

7 The self and social perception: Three kinds of questions in social cognitive neuroscience

*Matthew D. Lieberman
and Jennifer H. Pfeifer*

“A man has as many social selves as there are individuals who recognize him.”
William James (1890), *Principles of Psychology* (Vol. I, p. 294)

“The world is twofold for man in accordance with his twofold nature.”
Martin Buber (1937), *I and Thou* (p. 82)

A great deal of research in social psychology is motivated by one of two broad goals: (1) to understand the mental processes involved in how people make sense of the social world; (2) to understand how self-processes are shaped by the social world. In other words, social psychologists are deeply interested in the interplay between intrapersonal and interpersonal processes. In the final analysis, most social psychologists agree that neither can be understood in isolation. Though many naively take for granted a sovereign self that is inaccessible to others and independent of their influence, the opening quotation from William James, as well as the theoretical and empirical history of social psychology, suggests that the development and maintenance of the self is shaped by one's situational context. Alternatively, many believe that perceiving the social world is a relatively objective process akin to, albeit more complicated than, perceiving the nonsocial world. Endless evidence suggests that this, too, is a naive view, an issue addressed in the philosophy of Martin Buber. Perceiving the social world is a subjective process shaped by an individual's current motivation, emotion, and cognition, as well as his or her more long-standing traits such as personalities, self-schemas, and chronically accessible constructs. An even more extreme position was taken by the philosopher Nietzsche, who suggested that social perception is nothing but the projection of our own idiosyncratic representations onto the world in his claim, “Whoever thought that he had understood something of me had merely construed something out of me, after his own image” (Nietzsche, 1908/1969, p. 261).

Now that the methods of cognitive neuroscience are being applied to social psychological questions, it is to be expected that early research efforts turn to social perception and self-processes as appropriate starting points. In this

chapter, we review the work that has been done with human subject populations in the growing area of social cognitive neuroscience (Adolphs, 1999; Klein & Kihlstrom, 1998; Ochsner & Lieberman, 2001; for a review of the broader field of social neuroscience, see Cacioppo et al., 2002). This chapter is organized around three types of questions that can be addressed with a social cognitive neuroscience approach: (1) where is the process in the brain? (2) what kinds of computations or cognitions are involved in the process? (3) how does the process of interest change as a function of contextual factors, other ongoing mental activity, and learning history? To foreshadow one of our main conclusions, although it is natural to first study these topics separately, social cognitive neuroscientists must ultimately reconcile these initial accounts with the ways in which self-processes and social perception are bound up with one another. The separate study of each is merely a convenience carved out by our research methods, rather than an indication that either can be understood alone.

Three questions

We begin by laying out three questions that can be asked with a social cognitive neuroscience approach and what is gained from each type of question. The questions are presented in the order they are typically asked. Though questions 2 and 3 may be of greater intrinsic significance than question 1, typically, they cannot be empirically addressed until question 1 has been answered.

Question 1: Where is the process in the brain?

The best cognitive neuroscience research always informs psychological theories, and social cognitive neuroscience should aspire to no less. Yet, discovering the neural correlates of a social psychological process appears thoroughly irrelevant to theory at first blush. Indeed, if research began and ended with brain-mapping studies of this sort, psychologists would have little reason for enthusiasm. Straightforward brain-mapping studies serve an important purpose when considered from a broader context. Psychologists frequently generate novel hypotheses based upon the anecdotal experience of everyday life. Icheiser (1943) and Heider (1958) noticed that people tend to attribute observed behaviour to the enduring dispositions of the actor without sufficiently considering the impact of the situation on the actor's behaviour. These insights have guided over three decades of empirical work on attributional inference (Gilbert & Malone, 1995; Jones & Davis, 1965). In a similar manner, brain mapping provides us with anecdotes previously unavailable to the senses, as each study connects a group of brain regions to a particular macrolevel process or experience. The activation of region *x* while performing task *y* is not sufficient to identify region *x* with the performance of task *y*; however, each study provides a clue. For instance, region *x* may

make some minor contribution that is a necessary but nonobvious component of performing task y . Intuitions about the function(s) of region x will be built up as it is activated by a number of superficially different tasks. The payoff of brain-mapping studies is that we can begin to hypothesize about the computations common to the range of tasks and ultimately conduct targeted tests of this hypothesis. However, this requires us to know which candidate regions are potentially involved in a process of interest before we can begin more precise hypothesis testing. Determining these candidate regions is largely the result of brain-mapping studies.

Question 2: What kind of computations/cognitions are involved in the process?

From the extensive brain mapping done over the past decade, cognitive neuroscientists now have a catalogue of associations between cognition and brain regions. To name just a few, we know of specific regions of the brain associated with various forms of sensory processing, episodic memory, semantic memory, working memory, implicit learning, face processing, automatic affect, and conflict monitoring. Though this index of neurocognitive correlations is constantly being refined and updated, it allows us to test hypotheses regarding the different mental processes involved in performing any task. One might hypothesize that some task involves subprocesses $P1$, $P2$, and $P3$ which are known to be associated with neural regions $N1$, $N2$, and $N3$, respectively. While it is often difficult to assess the simultaneous contributions of multiple cognitive processes to a task in behavioural studies, neuroimaging methods are well suited for it. Looking for the presence or absence of five neural regions in the performance of a task is often no more difficult than looking for three or one. The presence or absence of $N1$, $N2$, and $N3$ thus allows us to infer the involvement of psychological processes $P1$, $P2$, and $P3$. Social psychologists often want to examine the extent to which their process of interest relies on automatic or controlled processing, “hot” affect or “cold” cognition, visual or linguistic processing, and episodic or semantic memory. Neuroimaging provides data bearing on each of these processes without having to interrupt a participant’s mental activity to assess dependent variables.

Question 3: How does the process of interest change as a function of contextual factors, other ongoing mental activity, and learning history?

The benefit of cognitive neuroscience cataloguing correlations between brain regions and cognitive operations is that we can now address question 2 and assess the common computations underlying different social psychological processes. The downside to this cataloguing process is that this enterprise leads to the compartmentalization of mental processing, such that it is easy to believe that these brain regions function independently of one another.

Compartmentalization of mental processes has an empirical legacy extending back to Wundt and the introspectionists (Kohler, 1947). For example, Ebbinghaus (1885/1964) examined memory processes by studying nonsense syllables. He believed that by avoiding content with affective and semantic meaning one could study the true roots of episodic memory. However, research from the past two decades has shown that affective processing is not an independent process added on top of, and therefore linearly separable from, episodic memory. Quite the contrary, episodic memory encoding in the medial temporal lobe is critically modulated by the affective significance of the stimulus by way of beta noradrenergic outputs from the amygdala (Cahill & McGaugh, 1998). Social psychological theory has a long tradition, inspired by Gestalt psychology, of emphasizing the interdependence of different social cognitive processes. For example, part of what makes the self an object of study for social psychology is its multifaceted nature, with different facets presenting themselves, depending on the social context. There is little reason to believe that the neural underpinnings of the self would be any less flexible in the face of social pressures than is the self's behavioural and cognitive expression.

One of the advantages of current fMRI data-analysis techniques is our ability to examine how and to what extent the interplay between different brain regions changes as a function of context, motivation, priming, and so on. Recall the above example involving neural processes *N1*, *N2*, and *N3*. It might be the case that these areas are always activated in a particular process, and yet the coordination of activity between the three regions may change as a function of current motivational and contextual conditions. The extent to which the regions operate together in lock step is termed *functional connectivity*, the analysis of which will prove critical to the social cognitive neuroscience approach. To elaborate, rather than merely showing whether two brain regions are activated by a task, functional connectivity analyses assess the extent to which the regions respond in concert to the task. The assessment of functional connectivity as motivation, cognition, or context changes allows us to determine how different aspects of social cognition affect one another.

The remainder of this review is divided into sections on the self and on social perception, including the following specific topics: self-knowledge, self-awareness, self-control, recognizing others, attribution, and stereotyping. We discuss how each of these topics has been investigated according to the three questions described above. Naturally, given the short history of neuroscientific research on these topics, there is a good deal more work in addressing the first question (brain mapping) than the second (function) and third (mediator/moderator processes), as each tends to build on the answers to the previous questions.

The self

Until the advent of neuroimaging in recent years, most insights into the neural bases of the self came from neuropsychological cases, the most famous of which is the case of Phineas Gage, which has been described at length elsewhere (Damasio, 1994; Macmillan, 2000). After a tamping iron was literally shot through Gage's skull, his cognitive, perceptual, and motor skills were left surprisingly unimpaired. However, Gage underwent a radical personality shift and, from all accounts, was no longer the same person. The damage also appeared to alter his social judgement, and Gage went on to make a string of poor social decisions that left him alone and penniless at the end of his life. The accident caused damage to his ventromedial prefrontal cortex, suggesting that this region of the brain might be critical for both social perception and self-processes.

Before proceeding with our discussion of a social cognitive neuroscience approach to the self, we must define and outline how modern social psychologists organize their study of the self. Rather than examine the self in its monolithic entirety, three aspects of self-processing are typically focused upon: self-knowledge, self-awareness, and self-control (Baumeister, 1998). Self-knowledge includes both the capacity to recognize oneself and the storehouse of information that one keeps regarding one's own personality, preferences, and autobiographical history. Self-awareness refers to the capacity to reflect upon and identify with one's own ongoing experience and actions. Self-control is the regulatory capacity to strategically overcome one's own impulses and habits.

