Instead, they would say things like "it's what my mom uses" or "I like the box". In many subsequent studies, Wilson (2002) and Schooler et al. (1993) have observed that the introspective act itself can change preferences or other mental representations being scrutinized and can do so in undesirable ways. In one study (Wilson et al., 1993), when individuals were asked to introspect on their reasons for liking or disliking two art prints before choosing one to keep and take home, they were more likely to make a choice that they would later regret, compared to those who had not introspected. Why this effect occurs is still unclear. Given this effect occurs reliably, it is surprising that people have such a strong sense of the infallibility and utility of introspection.

A similarly strange phenomenon of reflective thought occurs when one tries not to think about something in particular such as a white bear (Wegner, 1994). In the moment, one will report fewer occurrences of white bear thoughts than others who have been told to think freely about white bears; however, those suppressing thoughts of white bears will have more occurrences of white bear thoughts when the task is over compared to those who were not suppressing thoughts of white bears (Wyland et al., 2003). This rebound effect has also been shown in the context of stereotyping such that attempts to control stereotyping in the moment can lead to greater stereotyping later on (Macrae et al., 1994). This work clearly has implications for our understanding of executive processes but has not been integrated into work on the neurocognitive mechanisms of executive control.

Finally, when individuals try to reflect on how different possible outcomes in the future would affect them emotionally, people are systematically inaccurate. A sizable literature on affective forecasting (Gilbert et al., 1998; Kahneman et al., 1999; Liberman and Trope, 1998) has shown that people typically overestimate how long positive events will make them feel good and how long negative events will make them feel bad. A number of accounts for these and related affective forecasting errors have focused on the content that is considered when making these forecasts, but little has been said about the possibility that thoughts about feelings and actual feelings might be processed by distinct neurocognitive systems, such that the system involved in forecasting might be incapable of representing certain aspects of actual experience (Kahneman et al., 1993).

Automatic social cognition

Automatic social processes have been of interest, in part, because they demonstrate the ways in which the environment can control our behavior and thoughts without us even knowing this effect is occurring (in support of principles 1 and 2). They have also been of interest because they elicit *schadenfreude* when they demonstrate what often looks like irrational or undesirable thinking in people who assume that they themselves are not susceptible to these sorts of errors (Pronin et al., 2004). Most recently, they have been of interest because some of the effects are just plain bizarre.

A standard finding in automatic social cognition research is that primed categories will bias the social perception of ambiguous behaviors, as in the study by Higgins et al. (1977) described earlier.

Over the last decade, social psychologists have also found that a primed category often leads to category consistent behavior in the person who has been primed. In the most famous of these studies, Bargh et al. (1996) primed some individuals with words related to the category ‘elderly’ and found that these individuals subsequently walked more slowly when surreptitiously timed compared to those who had not been so primed. The standard explanation for this effect is that priming the category of ‘elderly’ activates motor imagery for slow behavior. The assumption that semantic (i.e., the concept ‘slow’) and motor associates (i.e., the motor program for going slow) are similarly activated or even represented in a common code is difficult to test behaviorally, but specific neural predictions can be generated to test whether this account is plausible (Crockett et al., 2005).

In related work, others have found that individuals score better or worse on tests of general knowledge depending on whether they have been primed with ‘professor’ or ‘supermodel’, respectively (Dijksterhuis and Van Knippenberg, 1998). On the other side of the coin, stereotype threat (Steele and Aronson, 1995) refers to a phenomenon whereby priming one’s group membership (e.g., race or gender) can raise the specter of confirming a negative stereotype about one’s group and can lead to impaired performances. Identifying the mediators and moderators of stereotype threat has been difficult suggesting an opportunity for social cognitive neuroscience to make a substantive contribution here.

There is now a sizable catalogue of behaviors that can be facilitated or inhibited by the appropriate category primes (Wheeler and Petty, 2001) and the list seems to be growing monthly. There have also been a number of studies that have examined the automatic priming of goals (Chartrand and Bargh, 1996) and motivations (Shah, 2003). In fact, people who have been unwittingly primed with a goal that they fail to achieve experience negative moods that they cannot explain (Chartrand and Jefferis, 2003). Despite all of these automatic behavior, goal, and motivation findings, virtually nothing is known about the representations and mechanisms involved in these processes.

Another perplexing aspect of automatic social cognition is that, in addition to being the putative basis for many of the errors and biases found in social judgment, it is also thought to be the source of many of our most accurate judgments. Ambady and Rosenthal (1993) observed that undergraduates could accurately identify a teacher’s teaching ability after observing just six seconds of video of the teacher with no sound and no knowledge of the material being taught. Testifying to the automaticity of these judgments, Ambady (1999) found that placing participants under cognitive load did not diminish the accuracy of their judgments. Behavioral science cannot currently give a good answer as to why some social automaticities are highly accurate while others are systematically inaccurate. Social cognitive neuroscience may be better positioned to investigate this (Lieberman, 2000) and tease apart the different forms of automatic social cognition.