Self-knowledge and self-recognition

Neural correlates

Because self-recognition depends on identification with an external aspect of the self that is intersubjectively available, it is perhaps the most empirically tractable of self-processes and thus has already been the subject of a number of cognitive neuroscience studies in recent years. Two neuropsychological investigations suggest that self-recognition is lateralized to one hemisphere but draw opposite conclusions as to which hemisphere is dominant. While undergoing a Wada test, which anaesthetizes the two hemispheres one at a time, patients were shown a picture that was created by morphing their own face with that of a famous person (Keenan, Nelson, O'Connor, & Pascual-Leone, 2001; see also Keenan, McCutcheon, Freund, Gallup, Sanders, & Pascual-Leone, 1999). Participants were then asked to choose which of the two unmorphed faces they had been shown. When only the right hemisphere was functioning, patients uniformly believed the face shown was their own, whereas the face was construed more often as the other famous individual when only the left hemisphere was functioning. Alternatively, a

split-brain patient demonstrated a distinct left hemisphere advantage for self-recognition (Turk, Heatherton, Kelley, Funnell, Gazzaniga, & Macrae, 2002). Two fMRI studies revealed a complex pattern of left and right hemisphere structures participating in self-recognition (Kircher et al., 2000). Both studies found right hemisphere activations in the hippocampus, anterior cingulate, precuneus, and cerebellum and left hemisphere activations in the lateral prefrontal cortex and inferior parietal lobe.

Self-knowledge, in terms of knowing one's own personality, preferences, and autobiography, is a more complicated topic of study, as these are conspicuously subjective features of the self. Everyone can agree that a photograph of a person taken yesterday is really of that person, but not everyone agrees on the personality traits of an individual—often, friends and loved ones can see a person quite differently than does the individual him- or herself. Despite this obstacle, self-knowledge has been the most frequently studied aspect of the self in cognitive neuroscience studies. Klein and colleagues (Klein, Cosmides, Costabile, & Mei, 2002; Klein, Loftus, & Kihlstrom, 1996) have examined patients with compromised declarative memory systems to determine whether these are critical to self-knowledge. Amazingly, they have found complete sparing of self-knowledge despite massive impairments to explicit memory for other phenomena. In one case (Klein, Loftus, & Kihlstrom, 1996), a woman with head trauma reported a given personality while unable to recall most of her life history, and still reported the same personality when her memory later returned. This would appear to fly in the face of Kant's theory of indirect self-knowledge, subsequently embodied in Bem's (1972) self-perception theory. Bem argued that people come to know themselves by observing their own behaviour, in the same way that people learn about the preferences and personality of others, rather than through introspection. Bem's theory suggests that knowing one's self is dependent on remembering one's own behaviour. One could argue that in the case of the head trauma patient, the woman had been able to reflect on her behaviour for many years before the injury and may have developed semantic self-knowledge that would have been unaffected by her temporary amnesia (McClelland, McNaughton, & O'Reilly, 1995). However, Klein et al. (2002) have shown that an autistic patient with compromised semantic memory still maintained accurate self-knowledge, suggesting that not all self-knowledge is developed explicitly.

Additionally, a neuropsychological study of 12 permanently amnesic patients (Lieberman, Ochsner, Gilbert, & Schacter, 2001) demonstrated that new behaviours led to updated self-knowledge even though the patients never had explicit memory of their behaviour. Social psychological research has shown that engaging in freely chosen counter-attitudinal behaviour is sufficient to change the attitudes and preferences of the actor, a process referred to as cognitive dissonance reduction. Noticing the conflict between one's behaviour and beliefs is thought to cause cognitive dissonance, a state of psychological distress that typically leads people to change their beliefs to fit

their behaviour. In other words, people engage in rationalization to keep up the appearance of self-consistency, both in their own eyes and those of others around them. Lieberman et al., however, found that amnesic patients incapable of forming new memories demonstrated as much attitude change after engaging in freely chosen counter-attitudinal behaviour as did healthy controls. Just seconds after performing it, these patients could not remember engaging in the counter-attitudinal behaviour—yet, their self was modified to take account of this behaviour. This suggests that if self-perception is truly a source of self-knowledge it cannot require elaborate, time-consuming, conscious consideration of one's behaviours; it must instead be a more automatic consequence of observing the self, at least some of the time (Shultz & Lepper, 1995).

Contrary to the neuropsychological findings, imaging studies of self-knowledge processes typically find activations associated with episodic retrieval and explicit thought. Three positron emission tomography (PET) studies (Craig et al., 1999; Fink, Markowitsch, Reinkemeier, Bruckbauer, Kessler, & Heiss, 1996; Kjaer, Nowak, & Lou, 2002), along with four fMRI studies (Johnson, Baxter, Wilder, Pipe, Heiserman, & Prigatano, 2002; Kelley, Macrae, Wyland, Caglar, Inati, & Heatheron, 2002; Kircher et al., 2000; Lieberman, Jarcho, & Satpute, 2003), have each used similar paradigms in which participants had to judge whether words were self-descriptive. As seen in Table 7.1, six of the seven studies found activation in the region of the precuneus and posterior cingulate. The medial prefrontal cortex (BA 9/10), extending into the ventromedial prefrontal cortex (BA 11), was active in four of the seven studies. Right inferotemporal cortex (BA 21/38) and inferior parietal (BA 40) activations were reported in three of the six studies. The ventrolateral prefrontal cortex (BA 44/45/47), basal ganglia, and insula were each reported in two of the studies.

Neurocognitive processes

Knowing which regions of the brain support self-knowledge can only be of theoretical value to the extent that the functions of these regions can be specified. The most frequently activated regions (precuneus, medial prefrontal cortex, and lateral temporal cortex) are far from completely understood, but do seem to perform functions consistent with explicit knowledge processes. The medial prefrontal cortex is just now beginning to be understood. As described later in the section on attribution, the medial prefrontal cortex is involved in making explicit attributions about the mental states of others (Castelli et al., 2002; Gallagher & Frith, 2003) and thus may be similarly involved in making attributions about the self (Bem, 1972). Mounting evidence suggests that the precuneus in the parietal lobe plays a role in perspective taking, differentiating self and other, and retrieval of episodic memories. The precuneus will be discussed at greater length below in the section on self-awareness. Finally, the right inferotemporal

Table 7.1 Neural correlates of self-knowledge retrieval from seven studies

<i>Brain region</i>	<i>Brodmann area(s)</i>	<i>Side</i>	<i>Kircher et al. (2000)</i>	<i>Johnson et al. (2002)</i>	<i>Craik et al. (1999)</i>	<i>Kelley et al. (2002)</i>	<i>Fink et al. (1996)</i>	<i>Lieberman et al. in press</i>	<i>Kjaer et al. (2002)</i>	<i>Total</i>
Precuneus and posterior cingulate	7/31	Both	•	•	•	•	•	•	•	6/7
Medial and ventromedial prefrontal	9/10/11	Both		•	•			•	•	4/7
Inferotemporal	21/38	Right		•			•			3/7
Inferior parietal	40	Both	•						•	3/7
Ventrolateral prefrontal	44/45/47	Right			•			•		2/7
Basal ganglia		Left	•					•		2/7
Insula		Both	•				•			2/7

cortex and temporal pole may play a role in the storage of declarative self-knowledge.

The putative functions of the areas most active when making self-knowledge judgements present a picture that is consistent with seventeenth-century philosopher John Locke's claim about the role of explicit thought and memory processes in the maintenance of self-knowledge. Locke (1690/1975) claimed that "as far as this consciousness can be extended backwards to any past action or thought, so far reaches the identity of that person" (Bk. II, Ch. 27–29). In other words, without being able to recall what one has done in the past, how can one answer questions about what kind of person one is in general? In this model of self-knowledge, we judge our own trait characteristics (such as "generous") by explicitly retrieving and considering exemplars of past behaviour in which we exhibited behaviour consistent or inconsistent with that trait. The activations in the precuneus, the medial prefrontal cortex, and the temporal lobe during self-knowledge judgements fit this model of *evidence-based* self-knowledge processing.

A reconciliation between Klein's work and the neuroimaging findings is possible if it is assumed that (1) there are multiple self-knowledge systems and (2) individuals access evidence-based self-knowledge when required to make self-judgements, even though they could use other sources of self-knowledge if they suffered from impairments to the neural systems involved in evidence-based self-knowledge.

Contextual factors

To this point, there is good evidence to suggest that evidence-based self-knowledge is not the only self-knowledge system that operates in humans. Other than the presumed existence of at least one other self-knowledge system, virtually nothing is known about the nature of this second system. Some have claimed that its representations are prototypes or schemas, but cognitive psychologists continue to debate whether these representations are truly distinct from exemplar models (Allen & Brooks, 1991; Nosofsky & Palmieri, 1997).

Lieberman, Jarcho, and Satpute (2003) recently conducted an fMRI study to identify the neural correlates of *intuition-based* self-knowledge, a possible second self-knowledge system. From previous behavioural work (Klein, Loftus, Trafton, & Fuhrman, 1992; Markus, 1977), it was assumed that the amount of experience in a domain would moderate a shift in reliance from evidence-based to intuition-based self-knowledge. More specifically, Klein et al. (1992) found support for evidence-based self-knowledge use when participants made self-judgements in a domain with which they had little experience, but no indication of evidence-based self-knowledge use in domains of high experience. Lieberman et al. (2003) hypothesized that the neural correlates of the two kinds of self-knowledge would be differentially activated as a function of domain experience.

Eleven professional actors and 11 college athletes were scanned while judging the self-descriptiveness of acting and athletic trait words. It was hypothesized that neural correlates of intuition-based self-knowledge would be more activated when participants made high-experience domain judgements, and the neural correlates of evidence-based self-knowledge would be more activated when participants made low-experience domain judgements. From previous research (Lieberman, 2003; Lieberman, Gaunt, Gilbert, & Trope, 2002), it was hypothesized that evidence-based self-processes would be associated with the medial temporal lobe, precuneus, and lateral prefrontal cortex, whereas intuition-based self-knowledge would be associated with the ventromedial prefrontal cortex, basal ganglia, and amygdala. Initial group analyses of self-judgements in a domain of high experience produced activations in the ventromedial prefrontal cortex, nucleus accumbens in the basal ganglia, amygdala, and posterior parietal cortex largely consistent with the hypotheses. No regions were more active in this group analysis of low-experience self-judgements. By the magnitude of each participant's reaction time advantage for high-experience domain judgements over low-experience domain judgements, participants were divided into two post hoc groups (*schematics* and *nonschematics*) for further analysis. Schematics, when analysed alone, produced high-experience self-judgement activations in the ventromedial prefrontal cortex, nucleus accumbens, amygdala, and precuneus along with a relative *deactivation* in hippocampus. Alternatively, nonschematics produced high-experience self-judgement activations in the hippocampus, precuneus, and lateral prefrontal cortex along with a relative *deactivation* in the ventromedial prefrontal cortex. In a between-groups analysis, schematics produced more activation in the ventromedial prefrontal cortex, whereas nonschematics produced more activation in the hippocampus (Figure 7.1).

This study demonstrates both that self-knowledge is instantiated in multiple networks in the brain that support different components of self-knowledge, and that these networks have qualitatively different operating characteristics. Future studies will be able to examine how various contextual factors, including the presence of other people, goals, emotion, motivation, and cognitive resources, determine when each of these networks is dominant in guiding our use of self-knowledge.

Self-awareness and introspection

William James (1890) described the self as composed of the "I" and the "me" (see also Mead, 1934; Sartre, 1936), and Allport (1955) similarly described the self dichotomously, as the knower and the known. The "me" or known self corresponds roughly to self-knowledge, as discussed in the previous section. It refers to relatively static information that one possesses regarding one's own personality, preferences, abilities, and the like. The composite "me" from all of these representations is relatively stable, such that people typically believe their "me" of last year and next year will be similar to their me of

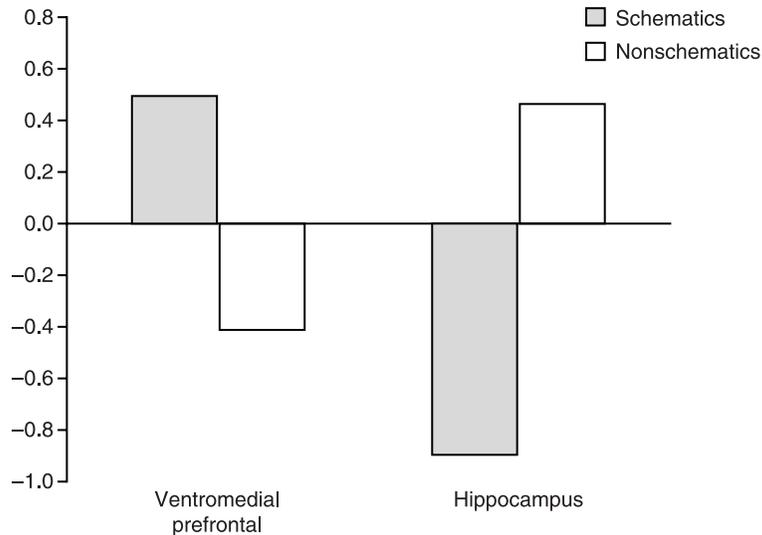


Figure 7.1 Relative activation in the ventromedial prefrontal cortex and hippocampus for schematics and nonschematics.

today (cf. Dweck, 1999). The knower or “I” refers to the feelings of spontaneous and active experience of the self in the moment. Inspired by cybernetic models of human cognition (Carver & Scheier, 1981; Miller, Galanter, & Pribram, 1960; Weiner, 1948), social psychologists have often divided the study of the “I” into two parts: self-awareness and self-control.

The simplest cybernetic models explain the activity of self-regulatory mechanisms such as thermostats. Thermostats computationally consist of two operations: *test* and *operate* (Miller et al., 1960). Thus, they test whether the ambient temperature is below some preset threshold; if so, the thermostat operates on the environment by turning on the heat until a new test comes back indicating the desirable temperature has been reached. Self-awareness, introspection, or self-focused attention seem to be the phenomenological analogues of the test function, whereas exerting self-control seems to be the phenomenological analogue to the operate function. In this section, we assess what is known about the neurocognitive underpinnings of self-awareness.

Neural correlates

There are few if any better examples of punctate self-awareness than the experience of pain. People do not feel pain and then wonder whose pain it is. Instead, the onset of pain leads to the immediate conclusion that the pain is one’s own and that something must be done to rectify the situation as quickly as possible. The subjective response to pain and the subsequent behaviour initiated by this experience nicely demonstrate the test–operate aspects of

cybernetic models. Though less dramatic, a plethora of other conflicts also produce momentary self-awareness. For instance, struggling with a Stroop task leads to awareness of one's own compulsion to say the colour named by the word rather than the colour of the ink that the word is written in as an answer. Realizing too late that one has just put the wrong answer on a test leads to self-awareness, and staring at a Necker cube may lead to self-focused attention on one's inability to hold the percept constant. Each of these conflict-ridden events prompting self-awareness activates the anterior cingulate cortex on the medial surface of the brain (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Eisenberger, Satpute, & Lieberman, 2003; Rainville et al., 1997). Most anterior cingulate activations appear to simultaneously produce an experience of self-awareness and set in motion cognitive, behavioural, and physiological responses to remove the conflict if possible. Furthermore, Lane et al. (1997) have found that the magnitude of anterior cingulate reactivity to emotional stimuli presentation is highly correlated with self-reported levels of chronic self-awareness.

Whereas the anterior cingulate has been implicated in bottom-up induced self-awareness, the posterior parietal cortex may be involved in the kind of indirect self-awareness proposed by Kant, Mead, and Bem. Patients with damage to the posterior parietal cortex, occurring frequently as a result of a stroke, have difficulty appreciating their own cognitive and behavioural impairments, a condition referred to as *anosognosia*. For example, Ramachandran (1995) described patients with right parietal damage who denied the paralysis of their left arm. When asked if her left arm was paralysed, one patient indicated that her left arm was fine—just as strong as the right one. When asked to point with her left hand, instead of admitting that she cannot move it, the patient claimed she decided she did not *want* to move her hand. Patients with anosognosia are known for generating contorted rationalizations for their deficits rather than acknowledging the deficit itself. Given that there are certain treatments that can temporarily remove the anosognosia, it appears that the patients are not using the rationalization as an explicit defence mechanism. Rather, it appears that damage to the parietal lobe truly prevents patients from experiencing and identifying with the deficit at all.

Neurocognitive processes

Self-awareness of a conflict and self-control exerted to overcome the conflict are naturally co-occurring events, as one typically leads to the other. Not surprisingly, it has been difficult to tease apart the neural contributors to these two processes. Most studies of tasks requiring self-control have shown activation of both the anterior cingulate and prefrontal cortex. Cohen and colleagues (Carter et al., 2000) have performed a series of elegant neuroimaging studies that suggest the anterior cingulate performs computations critical to self-awareness of conflict, without being directly involved in the exertion

of self-control. They used a modified Stroop-like task that allowed analysis of brain region activity covarying with the detection of the need for self-control separate from brain region activity covarying with the actual exertion of control (Carter et al., 1998, 2000). They found a double dissociation such that the anterior cingulate responded during the detection of the need for control, but not the exertion of control, whereas the lateral prefrontal cortex responded during the exertion of control, but not the detection of the need for control.

The neurocognitive properties of the posterior parietal cortex (including the wedge of cortex consisting of precuneus, intraparietal sulcus, and inferior parietal lobe) are just beginning to emerge with respect to self-awareness. The posterior parietal cortex has typically been assigned two functions, one as a nonexecutive maintenance component of working memory (Cabeza & Nyberg, 2000) and one as a spatial processing area (Goldman-Rakic, 1988). A variety of studies that have heretofore been unconnected suggest that the posterior parietal cortex may be the location where nonsymbolic, parallel, distributed representations are translated into symbolic, serial, local representations. It is these symbolic representations which we are aware of as distinct “figures” emerging from the tacit, unattended-to background that makes up the stream of consciousness. This claim is made plausible given the strong innervation of the posterior parietal cortex from ventral temporal and occipital areas, where visual information is analysed into objects and categories without necessarily reaching consciousness. A number of studies have shown that for stimuli presented near the subliminal/supraliminal border, the posterior parietal activation predicts when the stimuli will be perceived as a meaningful entity (Beck, Rees, Frith, & Lavie, 2001; Kjaer, Nowak, Kjaer, Lou, & Lou, 2001; Portas, Strange, Friston, Dolan, & Frith, 2000).