Social motives

People have motivations to be liked and accepted by others (Baumeister and Leary, 1995; Williams, 2001) and to be understood by others (Hardin and Higgins, 1996; Swann et al., 1989). To be disliked or unknown by others, both conferring an invisibility status on oneself, is a painful phenomena (Eisenberger et al., 2003; MacDonald and Leary, 2005) that may contribute

substantially to our willingness to conform to group norms that we know to be wrong (Asch, 1956; Williams et al., 2000). Although it can be argued that being known and liked by others is adaptive, the aspects of the human mind that promote these motives and how these motives are connected or separate from other mental processes are largely unknown. It has been argued that the need for social connection may overlap neurally with physical pain processes as well as cognitive processes that tend to set controlled processes in motion for some and self-focused processes for others (Eisenberger and Lieberman, 2004; Eisenberger et al., in press). The desire to be understood by others, one of our most profound motives, seems quite different and presumably relies on additional processes.

A large number of studies suggest that when we satisfy these social motives, as indicated by feeling social connected and having large social networks to rely upon, there are physical health benefits ranging from a diminished likelihood of contracting a cold (Cohen et al., 1997) to lower morbidity rates for the general population (House et al., 1988). At least some of these benefits do not seem to be mediated by supportive others promoting health behaviors leaving one to wonder how this apparent ‘action at a distance’ is occurring? How does the brain and body convert social support ‘out there’ into better health outcomes ‘in here’?

Conclusions

Each of the sections of this article is no doubt woefully incomplete and some of the phenomena that I have chosen to highlight would not make the list for many others (for more complete assessments of social cognition, see Barone et al., 1997; Fiske and Taylor, 1991; Kunda, 1999; Moskowitz, 2005). The goal of this article was neither to be exhaustive nor to necessarily present ideas in order of importance. Instead, the point of this article was to introduce cognitive neuroscientists to some of the rich complexity of social cognition as it has been studied by social psychologists for decades.

Much of social psychology is fundamentally paradoxical, at least to the western mind. We tend to believe that we are the captains of our destiny, and yet, time and time again, social psychology has shown that situational factors exert strong pressures on our behavior and often does so without our knowledge. The implications of these and other findings for social cognitive neuroscience are twofold. First, although social psychologists have established these various principles, understanding why humans are guided by these principles and when these principles apply remain largely unknown. If social cognitive neuroscience can help to answer these questions it would be a major contribution to our understanding of social cognition. Second, the principles of social psychology apply not only to the subjects in our investigations, but to us, the researchers as well. In the absence of understanding these principles, we are likely to generate social cognitive hypotheses that are unnecessarily naïve. If we are as blind to the power of situational forces and our own ability to construct social perceptions that do not feel constructed, we will be unable to generate experimental paradigms that take these factors into account. Ultimately, a successful social cognitive neuroscience should thoroughly integrate the methods of social cognition and cognitive neuroscience, and also rely in equal parts on the conceptual lexica of these two parent disciplines as well.

Acknowledgments

The author was supported by a grant from the National Institute of Mental Health (R21MH66709) during the writing of this article. The author wishes to thank Dan Gilbert, Roy Baumeister, Kevin Ochsner, and Naomi Eisenberger for helpful discussion of ideas relevant to this article and Tristen Inagaki for technical assistance with the article.