The capacity for true symbolic processing may be critical for perspective-taking capacity. Whereas distributed associative representations compute similarity functions, symbols can be used in propositions that explicitly represent asymmetric relations between entities (Holyoak & Hummel, 2000; Lieberman et al., 2002; Sloman, 1996). Perspective taking depends on representing asymmetric relationships, as the perspective taker must be able to differentiate his perspective from that of his target. A handful of recent studies have found that perspective-taking efforts are associated with posterior parietal activations (Chaminade & Decety, 2002; Farrer & Frith, 2002; Farrow et al., 2001; Ruby & Decety, 2001). Moreover, the precuneus in the posterior parietal cortex is consistently activated in episodic retrieval effort (Cabeza & Nyberg, 2000; Fletcher et al., 1995) and must be intact for successful episodic retrieval in rodents (Izquierdo et al., 1997). Combining these findings, some have suggested that during episodic retrieval one is attempting to take one’s own perspective from a previous point in time, indicating a form of self-perspective taking. This prospect highlights an interesting paradox about the self. For the current active self to retrieve information about its own past, it must suspend its current perspective. One might think that the

currently active self would, introspectively, be automatically in touch with the past self, but this turns out not to be the case. Indeed, this complicated process of taking a former perspective is not always successful, especially in the case of memory of prior beliefs and feelings (Bem & McConnell, 1970; Fredrickson, 2000; Hawkins & Hastie, 1990; Kahneman, Fredrickson, Schreiber, & Redelmeier, 1993; Ross, 1989).

Prominent social psychological theories have suggested that just as we must engage in perspective taking to retrieve our past selves, we must also engage in perspective taking to be aware of and understand our current self. Mead (1934) suggested that we come to know ourselves by seeing ourselves through the eyes of others around us. It is hard to directly size up our own abilities (“How can I figure out how smart I am on my own?”), but humans have an uncanny knack for automatically casting judgement on others (see section on attribution below). Even though people might not be able to know themselves directly, they can count on those around them to have already made these judgements. Learning about one’s self can then be as simple as learning what others have already learned about one’s self. To the extent that this process characterizes the acquisition of self-knowledge, it suggests: (1) self-knowledge generation is a fundamentally social process in contrast to the naive “introspective access” theory that most people hold; (2) the neural correlates of perspective taking should be central to the self-awareness that helps us to generate and retrieve self-knowledge.

Contextual factors

Recall that in the earlier section on self-knowledge, the precuneus was the most commonly activated area associated with self-judgements. We can now hypothesize that this activation reflects some combination of self-perspective taking to retrieve information about the self of the past and taking the views of important others on the self. To the extent that these perspective-taking processes are entangled, it may become increasingly difficult to dissociate self-processing from social processing. Future studies will benefit from independently manipulating these two processes, through priming and direct instruction, in order to get a clearer picture of the neurocognitive contributions of each to self-processing.

Self-control

Neural correlates

A variety of neuropsychological disorders suggest that self-control, the paradoxical ability to exert oneself in the service of overcoming one’s own impulses (Lakoff & Johnson, 1999; Turner, 1976), is associated with the lateral prefrontal cortex and the basal ganglia. Disorders of self-control including Parkinson’s disease, Huntington’s disease, obsessive-compulsive

disorder, and Tourette's syndrome each highlight the delicate neural balance that is necessary for healthy individuals to control their own thoughts and behaviour. All of these disorders have been found to affect the prefrontal cortex, basal ganglia, and anterior cingulate. We have already reviewed the test-operate cybernetic model of self-regulation, and the anterior cingulate's role is clearly relegated to the *test* function. The prefrontal cortex and basal ganglia are thus involved in the remaining operate function. The basal ganglia have been hypothesized to be involved in more automatic forms of self-control built up from habit (Lieberman, 2000), leaving the prefrontal cortex as the major brain region associated with effortful self-control.

Neurocognitive processes

There are at least three neurocognitive functions of the lateral prefrontal cortex that contribute to its role in effortful self-control. The lateral prefrontal cortex has consistently been associated with working memory and language, which together produce the ability to compose novel propositions from a series of symbolic representations and then hold these in mind. This planning capacity, which differentiates us from all other animals excepting a few primates who may have immature prototypes of these processes (Deacon, 1997), allows humans to imagine a variety of futures which have not yet come to pass and consider the consequences of following their impulses. Although we can choose alternative courses of action from those suggested by our impulses, this flexibility clearly comes at a cost. Exciting new data suggest that these prefrontal capacities have many properties of a muscle. Whereas self-control capacities may strengthen with use over time, they become "tired" and depleted in the short term with overuse such that rapid-fire bouts of self-control tend to produce increasingly poor results (Vohs & Heatherton, 2000).

There are three different computational mechanisms by which the propositional activation of the prefrontal cortex can lead to self-control. First, the products of prefrontal activity—in the form of conclusions and behavioural intentions—can activate the motor system directly, taking control away from more automatic neural processes. The downside to this mechanism is evident in the example of learning to touch-type. The individual must form an intention to type each letter, one at a time, a slow and fragile process that interferes with the ability to think about the content of what one would like to communicate. This direct guidance of behaviour is incredibly flexible and useful in novel circumstances, but it operates serially on behaviours and requires effort and attention. Moreover, the conclusions formed from the judgements and logic of the prefrontal cortex are far from perfect, often making the problem at hand worse (Wegner & Petty, 1995; Wegner, 1994). Indeed, decision making research suggests that explicit attempts at decision making tend to systematically leave out critical pieces of information even when they are available (Gilbert & Wilson, 2000; Tversky & Kahneman, 1972).

A second way the prefrontal cortex can exert self-control is by facilitating the activation of weaker processes and representations so that they can compete with more automatic ones (Shallice, 1988). Cohen and colleagues (Botvinick et al., 2001; Miller & Cohen, 2002; O'Reilly, Braver, & Cohen, 1999) suggest that this facilitation is the route to successful performance during the Stroop task; the prefrontal cortex promotes the less automatic process of colour naming in order to overcome the more automatic process of word reading. Note that neither colour naming nor word reading is presumed to be carried out by the prefrontal cortex; rather, the prefrontal cortex directs what phenomenologically corresponds to attention to the weaker process. Similarly, the prefrontal cortex may help resolve the identity of an object shown from a noncanonical viewpoint by activating mental imagery associated with the object's hypothesized identity (Kosslyn, Thomson, & Alpert, 1997). During attempts at memory retrieval, the prefrontal cortex has been shown to activate neurons in the temporal lobe corresponding to the appropriate representations in the temporal cortex to help boost these to a higher level than competing representations (Tomita et al., 1999).

The last way that the prefrontal cortex may exert control is by inhibiting the problematic impulses and representations. A number of rodent studies have suggested that the capacity to override pre-existing associative impulses is dependent on the lateral orbital frontal cortex (Baxter, Parker, Lindner, Izquierdo, & Murray, 2000; Elliott, Dolan, & Frith, 2000; Schoenbaum, Chiba, & Gallagher, 1999, 2000). This area of the rodent brain overlaps with what in human imaging research is often called the ventrolateral or inferior prefrontal cortex (Ongur & Price, 2000).

Neuropsychological and imaging studies have shown that attempts to inhibit prepotent behavioural responses produced ventrolateral prefrontal activations (Cools, Clark, Owen, & Robbins, 2002; Garavan, Ross, & Stein, 1999; Iversen & Mishkin, 1970; Jonides, Smith, Marshuetz, Koeppe, & Reuter-Lorenz, 1998; Konishi et al., 1999; Leung et al., 2000; Levine, Freedman, Dawson, Black, & Stuff, 1999; Monchi, Petrides, Petre, Worsley, & Dagher, 2001). Three recent studies have extended this research into the domain of explicit attempts at emotional self-regulation (Beauregard, Levesque, & Bourgouin, 2001; Ochsner, Bunge, Gross, & Gabrieli, 2002; Small, Zatorre, Dagher, Evans, & Jones-Gotman, 2001), and each study found that the attempt to override a prepotent affective response was associated with ventrolateral prefrontal activations. For instance, Ochsner et al. presented participants with images pretested to produce negative affect. In one condition, participants merely attended to the images, whereas in a second condition, participants were asked to actively reframe or "reappraise" the meaning of the images so as to diminish the negative affect associated with them. This method of reappraisal has long been the subject of study in social psychological theories of emotion (Gross, 1999; Lazarus, 1992; Schachter & Singer, 1962), and evidence suggests that as the meaning of an event changes

for an individual, so the affect changes. Whereas fear-related activations in the amygdala were prominent during the “attend” condition, these activations disappeared during the reappraisal trials. Moreover, in the reappraise condition, a strong negative correlation emerged between the amygdala and ventrolateral prefrontal cortex, suggesting that the latter was inhibiting the former. Also suggestive are two PET studies of bipolar patients during their manic phase. These studies revealed decreased ventrolateral prefrontal activity during decision making, suggesting that one reason these patients become increasingly impulsive during their manic phase is that the ventrolateral prefrontal cortex is temporarily unable to inhibit these impulses (Blumberg et al., 1999; Rubinsztein et al., 2001).

Self-regulation by way of placebos and hypnotic suggestion also implicates the ventrolateral prefrontal cortex. In a recent PET study of placebo analgesia (Lieberman, et al., 2004), patients with irritable bowel syndrome evidenced symptom-related decreases in anterior cingulate responses to painful stimulation. Anterior cingulate activity in response to pain is a remarkably good index of the affective distress given in self-reports (Rainville et al., 1997), and thus placebo analgesia would be expected to reduce anterior cingulate activity. It remains unclear how suggestion and expectation associated with placebos bring about this change. Lieberman et al. ran functional connectivity analyses and found the magnitude of anterior cingulate attenuation after taking placebo was well predicted by the magnitude of right ventrolateral prefrontal activations ($r = -.52$).

Whereas the preceding review of self-control relates effortful attempts at self-regulation with the ventrolateral prefrontal cortex, we have been led to hypothesize that there might also be automatic self-regulatory consequences of engaging the parts of the brain that typically promote self-control. In other words, in most of the previous studies, people were quite intentionally engaging in self-control, the success of which appears to be mediated by the ventrolateral prefrontal cortex. If the ventrolateral prefrontal cortex is capable of inhibiting unwanted impulses in the amygdala and elsewhere, perhaps the mere activation of the ventrolateral prefrontal cortex is sufficient to inhibit automatic affective impulses *even when the activating thoughts are not self-regulatory* (Hariri, Bookheimer, & Mazziotta, 2000; Lieberman, 2003; Lieberman, Hariri, Jarcho, & Bookheimer, 2002).

Contextual factors

Though social psychologists have long been interested in the effects of social and motivational factors on the deployment and efficacy of self-control resources (Baumeister, 1984; Fiske & Neuberg, 1990; Kruglanski & Freund, 1983; Tetlock, 1983), no work has been done within a social cognitive neuroscience framework assessing these interactions. Hopefully, work along these lines will be conducted in the near future.

Social perception

At times, social cognition has been almost synonymous with social perception. Understanding the personality, intentions, beliefs, and identity of others may be the most important form of perception humans engage in. Rhesus monkeys who are unable to appreciate where they stand in the dominance hierarchy, relative to others they interact with, are expelled from the group before they are old enough to fend for themselves (Suomi, in press). Decades of research in social cognition has detailed the processes by which these inferences are made, most often highlighting the common errors that result from these processes. Attribution and stereotyping processes both involve making inferences about the personality, attitudes, intentions, and moral status of others, with attribution focused on behaviour and stereotyping focused on group membership. Both of these areas have begun to be examined from a social cognitive neuroscience approach. In addition, this section will also review work on recognizing others that has emerged primarily from the cognitive neuroscience literature on face processing, as the identity of others is relevant to our production of attributions or access to stereotypes. Though the decoding of emotional expressions is another important domain of social perception, this area has received enough attention to warrant reviews of its own, and thus we direct the reader to them (Lane & Nadel, 2000).

Recognizing others

Neural correlates

Recognizing familiar others seems so basic that social psychologists have rarely ever considered how humans perform this task. Neuropsychological damage impairing face-recognition abilities provides us with dramatic examples of how complex and delicate this recognition process really is. Perhaps the strangest of social perception impairments is Capgras syndrome (also known as reduplicative amnesia). Capgras patients suffer from the delusion that emotionally close, familiar others have been replaced with exact duplicates. These patients admit that the impostors look and act in the same way as the people they replaced, yet despite this evidence, the belief persists. An extraordinary example of the beliefs of a Capgras patient can be found in a client–patient dialogue reprinted in Stuss (1991, pp. 71–73). Though the neural basis of Capgras syndrome has yet to be precisely characterized, research suggests that right hemisphere dysfunction (right orbitofrontal, cingulate, parietal, basal ganglia, and amygdala) is indicative of Capgras and other delusions of misidentification (Breen, Caine, Coltheart, Hendy, & Roberts, 2000; Edelstyn & Oyebode, 1998; Malloy et al., 1992; Papageorgiou et al., 2002).

Another striking disorder of face recognition is prosopagnosia. Prosopagnosics are impaired in their ability to interpret the elements of faces (eyes,

ears, nose, and mouth) coherently, and as such are often incapable of recognizing familiar others by sight alone. However, prosopagnosics are not often impaired in their ability to recognize objects other than faces, which distinguishes them from individuals with other visual agnosias (Farah, 1996). Prosopagnosia results from damage to the occipitotemporal cortex, a fact which led some to speculate that this region may be selective for recognizing faces (Damasio et al., 1982, 1990; Farah, 1994; Meadows, 1974). Face recognition has now been extensively studied in normal populations, and is strongly associated with activity in the mid-fusiform gyrus, on the ventral surface of the brain near the juncture of the temporal and occipital lobes (fusiform face area (FFA) (Kanwisher, McDermott, & Chun, 1997; see also McCarthy, Puce, Gore, & Allison, 1997; Puce et al., 1995). Recent research also suggests areas of the superior temporal sulcus may be face selective, attending to details of facial structure (e.g. Allison et al., 2000; Andrews et al., 2002; Haxby et al., 2000).

Neurocognitive processes

Although it has been traditionally accounted for by psychodynamic explanations, the discovery of Capgras syndrome resulting from neurological conditions has had implications for our understanding of normal face-recognition processes (Ellis & Lewis, 2001). This greater understanding has arisen from comparisons between Capgras patients and prosopagnosics. Ellis and Young (1990) suggested that prosopagnosia and the Capgras delusion are mirror images of each other. Prosopagnosics, while unable to overtly recognize familiar others, often show signs of covert face-recognition abilities which Capgras patients lack, such as differential skin conductance responses (SCRs) to familiar faces (Ellis, Burton, Young, & Flude, 1997; Ellis, Lewis, Moselhy & Young, 2000; Hirstein & Ramachandran, 1997). Clearly, however, Capgras patients must still possess some overt face-recognition abilities: While denying its authenticity, they report that a familiar other's face *looks exactly like* their loved one's.

This differentiation led researchers to suggest that normal face-recognition processes include both overt and covert components. Processes of overt face recognition, as described above, are thought to take place in the fusiform gyrus. Covert face recognition includes both affective responses to familiar faces and behavioural or cognitive recognition as reflected in visual or semantic priming (Ellis et al., 2000). The affective response to familiar faces is associated with amygdala activity, particularly in the right hemisphere (Breen, Caine, & Coltheart, 2000; Ellis et al., 2000; Seeck et al., 1993). It is probably this affective response to familiar faces that Capgras patients are missing, and that many prosopagnosics retain. Ellis and Lewis (2001) speculate that the delusion is thus driven by a lack of affective "glow", which James (1890) postulated was the basis of our sense of familiarity when encountering someone we know well.

Capgras patients, unlike prosopagnosics, are able to recognize faces by their intact ability to interpret the face as a whole, a process some deem critical to face recognition. Proponents of face recognition via holistic processing note that recognition of facial features is much more successful when encoded in the context of the whole face, rather than alone (Tanaka & Farah, 1993). Configural, or view-dependent, processing is often named either concurrently with or as an alternative to holistic processing (Diamond & Carey, 1986). Both were proposed to account for the inversion effect in face recognition first discovered by Yin (1969), wherein recognition of faces is disproportionately impaired when they are presented inverted (upside-down), in comparison to recognition impairments caused when other categories of objects are inverted (Farah, Tanaka, & Drain, 1995).

Evidence supporting the configural or holistic processing function of the FFA includes that experts with some category of nonface objects (such as birds or dogs) demonstrate both inversion effects for recognition of these objects and neural activation comparable to that seen in normal humans viewing faces (Diamond & Carey, 1986; Gauthier, Skudlarski, Gore, & Anderson, 2000; Tanaka & Curran, 2001). Although researchers have found, relative to upright faces, only a slight reduction of activity in the FFA for viewing inverted faces compared to objects, these results do not necessarily disconfirm this proposed functional explanation (Tong, Nakayama, Moscovitch, Weinrib, & Kanwisher, 2000; see also Kanwisher, Tong, & Nakayama, 1998). For example, Haxby et al. (Haxby, Ungerleider, Clark, Schouten, Hoffman, & Martin, 1999) suggest that, first, nearby regions involved in parts-based processing of objects are activated by inverted faces; after determining the object is indeed a face, information is shunted to the FFA for further analysis. This proposal was supported by greater activation of object-processing regions when viewing inverted faces relative to upright faces, and relatively later activation of the FFA when viewing inverted faces relative to upright faces.

Gauthier and her colleagues (Gauthier, Anderson, Tarr, Skudlarski, & Gore, 1997; Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999; Gauthier, Tarr, Moylan, Anderson, Skudlarski, & Gore, 2000) have proposed a counterargument to the domain-specific distinction between parts-based and configural or holistic processing. Her studies have suggested that the unique processes performed in the fusiform gyrus and surrounding regions may be better described as subordinate-level identification. The criticism of previous work supporting a mechanism—and region—selectively engaged in face recognition is that control stimuli in these experiments were not best suited to make such a determination (Gauthier & Tarr, 1997). Specifically, control stimuli in previous behavioural and fMRI studies were members of categories typically recognized at the *basic level* (“That’s a radio”), whereas faces are typically recognized at the *subordinate level* (“That’s John”). Compared to the basic level, participants recognizing objects at the subordinate level were shown to activate the same region of fusiform and inferotemporal cortex

used for face recognition (defined independently for each subject) (Gauthier et al., 2000).