References

- Ajzen, I., 1991. The theory of planned behavior. *Organizational Behavior and Human Decision Processes. Special Theories of Cognitive Self-Regulation* 50, 179–211.
- Allport, G.W., 1954. The nature of prejudice. *Psychology* 54, 193–202.
- Ambady, N., 1999. The perils of pondering: distraction, deliberation, and thin slice judgments. Unpublished manuscript, Harvard University.
- Ambady, N., Rosenthal, R., 1993. Half a minute: predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *J. Pers. Soc. Psychol.* 64, 431–441.
- Amodio, D.M., Harmon-Jones, E., Devine, P.G., 2003. Individual differences in the activation and control of affective race bias as assessed by startle eyeblink response and self-report. *J. Pers. Soc. Psychol.* 84, 738–753.
- Anderson, J.R., Qin, Y., Sohn, M.H., Stenger, V.A., Carter, C.S., 2003. An information-processing model of the BOLD response in symbol manipulation tasks. *Psychon. Bull. Rev.* 10, 241–261.
- Aron, A., Fisher, H.E., Mashek, D.J., Strong, G., Li, H.F., Brown, L.L., 2005. Reward, motivation and emotion systems associated with early-stage intense romantic love. *J. Neurophysiol* (electronic publication ahead of print).
- Asch, S.E., 1956. Studies of independence and conformity: a minority of one against a unanimous majority. *Psychol. Monogr.* 70, 416.
- Ashby, F.G., Isen, A.M., Turken, A.U., 1999. A neuropsychological theory of positive affect and its influence on cognition. *Psychol. Rev.* 106, 529–550.
- Bargh, J.A., 1989. Conditional automaticity: varieties of automatic influence in social perception and cognition. In: Uleman, J.S., Bargh, J.A. (Eds.), *Unintended Thought*. Guilford, New York, NY, USA, pp. 3–51.
- Bargh, J.A., 1992. Does subliminality matter to social psychology? Awareness of the stimulus versus awareness of its influence. In: Bornstein, R.F., Pittman, T.S. (Eds.), *Perception Without Awareness: Cognitive, Clinical, and Social Perspectives*. Guilford Press, New York, NY, USA, pp. 236–255.
- Bargh, J.A., Chaiken, S., Govender, R., Pratto, F., 1992. The generality of the automatic attitude activation effect. *J. Pers. Soc. Psychol.* 62, 893–912.
- Bargh, J.A., Chen, M., Burrows, L., 1996. Automaticity of social behavior: direct effects of trait construct and stereotype activation on action. *J. Pers. Soc. Psychol.* 71, 230–244.
- Baron-Cohen, S., Ring, H., Moriarty, J., Shmitz, P., Costa, D., Ell, P., 1994. Recognition of mental state terms: a clinical study of autism, and a functional imaging study of normal adults. *Br. J. Psychiatry* 165, 640–649.
- Barone, D.F., Maddux, J.E., Snyder, C.R., 1997. *Social Cognitive Psychology: History and Current Domains*. Plenum, New York, NY.
- Baumeister, R.E., 1984. Choking under pressure: self-consciousness and paradoxical effects of incentives on skillful performance. *J. Pers. Soc. Psychol.* 46, 610–620.
- Baumeister, R.F., 1990. Anxiety and deconstruction: on escaping the self. In: Olson, J.M., Zanna, M.P. (Eds.), *Self-Inference Processes: The Ontario Symposium, Ont. Symp. Pers. Soc. Psychol.*, vol. 6. Lawrence Erlbaum Associates, Inc., Hillsdale, NJ, pp. 259–291.
- Baumeister, R.F., 1998. The self. In: Gilbert, D.T., Fiske, S.T., Lindzey, G. (Eds.), *The Handbook of Social Psychology*. McGraw-Hill, Boston, MA, pp. 680–740.
- Baumeister, R.F., Leary, M.R., 1995. The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychol. Bull.* 117, 497–529.
- Beauregard, K.S., Dunning, D., 1998. Turning up the contrast: self-enhancement motives prompt egocentric contrast effects in social judgments. *J. Pers. Soc. Psychol.* 74, 606–621.
- Beauregard, M., Levesque, J., Bourgouin, P., 2001. Neural correlates of conscious self-regulation of emotion. *J. Neurosci.* 21, 6993–7000.
- Bem, D.J., 1972. Self-perception theory. In: Berkowitz, L. (Ed.), *Advances in Experimental Social Psychology*. Academic Press, New York, NY, USA, pp. 1–62.
- Blackwood, N.J., Bentall, R.P., Ffytche, D.H., Simmons, A., Murray, R.M., 2003. Self-responsibility and the self-serving bias: an fMRI investigation of causal attributions. *NeuroImage* 20, 1076–1085.
- Blascovich, J., Mendes, W.B., 2000. Challenge and threat appraisals: the role of affective cues. In: Forgas, J.P. (Ed.), *Feeling and Thinking: The Role of Affect in Social Cognition. Studies in Emotion and Social Interaction, Second Series*. Cambridge Univ. Press, New York, NY, USA, pp. 59–82.
- Botvinick, M., Nystrom, L.E., Fissell, K., Carter, C.S., Cohen, J.D., 1999. Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature* 402, 179–181.
- Brown, J.D., Gallagher, K.M., 1992. Coming to terms with failure: private self-enhancement and public self-effacement. *J. Exp. Soc. Psychol.* 28, 3–22.
- Brown, J.D., Smart, S.A., 1991. The self and social conduct: linking self-representations to prosocial behavior. *J. Pers. Soc. Psychol.* 60, 368–375.
- Bushman, B.J., Baumeister, R.F., 1998. Threatened egotism, narcissism, self-esteem, and direct and displaced aggression: does self-love or self-hate lead to violence? *J. Pers. Soc. Psychol.* 75, 219–229.
- Cacioppo, J.T., Crites, S.L., Gardner, W.L., 1996. Attitudes to the right: evaluative processing is associated with lateralized late positive event-related brain potentials. *Pers. Soc. Psychol. Bull.* 22, 1205–1219.
- Cacioppo, J.T., Bernston, G.G., Adolphs, R., Carter, C.S., Davidson, R.J., McClintock, M.K., McEwen, B.S., Meaney, M.J., Schacter, D.L., Sternberg, E.M., Suomi, S.S., Taylor, S.E., 2002. *Foundations in Social Neuroscience*. MIT Press, Cambridge, MA.
- Carr, L., Iacoboni, M., Dubeau, M.C., Mazziotta, J.C., Lenzi, G.L., 2003. Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. *Proc. Natl. Acad. Sci. U. S. A.* 100, 5497–5502.
- Carver, C.S., Scheier, M.F., 1981. *Attention and Self-Regulation: A Control Theory Approach to Human Behavior*. Springer-Verlag, New York, NY.
- Chaiken, S., Trope, Y., 1999. *Dual-Process Theories in Social Psychology*. Guilford Press, New York, NY, USA.
- Chaiken, S., Liberman, A., Eagly, A.H., 1989. Heuristic and systematic processing within and beyond the persuasion context. In: Uleman, J.S., Bargh, J.A. (Eds.), *Unintended Thought*. Guilford Press, New York, NY, USA, pp. 212–252.
- Chaminade, T., Decety, J., 2002. Leader or follower? Involvement of the inferior parietal lobule in agency. *NeuroReport* 13, 1975–1978.
- Chartrand, T.L., Bargh, J.A., 1996. Automatic activation of impression formation and memorization goals: nonconscious goal priming reproduces effects of explicit task instructions. *J. Pers. Soc. Psychol.* 71, 464–478.
- Chartrand, T.L., Jefferis, V.E., 2003. Consequences of automatic goal pursuit and the case of nonconscious mimicry. In: Forgas, J.P., Williams, K.R., von Hippel, W. (Eds.), *Social Judgments: Implicit and Explicit Processes*. Cambridge Univ. Press, New York, NY, USA, pp. 290–305.
- Christoff, K., Prabhakaran, V., Dorfman, J., Zhao, Z., Kroger, J.K., Holyoak, K.J., Gabrieli, J.D., 2001. Rostrolateral prefrontal cortex involvement in relational integration during reasoning. *NeuroImage* 14, 1136–1149.

- Cohen, S., Doyle, W.J., Skoner, D.P., Rabin, B.S., Gwaltney Jr., J.M., 1997. Social ties and susceptibility to the common cold. *J. Am. Med. Assoc.* 277, 1940–1944.
- Collins, A.M., Loftus, E.E., 1975. A spreading-activation theory of semantic processing. *Psychol. Rev.* 82, 407–428.
- Cooley, C.H., 1902. *Human Nature and Social Order*. Scribner's, New York, NY.
- Cooper, J., Fazio, R.H., 1984. A new look at dissonance theory. In: Berkowitz, L. (Ed.), *Advances in Experimental Social Psychology*. Academic Press, New York, NY, pp. 229–266.
- Crockett, M.J., Eisenberger, N.I., Lieberman, M.D., 2005. Stereotype activation or behavior regulation? An fMRI study of automatic behavior effects. Poster Presented at the 6th Annual Meeting of the Society for Personality and Social Psychology. New Orleans, LA (January).
- Csikszentmihalyi, M., 2000. *Beyond Boredom and Anxiety: Experiencing Flow in Work and Play*. Jossey-Bass, New York, NY.
- Cunningham, W.A., Johnson, M.K., Gatenby, J.C., Gore, J.C., Banaji, M.R., 2003. Neural components of social evaluation. *J. Pers. Soc. Psychol.* 85, 639–649.
- De Quervain, D.J., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., Fehr, E., 2004. The neural basis of altruistic punishment. *Science* 305, 1254–1258.
- Devine, P.G., 1989. Stereotypes and prejudice: their automatic and controlled components. *J. Pers. Soc. Psychol.* 56, 680–690.
- Dijksterhuis, A., Van Knippenberg, A., 1998. The relation between perception and behavior, or how to win a game of trivial pursuit. *J. Pers. Soc. Psychol.* 74, 865–877.
- Dixon, J.C., 1957. Development of self-recognition. *J. Genet. Psychol.* 91, 251–256.
- Dunning, D., Cohen, G.L., 1992. Egocentric definitions of traits and abilities in social judgment. *J. Pers. Soc. Psychol.* 63, 341–355.
- Eisenberger, N.I., Lieberman, M.D., 2004. Why it hurts to be left out: the neurocognitive overlap between physical and social pain. *Trends Cogn. Sci.* 8, 294–300.
- Eisenberger, N.I., Lieberman, M.D., Williams, K.D., 2003. Does rejection hurt? An fMRI study of social exclusion. *Science* 302, 290–292.
- Eisenberger, N.I., Lieberman, M.D., Satpute, A.B., 2005. Personality from a controlled processing perspective: an fMRI study of neuroticism, extraversion, and self-consciousness. *Cogn. Affect. Behav. Neurosci.* 5, 169–181.
- Fazio, R.H., Towles-Schwen, T., 1999. The MODE model of attitude-behavior processes. In: Chaiken, S., Trope, Y. (Eds.), *Dual-Process Theories in Social Psychology*. Guilford Press, New York, NY, USA, pp. 117–138.
- Fazio, R.H., Sanbonmatsu, D.M., Powell, M.C., Kardes, F.R., 1986. On the automatic activation of attitudes. *J. Pers. Soc. Psychol.* 50, 229–238.
- Festinger, L., 1957. *A Theory of Cognitive Dissonance*. Row, Peterson, Evanston, IL.
- Fiske, S.T., Taylor, S.E., 1991. *Social Cognition*, 2nd ed. McGraw-Hill Book Company, New York.
- Forgas, J.P., 1995. Mood and judgment: the affect infusion model (AIM). *Psychol. Bull.* 117, 39–66.
- Frith, C., Frith, U., 1999. Interacting minds—A biological basis. *Science* 286, 1692–1695.
- Gilbert, D.T., 1989. Thinking lightly about others. Automatic components of the social inference process. In: Uleman, J.S., Bargh, J.A. (Eds.), *Unintended Thought*. Guilford, New York, NY, USA, pp. 189–211.
- Gilbert, D.T., Malone, P.S., 1995. The correspondence bias. *Psychol. Bull.* 117, 21–38.
- Gilbert, D.T., Pelham, B.W., Krull, D.S., 1988. On cognitive busyness: when person perceivers meet persons perceived. *J. Pers. Soc. Psychol.* 54, 733–740.
- Gilbert, D.T., Pinel, E.C., Wilson, T.D., Blumberg, S.J., Wheatley, T.P., 1998. Immune neglect: a source of durability bias in affective forecasting. *J. Pers. Soc. Psychol.* 75, 617–638.