Face detection is yet another functional explanation for the activity seen in the FFA, whether construed broadly as responsiveness to a variety of face-like stimuli or strictly as conscious awareness of faces. Tong et al. (2000) suggested the former role on the basis of research showing FFA activation for a variety of face-like stimuli including human faces, schematic faces, cat faces, and caricatures. The latter functional explanation is suggested by recent research utilizing Rubin's face/vase illusion (Andrews, Schluppeck, Homfray, Matthews, & Blakemore, 2002; Hasson, Hendler, Bashat, & Malach, 2001). Rubin's face/vase illusion is an ambiguous stimulus, which viewers perceive as alternating between a face and a vase. Using an event-related design, Andrews et al. (2002) found that activity in the fusiform gyrus predicted participants' self-reported conscious awareness of perceiving a face in the illusion. They argue that this cortical area explicitly interprets the stimulus as a face, bringing it to conscious awareness, and this process is associated with fusiform activity.

Contextual factors

Although contextual factors have been only recently explored in this realm of research, there are several examples of their application to our understanding of face recognition. One such example is the contrast between the race or ethnicity of the viewer and that of the individual whose face is being recognized. A well-studied and robust phenomenon in the face-recognition literature, the own-race bias (ORB), is that same-race faces are recognized more accurately than other-race faces (also known as other-race effect or cross-race effect) (Malpass & Kravitz, 1969). Golby, Gabrieli, Chiao, & Eberhardt (2000) recently demonstrated that when regions of interest preferentially responsive to faces were individually defined, these regions (left fusiform, right parahippocampus, and hippocampus) were shown to be more reactive to same-race faces. Greater activation in these regions during encoding predicted more accurate recognition outside the scanner.

ORB is a concern in the legal system, as it has implications for eyewitness identification (Meissner & Brigham, 2001). Interestingly, the context of many eyewitness identification scenarios illuminates other factors for future study. For example, face-recognition accuracy is also diminished when weapons are present, because the weapon diverts the observer's attention from the face of the perpetrator ("weapons focus effect") (Kramer, Buckhout, & Eugenio, 1990; Loftus, Loftus, & Messo, 1987). Wojciulik, Kanwisher, & Driver (1998) found that attention (whether internally generated or elicited by factors such as stimulus quality and location) modulated activity in the FFA. Other related contextual factors, which remain to be studied, are the effects of extreme arousal, stress, and motivation on the neural processes of face recognition.

Although we may be suspicious of identifications of familiar others when the eyewitness is motivated to lie, the accuracy of eyewitness testimony is probably most extensively questioned when the perpetrator is unfamiliar to the eyewitness. The differing neural systems involved in recognition of familiar and unfamiliar faces is another prime illustration of the role contextual factors play in the process of face recognition (Leveroni, Seidenberg, Mayer, Mead, Binder, & Rao, 2000). Leveroni et al. found that recognition of famous faces activated additional temporofrontal regions including the superior, middle, and inferior prefrontal areas and the lateral surface of the middle temporal gyrus. They speculate that this reflects the retrieval of additional semantic information associated with the familiar individual's personal identity (the person-identity system; see also Haxby, Hoffman, & Gobbini, 2002; Reinkemeier, Markowitsch, Rausch, & Kessler, 1997).

Attribution

Attributions are a hallmark of the unique human capacity for appreciating the intentions, beliefs, desires, and enduring psychological traits of others. When students raise their hand in class, the teacher easily infers the students' desire to speak. When a student continuously speaks loudly in a library, the librarian has no trouble inferring that this person has a loud disposition, is disrespectful, or both. Social psychological theories have delineated the appropriate rules for inferring mental states and traits from observed behaviours (Jones & Davis, 1965; Kelley, 1967). For instance, it is more appropriate to infer that an enduring disposition was guiding a person's behaviour when the behaviour is counternormative (such as a loud student in the library). To the extent that a behaviour was provoked by situational norms, it is impossible to disentangle the contributions of the norm and a potential internal disposition (for example, a quiet student in the library).

We have introduced most of the previous sections with a neuro-psychological condition that produces impaired functioning in the relevant area of social cognition. In the case of attribution, both normals and neurologically affected patients have difficulty, but in very different ways. Autism is a developmental disorder associated with a reduced capacity to infer the mental states of others and is reviewed at length elsewhere in this volume (Chapters 6 and 8). Healthy adults, on the other hand, make the error of too often assuming that the behaviour of others is the product of corresponding dispositions (Gilbert & Malone, 1995; Jones & Davis, 1965). Icheiser (1943) first observed that "Instead of saying, for instance, the individual X acted in a certain way because he was in a certain situation, we are prone to believe that he behaved in a certain way because he possessed certain personal qualities" (p. 152). Jones and Davis (1965) first provided evidence for this phenomenon when subjects indicated that a target's essay indicated the target's true beliefs even when subjects were informed that the target had been forced to espouse certain beliefs in the essay. This phenomenon, the *correspondence bias* (also

called the *fundamental attribution error*; Ross, 1977), occurs when the disposition corresponding to a behaviour is inferred from the behaviour even when a clear situational explanation for the behaviour is present.

It is fair to say, then, that if autistics have an *underdeveloped* capacity for inferring the mental states of others as causing behaviour, healthy adults have an *overdeveloped* capacity for inferring these mental states. Though the bias is obvious in autistics, it may be less immediately debilitating in adults (Funder, 1987; Swann, 1984). Nevertheless, social psychologists have argued that the correspondence bias leads us to hold others personally responsible when they are mere victims of circumstance. The scale of this problem is global, as nations in long-standing disputes see themselves as reacting to the provocations of the other country while assuming the other nations's actions are driven by the immoral character of its leaders (Griffin & Ross, 1991). Because of the treatment given to attributional difficulties in autistics in other parts of this volume, we will focus primarily on the neural bases of normal attributional processes and why these "normal" processes produce systematically biased attributions.

Neural correlates

The tendency for normal individuals to overattribute behaviour to the internal dispositions of the actor was clearly demonstrated by Heider and Simmel (1944). They showed subjects short videoclips of "fighting triangles" and circles moving around in a coordinated fashion. Though it is immediately obvious upon a moment's reflection that these shapes have no minds at all, it is irresistible to see and describe the movement of the shapes in terms of goals, beliefs, and personality. It is appropriate then, that, modern versions of these clips have been used in many studies to identify a common set of neural regions associated with social perception. Two laboratories (Castelli, Frith, Happe, & Frith, 2002; Schultz et al., 2003) have each found activations in the bilateral superior temporal sulcus, right temporal pole and amygdala, fusiform gyrus, and medial prefrontal cortex in conjunction with watching these videoclips. Iaconi et al. (2004) also found activation in each of these areas when participants watched videoclips of actual people having different kinds of social interactions.

Neurocognitive processes

Gilbert's sequential operations model (1989; see also Trope, 1986) suggests that the attribution process can be decomposed into three stages that occur in sequence: behavioural categorization, dispositional characterization, and correction. According to this model, observed behaviour is initially identified as some specific kind of behaviour ("He is talking in a loud voice in the library"), followed by an inference that this behaviour was caused by an internal disposition ("He is a loud person in general"). In some cases,

the dispositional inference will be subsequently corrected to take account of situational factors that could have provoked the behaviour. (“He is talking to someone who is hard of hearing. Maybe he is not as dispositionally loud as I first thought.”) Critical to this model, the correction stage is presumed to require greater cognitive and motivational resources than the first two stages and will fail if those resources are not available. In other words, it takes mental effort and a desire for accuracy before mitigating situational factors are taken into account. Gilbert suggested that because these resources are often absent in daily life, people would be expected to show the correspondence bias. In a series of studies, Gilbert, Pelham, and Krull (1988) found that when individuals are placed under cognitive load the correspondence bias is enhanced, suggesting: (1) the first two stages of the attribution process are automatic enough to occur while still under cognitive load; (2) the correction stage that incorporates situational factors does not occur under cognitive load.

In a social cognitive neuroscience model of the attribution process (Lieberman, Gaunt, Gilbert, & Trope, 2002), we have suggested that this correction stage depends on the lateral prefrontal cortex, as it is associated with effortful mental processing and the use of propositional logic (Bunge, Hazeltine, Scanlon, Rosen, & Gabrieli, 2002) that mental correction would require. Recent fMRI evidence also suggests that frontal poles may be involved in the correction process (Christoff & Gabrieli, 2000; Kroger et al., 2002). It is not surprising that fMRI studies of attribution have not observed lateral prefrontal activations, because in these studies there was no need for attributional correction and no situational factors relevant to the attributions being made. Presumably, future studies will examine this correction process more directly.

Lieberman et al. (2002) theorized that the lateral and ventral temporal cortex, whose representations are bidirectionally linked in an associative network (Rolls, 1999; Suzuki, Saleem, & Tanaka, 2000), contributes directly to the tendency to conflate behaviours with the corresponding dispositions. These regions of the temporal lobe, from the fusiform gyrus posteriorly to the temporal pole anteriorly, are involved in object identity, categorization, and linking semantic attributes to an object. As described in the previous section, prosopagnosics with fusiform gyrus damage have difficulty linking facial perceptual information to a person’s identity. Alternatively, patients with the temporal variant of frontotemporal dementia (FTD_T; also called semantic dementia) have difficulty retrieving semantic facts about people (Hodges & Graham, 1998). FTD_T is thought to begin by damaging the temporal poles and then slowly work back through the lateral sections of the temporal lobe (Garrard & Hodges, 2000; Mummery, Patterson, Hodges, & Wise, 2000).

The superior temporal sulcus (and gyrus) across numerous studies has been associated with behavioural identification (Alison, Puce, & McCarthy, 2000; Haxby, Hoffman, & Gobbini, 2000), particularly for behaviours that *could* be described as intentional. For instance, Perrett, Jellema, Frigerio, and Burt

(2001) have shown that viewing the same intention enacted different ways activates the same neurons in the primate superior temporal sulcus. However, behaviours that are not seen as intentional (such as falling down) do not activate this region. Some have suggested that this area is central to extracting the possible intentions from perceived biological motion (Gallagher & Frith, 2003); however, the activation of this area in the studies with Heider and Simmel's (1944) "fighting triangles" suggests a more general role in identifying intentions from behaviour. When the triangles in these videos appear to be fighting, there is little in common visually with the biological motion one might observe in an actual fight. It is, however, possible that individuals imagine actual human behaviour while watching these clips and activate the superior temporal sulcus in a top-down, rather than bottom-up, manner.

Because the lateral temporal cortex contains more distributed representations than regions such as the medial temporal cortex and prefrontal cortex (O'Reilly, Braver, & Cohen, 1999), its representations are more like statistical generalizations than distinct symbolic entities (McClelland, McNaughton, & O'Reilly, 1995; O'Reilly & Norman, 2003) and consequently are associatively, rather than propositionally, linked (Holyoak & Hummel, 2000; Sloman, 1996; Smith & DeCoster, 2000). This arrangement may promote the blurring between an actor's current behaviour and the actor's dispositional propensity to that kind of behaviour. In essence, this blurring between action and actor is a neural account of the correspondence bias.

Contextual factors

Social psychologists have long known that various contextual factors can change the attribution process. For instance, people are far more likely to consider the situational antecedents of their own negative behaviour than those of the negative behaviour of others (Dunning, 2003; Ross & Sicoly, 1979). This *self-serving bias* is so named because the self is given the benefit of the doubt, but others are not. The self-serving bias may promote *positive illusions* (Taylor & Brown, 1988) in which views of the self in various domains are aggrandized. Recently, Blackwood et al. (2003) used a vignette-based attribution methodology in an fMRI study to assess the neural correlates of the self-serving bias in attribution. In their study, bilateral basal ganglia activations were associated with this bias. This finding fits well with previous work suggesting the basal ganglia's involvement in automatic positive affect (Depue & Collins, 1999; Knutson, Adams, Hong, & Hommer, 2001; Lieberman, 2000) and automatic self-knowledge processes (Lieberman, Jarcho, & Satpute, 2003).

People show evidence of a positive motivational bias in drawing inferences about others if those others are part of the observers' in-group. In one study (Pettigrew, 1979), Pakistani and Indian subjects observed either a Pakistani or an Indian target engage in positive or negative behaviour. Participants inferred that the negative behaviour of a racial out-group member was more

indicative of their internal dispositions than the same behaviour enacted by a member of their racial in-group. This study suggests that the self-serving bias extends beyond the corporeal self to others that the self shares group membership with or empathizes with.

Empathy may then be considered a key contextual factor moderating attribution processes. As reported above, Iacoboni et al. (2004) found similar activations as the studies that have adapted Heider and Simmel's (1944) "fighting triangles" paradigm. In addition to these shared activations, Iacoboni et al. also found substantial activation in the precuneus. Recall that the main difference between Iacoboni et al. and the other studies was that Iacoboni et al. had participants watch videoclips of real people interacting. Perhaps these videoclips allowed people to identify or empathize with the actors, whereas identification with triangles is unlikely despite the ability of participants to draw inferences about those triangles.

Empathy has just begun to be studied within cognitive neuroscience, but it has a long history within philosophy and social psychology as a critical factor in producing prosocial behaviour and cognition (Batson, 1991; Buber, 1937). Farrow et al. (2001) recently found that both an empathy task and forgiveness judgements, which would seem to require empathy, activated the precuneus whereas other social judgements did not. Earlier, in the section on self-awareness, we described social psychological theories (Mead, 1934) that proposed that each self comes into existence as one takes the perspective of others on oneself. From this view, the self is ontologically dependent on other people treating us as if we have a self and our ability to appreciate that they have this view. The precuneus, associated with perspective taking, a prerequisite for empathy, has been found to be activated in the majority of self-knowledge studies. Here, in this section, work by Iacoboni et al. (2004) and Farrow et al. suggests that this same process may work in reverse such that empathy projects the self onto others, allowing more benevolent attributions about the intentions behind their behaviour.

While self and social perception are not the same thing, they are deeply intertwined. Ultimately, neither can be understood in the absence of the other. Social cognitive neuroscience research has not yet produced a programme of research in which the two are systematically manipulated with respect to each other. Social psychologists, following the tradition of Mead, have created numerous experimental paradigms that activate different social reference points while subjects are making self-judgements. For instance, Baldwin and Holmes (1987) found that subjects produced different self-judgements after being prompted to think about their friends or a family member. Duval and Wicklund (1973) found that participants made different self-judgements while sitting in front of a mirror, as this served as a reminder of the self that is seen by others. Seen from this vantage point, participants tended to evaluate their own behaviour the way a typical observer would (see also Storms, 1973). It is as yet unknown how the neural processes of self-judgement might differ as a result of these activated social referents.

Similarly, it is unknown how the attributions made about others might differ neurally as a function of relationship between the target and the participant.

Before leaving the section on attribution, it is important to briefly mention neuroimaging research on moral attributions (for reviews, see Greene & Haidt, 2002; Moll, Oliveira-Souza, & Eslinger, 2003), which provides another example of how contextual factors can modify the neural structures recruited in a task. For many years, attribution theorists assumed that the rules of logical inference applied to all domains of dispositional inference. Reeder (1993) has demonstrated that there are domain-specific changes in the rules that are applied. For instance, in a skill domain, only successes are construed as diagnostic of ability because ability is considered to be a necessary condition for successful performances, whereas failure might be the result of numerous situational factors. On the other hand, in moral domains, only failures are diagnostic—knowing that someone has lied once tells us more about their moral character than knowing that they told the truth once. The moral domain also differs from other dispositional domains in the corrections made for situational provocations (Reeder et al., 2002). Knowing that Jack tripped Jill, but only did it because he was paid \$100, will lead people to infer that Jack is not as dispositionally aggressive as someone who enacted the same behaviour without financial inducement. However, Jack will be seen as equally immoral in both cases.

As with the behavioural work differentiating moral attributions from other kinds of dispositional attributions, fMRI studies by Moll and colleagues suggest that moral attributions are neurally distinct from other dispositional attributions. In their work, moral attributions activate much of the same network described above, including the superior temporal sulcus, medial prefrontal cortex, and amygdala (Moll, Oliveira-Souza, & Eslinger, 2003), but in several studies moral attributions also activated regions in orbitofrontal cortex (de Oliveira & Moll, 2000; Moll, Oliveira-Souza, Bramati, & Grafman, 2002a; Moll et al., 2002b). Additionally, in these studies, the orbitofrontal cortex exhibited strong functional connectivity with other attributional regions, suggesting a possible modulation of activity in these other regions by the orbitofrontal cortex.

Stereotyping

Humans use category membership to guide behaviour in almost every aspect of life. Learning to use a rental car takes little effort because the assumptions we make about cars in general largely hold true for any rental car. Stereotypes about people in groups are similar to other forms of categorization in many ways; for instance, they increase the efficiency with which we process information about members of those groups (Macrae, Milne, & Bodenhausen, 1994). Social group-based stereotyping is, however, different from object categorization in important ways. Stereotypes often involve negative generalizations about the intelligence, ability, and moral character of group members. That is

to say, that stereotypes of race, gender, age, and sexual orientation typically involve an affective component that is absent in our stereotypes of paper clips and light bulbs. Moreover, unlike light bulbs, humans can react to being stereotyped. Stigmatized groups often fight publicly against the stereotypes applied to them. Unfortunately, the self-awareness that stigmatized group members have of their group's stereotype can actually lead to self-fulfilling prophecies in which the individuals end up confirming stereotypes that were not originally true (Davies, Spencer, Quinn, & Gerhardstein, 2002; Steele & Aronson, 1995). Additionally, in a liberal society, people are aware that most disapprove of these negative social generalizations, and thus people often will not admit their prejudices publicly, even though, cognitively and behaviourally, there is ample evidence of their continued existence. Thus, a great deal of effort may go into self-regulating behaviour that could be perceived as prejudiced. Finally, some stereotypes appear to be so pervasive that people with otherwise egalitarian views may still show evidence of implicit stereotyping even though these beliefs are not held consciously. Implicit stereotyping can be just as pernicious as explicit stereotyping (Word, Zanna, & Cooper, 1974), even more so in some ways, as it may go unchecked.