- Gollwitzer, P.M., Heckhausen, H., Steller, B., 1990. Deliberative and implemental mind-sets: cognitive tuning toward congruous thoughts and information. *J. Pers. Soc. Psychol.* 59, 1119–1127.
- Gray, J.R., 1999. A bias toward short-term thinking in threat-related negative emotional states. *Pers. Soc. Psychol. Bull.* 25, 65–75.
- Greenwald, A.G., 1980. The totalitarian ego. *Am. Psychol.* 35, 603–618.
- Griffin, D.W., Ross, L., 1991. Subject construal, social inference, and human misunderstanding. In: Zanna, M.P. (Ed.), *Adv. Exp. Soc. Psychol.* Academic Press, San Diego, pp. 319–359.
- Gusnard, D.A., Akbudak, E., Shulman, G.L., Raichle, M.E., 2001. Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proc. Natl. Acad. Sci.* 98, 4259–4264.
- Hardin, C.D., Higgins, E.T., 1996. Social reality: how social verification makes the subjective objective. In: Sorrentino, R.M., Higgins, E.T. (Eds.), *The Interpersonal Context. Handbook of Motivation and Cognition, Handb. Motiv. Cogn.*, vol. 3. Guilford Press, New York, NY, USA, pp. 28–84.
- Harmon-Jones, E., Mills, J., 1999. Cognitive dissonance: progress on a pivotal theory in social psychology. In: Harmon-Jones, E., Mills, J. (Eds.), *Science Conference Series. American Psychological Association*, Washington, DC, USA, p. 411.
- Hart, A.J., Whalen, P.J., Shin, L.M., McInerney, S.C., Fischer, H., Rauch, S.L., 2000. Differential response in the human amygdala to racial outgroup vs. ingroup face stimuli. *NeuroReport* 11, 2351–2355.
- Harter, S., 1999. *The Construction of the Self: A Developmental Perspective*. Guilford Press, New York.
- Hasson, U., Hendler, T., Bashat, D.B., Malach, R., 2001. Vase or face? A neural correlates of shape-selective grouping processes in the human brain. *J. Cogn. Neurosci.* 13, 744–753.
- Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., Malach, R., 2004. Synchronization of cortical activity during natural vision. *Science* 303, 1634–1640.
- Hastorf, A., Cantril, H., 1954. They saw a game: a case study. *J. Abnorm. Soc. Psychol.* 49, 129–134.
- Heatherton, T.F., Baumeister, R.F., 1991. Binge eating as escape from self-awareness. *Psychol. Bull.* 110, 86–108.
- Heine, S.H., Lehman, D.R., Markus, H.R., Kitayama, S., 1999. Is there a universal need for positive self-regard? *Psychol. Rev.* 106, 766–794.
- Herrigel, E., 1953. *Zen in the Art of Archery*. Vintage, New York.
- Higgins, E.T., 1987. Self-discrepancy: a theory relating self and affect. *Psychol. Rev.* 94, 319–340.
- Higgins, E.T., 1996. Knowledge activation: accessibility, applicability, and salience. In: Higgins, E.T., Kruglanski, A.W. (Eds.), *Social Psychology: Handbook of Basic Principles*. Guilford, New York, NY, USA, pp. 133–168.
- Higgins, E.T., 1997. Beyond pleasure and pain. *Am. Psychol.* 52, 1280–1300.
- Higgins, E.T., Rholes, W.S., Jones, C.R., 1977. Category accessibility and impression formation. *J. Exp. Soc. Psychol.* 13, 141–154.
- House, J.S., Landis, K.R., Umberson, D., 1988. Social relationships and health. *Science* 241, 540–545.
- Hume, D., 1739/1890. *A Treatise on Human Nature*. Langmans, Green, and Co., Cleveland, OH.
- Iacoboni, M., Lieberman, M.D., Knowlton, B.J., Molnar-Szakacs, I., Moritz, M., Throop, C.J., Fiske, A.P., 2004. Watching social interactions produces dorsomedial prefrontal and medial parietal BOLD fMRI signal increases compared to a resting baseline. *NeuroImage* 21, 1167–1173.
- James, W., 1890/1950. *The Principles of Psychology*. Dover, New York, NY.
- Jones, E.E., Harris, V.A., 1967. The attribution of attitudes. *J. Exp. Soc. Psychol.* 3, 1–24.
- Kahneman, D., Fredrickson, B.L., Schreiber, C.A., Redelmeier, D.A., 1993. When more pain is preferred to less: adding a better ending. *Psychol. Sci.* 4, 401–405.
- Kahneman, D., Diener, E., Schwarz, N. (Eds.), *Well-Being: The Foundations of Hedonic Psychology*. Russell Sage Foundation, New York, NY.

- Kant, I., 1781/1965. *Critique of Pure Reason*. St. Martin's Press, New York.
- Keenan, J.P., Nelson, A., O'Connor, M., Pascual-Leone, A., 2001. Self-recognition and the right hemisphere. *Nature* 409, 305.
- Kosfeld, M., Heinrichs, M., Zak, P.J., Fischbacher, U., Fehr, E., 2005. Oxytocin increases trust in humans. *Nature* 435, 673–676.
- Kelley, W.M., Macrae, C.N., Wyland, C.L., Caglar, S., Inati, S., Heatherton, T.F., 2002. Finding the self? An event-related fMRI study. *J. Cogn. Neurosci.* 14, 785–794.
- Klein, S.B., Loftus, J., Kihlstrom, J.F., 1996. Self-knowledge of an amnesic patient: toward a neuropsychology of personality and social psychology. *J. Exp. Psychol. Gen.* 125, 250–260.
- Kraemer, D.J., Macrae, C.N., Green, A.E., Kelley, W.M., 2005. Musical imagery: sound of silence activates auditory cortex. *Nature* 434, 158.
- 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. *Behav. Brain Sci.* 27, 313–322.
- Kunda, Z., 1990. The case for motivated reasoning. *Psychol. Bull.* 108, 480–498.
- Kunda, Z., 1999. *Social Cognition: Making Sense of People*. MIT Press, Cambridge, MA.
- Lakoff, G., Johnson, M., 1999. *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*. Basic Books, New York, NY.
- Latane, B., Darley, J.M., 1970. *The Unresponsive Bystander: Why Doesn't He Help?* Prentice-Hall, Englewood Cliffs, NJ.
- Leary, M.R., Baumeister, R.F., 2000. The nature and function of self-esteem: sociometer theory. In: Zanna, M.P. (Ed.), *Advances in Experimental Social Psychology*. Academic Press, San Diego, CA, pp. 1–62.
- Lieberman, M.D., 2000. Intuition: a social cognitive neuroscience approach. *Psychol. Bull.* 126, 109–137.
- Lieberman, M.D., Pfeifer, J.H., 2005. The self and social perception: three kinds of questions in social cognitive neuroscience. In: Easton, A., Emery, N. (Eds.), *Cognitive Neuroscience of Emotional and Social Behavior*. Psychology Press, Philadelphia, PA, pp. 195–235.
- Lieberman, N., Trope, Y., 1998. The role of feasibility and desirability considerations in near and distant future decisions: a test of temporal construal theory. *J. Pers. Soc. Psychol.* 75, 5–18.
- Lieberman, M.D., in press. The X- and C-systems: the neural basis of reflexive and reflective social cognition. In: Harmon-Jones E., Winkelman, P. (Eds.), *Fundamentals of Social Neuroscience*. Guilford, New York, NY.
- Lieberman, M.D., Ochsner, K.N., Gilbert, D.T., Schacter, D.L., 2001. Do amnesics exhibit cognitive dissonance reduction? The role of explicit memory and attention in attitude change. *Psychol. Sci.* 12, 135–140.
- Lieberman, M.D., Gaunt, R., Gilbert, D.T., Trope, Y., 2002. Reflection and reflexion: a social cognitive neuroscience approach to attributional inference. *Adv. Exp. Soc. Psychol.* 34, 199–249.
- Lieberman, M.D., Jarcho, J.M., Berman, S., Naliboff, B., Suyenobu, B.Y., Mandelkern, M., Mayer, E., 2004. The neural correlates of placebo effects: a disruption account. *NeuroImage* 22, 447–455.
- Lieberman, M.D., Hariri, A., Jarcho, J.J., Eisenberger, N.I., Bookheimer, S.Y., 2005. An fMRI investigation of race-related amygdala activity in African-American and Caucasian-American individuals. *Nat. Neurosci.* 8, 720–722.
- MacDonald, G., Leary, M.R., 2005. Why does social exclusion hurt? The relationship between social and physical pain. *Psychol. Bull.* 131, 202–223.
- Macrae, C.N., Milne, A.B., Bodenhausen, G.V., 1994. Stereotypes as energy-saving devices: a peek inside the cognitive toolbox. *J. Pers. Soc. Psychol.* 66, 37–47.
- Markus, H.R., 1977. Self-schemata and processing information about the self. *J. Pers. Soc. Psychol.* 35, 63–78.
- Mead, G.H., 1934. *Mind, Self, and Society*. University of Chicago Press, Chicago, IL.
- Milgram, S., 1963. Behavioral study of obedience. *J. Abnorm. Soc. Psychol.* 67, 371–378.
- Miller, D.T., Ratner, R.K., 1998. The disparity between the actual and assumed power of self-interest. *J. Pers. Soc. Psychol.* 74, 53–62.
- Mitchell, J.P., Macrae, C.N., Banaji, M.R., 2004a. Encoding-specific effects of social cognition on the neural correlates of subsequent memory. *J. Neurosci.* 24, 4912–4917.
- Mitchell, K.J., Johnson, M.K., Raye, C.L., Greene, E.J., 2004b. Prefrontal cortex activity associated with source monitoring in a working memory task. *J. Cogn. Neurosci.* 16, 921–934.
- Moskowitz, G.B., 2005. *Social Cognition: Understanding Self and Others*. Guilford, New York, NY.
- Moscowitz, G.B., Roman, R.J., 1992. Spontaneous trait inferences as self-generated primes: implications for conscious social judgment. *J. Pers. Soc. Psychol.* 62, 728–738.
- Neisser, U., 1976. *Cognition and Reality: Principles and Implications of Cognitive Psychology*. Freeman, San Francisco.
- Nietzsche, F.W., 1872/1999. *The Birth of Tragedy*. Cambridge Univ. Press, Cambridge, England.
- Nisbett, R., Ross, L., 1980. *Human Inference: Strategies and Shortcomings of Social Judgment*. Prentice Hall, Englewood Cliffs, NJ.
- Nisbett, R.E., Wilson, T.D., 1977. Telling more than we can know: verbal reports on mental processes. *Psychol. Rev.* 84, 231–259.
- Ochsner, K.N., Lieberman, M.D., 2001. The emergence of social cognitive neuroscience. *Am. Psychol.* 56, 717–734.
- Ochsner, K.N., Bunge, S.A., Gross, J.J., Gabrieli, J.D.E., 2002. Rethinking feelings: an fMRI study of the cognitive regulation of emotion. *J. Cogn. Neurosci.* 14, 1215–1229.
- Petrovic, P., Kalso, E., Petersson, K.M., Ingvar, M., 2002. Placebo and opioid analgesia—Imaging a shared neuronal network. *Science* 295, 1737–1740.
- Petty, R.E., Cacioppo, J.T., 1986. The elaboration likelihood model of persuasion. *Adv. Exp. Soc. Psychol.* 19, 123–205.
- Phelps, E.A., O'Connor, K.J., Cunningham, W.A., Funayama, E.S., Gatenby, J.C., Gore, J.C., Banaji, M., 2000. Performance on indirect measures of race evaluation predicts amygdala activation. *J. Cogn. Neurosci.* 12, 729–738.
- Pipp, S., Fischer, K.W., Jennings, S., 1987. Acquisition of self- and mother knowledge in infancy. *Dev. Psychol.* 23, 86–96.
- Pipp-Siegel, S., Easterbrooks, M.A., Brown, S.R., Harmon, R.J., 1995. The relation between infants' self mother knowledge and three attachment categories. *Infant Ment. Health J.* 16, 221–232.
- Pronin, E., Gilovich, T., Ross, L., 2004. Objectivity in the eye of the beholder: divergent perceptions of bias in self versus others. *Psychol. Rev.* 111, 781–799.
- Quattrone, G.A., 1982. Overattribution and unit formation: when behavior engulfs the person. *J. Pers. Soc. Psychol.* 42, 593–607.
- Richeson, J.A., Baird, A.A., Gordon, H.L., Heatherton, T.F., Wyland, C.L., Trawalter, S., Shelton, J.N., 2003. An fMRI investigation of the impact of interracial contact on executive function. *Nat. Neurosci.* 6, 1323–1328.
- Rilling, J.K., Sanfey, A.G., Aronson, J.A., Nystrom, L.E., Cohen, J.D., 2004. The neural correlates of theory of mind within interpersonal interactions. *NeuroImage* 22, 1694–1703.
- Robinson, R.J., Keltner, D., Ward, A., Ross, L., 1995. Actual versus assumed differences in construal: “Naive realism” in intergroup perception and conflict. *J. Pers. Soc. Psychol.* 68, 404–417.
- Rosenthal, R., Jacobson, L., 1968. *Pygmalion in the Classroom: Teacher Expectation and Student Intellectual Development*. Holt, Rinehart and Winston, New York, NY.
- Ross, L., 1977. The intuitive psychologist and his shortcomings. In: Berkowitz, L. (Ed.), *Advances in Experimental Social Psychology*. Academic Press, New York, NY, USA.
- Ross, M., Sicoly, E., 1979. Egocentric biases in availability and attribution. *J. Pers. Soc. Psychol.* 37, 322–336.
- Sanfey, A.G., Rilling, J.K., Aronson, J.A., Nystrom, L.E., Cohen, J.D., 2003. The neural basis of economic decision-making in the Ultimatum Game. *Science* 300, 1755–1758.
- Schooler, J.W., Ohlsson, S., Brooks, K., 1993. Thoughts beyond words:

- when language overshadows insight. *J. Exp. Psychol. Gen.* 122, 166–183.
- Shah, J., 2003. The motivational looking glass: how significant others implicitly affect goal appraisals. *J. Pers. Soc. Psychol.* 85, 424–439.
- Shultz, T.R., Lepper, M.R., 1995. Cognitive dissonance reduction as constraint satisfaction. *Psychol. Rev.* 103, 219–240.
- Simon, D., Krawczyk, D.C., Holyoak, K.J., 2004. Construction of preferences by constraint satisfaction. *Psychol. Sci.* 15, 331–336.
- Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R.J., Frith, C.D., 2004. Empathy for pain involves the affective but not sensory components of pain. *Science* 303, 1157–1162.
- Smith, E., Jonides, J., 1999. Storage and executive processes in the frontal lobes. *Science* 283, 657–660.
- Snyder, M., Swann, W.B., 1978. Hypothesis-testing processes in social interaction. *J. Pers. Soc. Psychol.* 36, 1202–1212.
- Steele, C.M., 1988. The psychology of self-affirmation: sustaining the integrity of the self. In: Berkowitz, L. (Ed.), *Advances in Experimental Social Psychology*. Academic Press, San Diego, CA, pp. 261–302.
- Steele, C.M., Aronson, J., 1995. Stereotype threat and the intellectual test performance of African Americans. *J. Pers. Soc. Psychol.* 69, 797–811.
- Suzuki, D.T., 1964. *An Introduction to Zen Buddhism*. Grove Press, New York.
- Swann, W.B. Jr., Pelham, B.W., Krull, D.S., 1989. Agreeable fancy or disagreeable truth? Reconciling self-enhancement and self-verification. *J. Pers. Soc. Psychol.* 57, 782–791.
- Taylor, S.E., Brown, J.D., 1988. Illusion and well-being: a social psychological perspective on mental health. *Psychol. Bull.* 103, 193–210.
- Taylor, S.E., Klein, L.C., Lewis, B.P., Gruenewald, T.L., Gurung, R.A., Updegraff, J.A., 2000. Biobehavioral responses to stress in females: tend-and-befriend, not fight-or flight. *Psychol. Rev.* 107, 411–429.
- Tesser, A., 1988. Toward a self-evaluation maintenance model of social behavior. In: Berkowitz Jr., L. (Ed.), *Advances in Experimental Social Psychology*. Academic Press, New York, NY, USA, pp. 181–227.
- Tesser, A., Cornell, D.P., 1991. On the confluence of self processes. *J. Exp. Soc. Psychol.* 64, 85–96.
- Thurstone, L.L., 1931. The measurement of attitudes. *J. Abnorm. Soc. Psychol.* 26, 249–269.
- Trope, Y., 1986. Identification and inferential processes in dispositional attribution. *Psychol. Rev.* 93, 239–257.
- Tversky, A., Kahneman, D., 1974. Judgment under uncertainty: heuristics and biases. *Science* 185, 1124–1131.
- Tversky, A., Kahneman, D., 1981. The framing of decisions and the psychology of choice. *Science* 211, 453–458.
- Wager, T.D., Rilling, J.K., Smith, E.E., Sokolik, A., Casey, K.L., Davidson, R.J., Kosslyn, S.M., Rose, R.M., Cohen, J.D., 2004. Placebo-induced changes in fMRI in the anticipation and experience of pain. *Science* 303, 1162–1167.
- Wegner, D.M., 1994. Ironic processes of mental control. *Psychol. Rev.* 101, 34–52.
- Wegner, D.M., Bargh, J.A., 1998. Control and automaticity in social life. In: Gilbert, D.T., Fiske, S.T., Lindzey, G. (Eds.), *The Handbook of Social Psychology*, 4th ed. Oxford Univ. Press, New York, NY, USA, pp. 446–496.
- Wegner, D.M., Gilbert, D.T., 2000. Social psychology—The science of human experience. In: Bless, H., Forgas, J.P. (Eds.), *The Message Within: The Role of Subjective Experience in Social Cognition and Behavior*. Psychology Press, Philadelphia, PA.
- Wheeler, M.E., Fiske, S.T., 2005. Controlling racial prejudice: social-cognitive goals affect amygdala and stereotype activation. *Psychol. Sci.* 16, 56–63.
- Wheeler, S.C., Petty, R.E., 2001. The effects of stereotype activation on behavior: a review of possible mechanisms. *Psychol. Bull.* 127, 797–826.
- Wicklund, R.A., 1975. Objective self-awareness. In: Berkowitz, L. (Ed.), *Advances in Experimental Social Psychology*. Academic Press, New York, NY.
- Williams, K.D., 2001. *Ostracism: The Power of Silence. Emotions and Social Behavior*. Guilford Press, New York, NY, USA, p. 282.
- Williams, K.D., Cheung, C.K.T., Choi, W., 2000. Cyerostracism: effects of being ignored over the internet. *J. Pers. Soc. Psychol.* 79, 748–762.
- Wilson, T.D., 2002. *Strangers to Ourselves. Discovering the Adaptive Unconscious*. Harvard Univ. Press, Cambridge, MA.
- 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. *Pers. Soc. Psychol. Bull.* 19, 331–339.
- Wood, J.N., Romero, S.G., Knutson, K.M., Grafman, J., 2005. Representation of attitudinal knowledge: role of prefrontal cortex, amygdala and parahippocampal gyrus. *Neuropsychologia* 43, 249–259.
- Wyland, C.L., Kelley, W.M., Macrae, C.N., Gordon, H.L., Heatherton, T.F., 2003. Neural correlates of thought suppression. *Neuropsychologia* 41, 1863–1867.