Though stereotyping processes have begun to be studied with the techniques of cognitive neuroscience, the work thus far has been primarily at the level of brain mapping, in determining the regions of interest that will be probed in later experiments. Here we briefly summarize the results of these studies. A number of studies have shown that visual presentations of black faces activate the amygdala to a greater degree than the presentation of white faces (Cunningham, 2001; Hart et al., 2000; Lieberman, Hariri, Jarcho, Eisenberger, & Bookheimer, 2004; Phelps et al., 2000). Phelps et al. found that the magnitude and extent of amygdala activity correlated with a behavioural measure of implicit stereotypes obtained outside the scanner. Given that implicit stereotypes are thought to be largely the product of repetitive negative presentations of stigmatized groups in schools, media, and elsewhere (Banaji & Greenwald, 1994), it has been suggested that both blacks and whites have negative implicit stereotypes of blacks. Lieberman et al. ran both white ($n=11$) and black ($n=8$) participants to examine this hypothesis and found that both whites and blacks produced more amygdala activity to black faces than to white faces. Milne and Grafman (1999) tested patients with lesions in the ventromedial prefrontal cortex and found no evidence of implicit gender stereotypes. This finding parallels the self-schema work (Lieberman et al., 2004) described above, in which the ventromedial prefrontal cortex was also activated, especially given that self-schemas could easily be described as implicit self-stereotypes. Amodio, Harmon-Jones, Devine, Curtin, Hartley, & Covert (2004) found increased anterior cingulate activation when subjects performed a task designed to make participants feel self-conscious about using their stereotypes. Lastly, Richeson & Shelton (in press) have found that interracial interactions deplete controlled processing resources, leading to poorer performance on a subsequent controlled

processing task. They have followed this study up (Richeson et al., 2003) with an fMRI study suggesting that prefrontal activity involved in self-regulating one's own prejudicial responses may mediate the later controlled processing depletion. This fits with the work by Lieberman et al. (2004) that found that the degree of amygdala suppression is highly correlated with prefrontal activity.

Conclusion

The purpose of this chapter was twofold. First, we wanted to review human research in the social cognitive neuroscience of self-processing and social perception. In just the past few years, there has been a rapid increase in research in these areas, and we suspect that in a few years it will be no more possible to review the two topics together comprehensively than it is to review vision and memory together. We reviewed them together here in the hope of raising awareness of some of the ways that the two areas are linked. Without social perception, it is possible there might not be self-processing. Social perception and the ability to take the perspective of others are potentially critical not only to the inception of the self, but also to the way the self is constructed and maintained at different moments in time. There is a temptation to think of the self as an object with stable attributes—indeed, this is a temptation not just for scientists but for all people who are attached to their sense of self and independence. However, the self appears to be at least partly constructed and reconstructed over time as a function of situational and interpersonal constraints. A complete understanding of the neural bases of self-processes must necessarily take account of these social factors. Similarly, social perception is quite different depending on whether perceivers put themselves in the shoes of the perceived. Unlike some types of perception, social perception is largely motivational. Social stimuli are often ambiguous and are often interpreted in accordance with our self-serving biases. Consequently, neither the study of the self nor the study of social perception can proceed for too long without consideration of the other.

The second purpose of this chapter was to describe three different kind of questions that social psychologists can ask by the methods of cognitive neuroscience. The first question, that of brain mapping, is one for which social psychologists might possess some scepticism. What is the value of knowing where in the brain a social psychological process lives? We have suggested that this answer has value, but that its value is as a means rather than an end. Knowing where a phenomenon is allows us to ask the other questions about cognitive operations, mediators, and moderators. Social psychologists are deeply interested in the questions of how social psychological processes work (that is, the cognitive and affective characteristics) and how these processes interact with other processes and contextual factors (that is, mediators and moderators). To the extent that cognitive and behavioural

testing can answer these questions, neuroscience is unneeded. It would be foolish to think that lots of money needs to be spent on expensive technology just to confirm what we already know is true from behavioural work that has greater statistical power and is frequently more tightly controlled. But there are questions that remain despite our best efforts with cognitive and behavioural measures. Here the methods of cognitive neuroscience may add an important piece to the puzzle. We cannot imagine that the three types of questions we have put forward here are the only kinds that social cognitive neuroscientists will ask in the coming years. Rather, we hope they are a good starting point for thinking broadly about why social psychologists should care about the brain.

Acknowledgements

This work was supported in part by a National Science Foundation grant (BCS-0074562) and a McDonnell-Pew grant (JSMF 99-25 CN-QUA.05) to Matthew Lieberman, and a National Science Foundation graduate fellowship to Jennifer Pfeifer.

References

- Adolphs, R. (1999). Social cognition and the human brain. *Trends in Cognitive Sciences*, 3, 469–479.
- Allen, S. W., & Brooks, L. R. (1991). Specializing the operation of an explicit rule. *Journal of Experimental Psychology: General*, 120, 3–19.
- Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences*, 4, 267–278.
- Allport, G. W. (1955). *Becoming: Basic considerations for a psychology of personality*. New Haven, CT: Yale University Press.
- Amodio, D. M., Harmon-Jones, E., Devine, P. G., Curtin, J. J., Hartley, S. L., & Covert, A. E. (2004). Neural signals for the detection of unintentional race bias. *Psychological Science*, 15, 88–93.
- Baldwin, M. W., & Holmes, J. G. (1987). Salient private audiences and awareness of the self. *Journal of Personality and Social Psychology*, 52, 1087–1098.
- Banaji, M. R., & Greenwald, A. G. (1994). Implicit stereotyping and prejudice. In M. P. Zanna & J. M. Olson (Eds.), *The psychology of prejudice: The Ontario symposium* (Vol. 7, pp. 55–76). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Batson, C. D. (1991). *The altruism question: Toward a social-psychological answer*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Baumeister, R. F. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology*, 46, 610–620.
- Baumeister, R. F. (1998). The self. In D. T. Gilbert, S. T. Fiske & G. Lindzey (Eds.), *The handbook of social psychology* (pp. 680–740). Boston: McGraw-Hill.
- Baxter, M. G., Parker, A., Lindner, C. C., Izquierdo, A. D., & Murray, E. A. (2000). Control of response selection by reinforcer value requires interaction of amygdala and orbital prefrontal cortex. *Journal of Neuroscience*, 20, 4311–4319.

- Beauregard, M., Levesque, J., & Bourgouin, P. (2001). Neural correlates of conscious self-regulation of emotion. *Journal of Neuroscience*, *21*, 6993–7000.
- Beck, D. M., Rees, D. M., Frith, C. D., & Lavie, N. (2001). Neural correlates of change detection and change blindness. *Nature Neuroscience*, *4*, 645–650.
- Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 1–62). New York: Academic Press.
- Bem, D. J., & McConnell, H. K. (1970). Testing the self-perception explanation of dissonance phenomena: On the salience of premanipulated attitudes. *Journal of Personality and Social Psychology*, *14*, 23–31.
- 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.
- Blumberg, H. P., Stern, E., Ricketts, S., Martinez, D., de Asis, J., White, T., et al. (1999). Rostral and orbital prefrontal cortex dysfunction in the manic state of bipolar disorder. *American Journal of Psychiatry*, *156*, 1986–1988.
- Botvinick, M. M., Braver, T. D., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*, 624–652.
- Breen, N., Caine, D., & Coltheart, M. (2000). Models of face recognition and delusional misidentification: A critical review. *Cognitive Neuropsychology*. Special Issue: *The cognitive neuroscience of face processing*, *17*, 55–71.
- Breen, N., Caine, D., Coltheart, M., Hendy, J., & Roberts, C. (2000). Towards an understanding of delusions of misidentification: Four case studies. *Mind and Language*. Special Issue: *Pathologies of belief*, *15*, 74–110.
- Buber, M. (1937). *I and thou*. New York: Scribners & Sons.
- Bunge, S. A., Hazeltine, E., Scanlon, M. D., Rosen, A. C., & Gabrieli, D. E. (2002). Dissociable contributions of prefrontal and parietal cortices to response selection. *NeuroImage*, *17*, 1562–1571.
- Cabeza, R., & Nyberg, L. (2000). Imaging cognition. II. An empirical review of 275 PET and fMRI studies. *Journal of Cognitive Neuroscience*, *12*, 1–47.
- Cacioppo, J. T., Berntson, G. G., Adolphs, R., Carter, C. S., Davidson, R. J., McClinton, M., et al. (2002). *Foundations in social neuroscience*. Cambridge, MA: MIT Press.
- Cahill, L., & McGaugh, L. (1998). Mechanisms of emotional arousal and lasting declarative memory. *Trends in Neuroscience*, *21*, 294–299.
- Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*, *280*, 747–749.
- Carter, C. S., MacDonald, A. W., Botvinick, M. M., Ross, L. L., Stenger, V. A., Noll, D., et al. (2000). Parsing executive processes: Strategic vs. evaluative functions of the anterior cingulate cortex. *Proceedings of National Academy of Sciences of the USA*, *97*, 1944–1948.
- Carver, C. S., & Scheier, M. F. (1981). *Attention and self-regulation: A control theory approach to human behavior*. New York: Springer-Verlag.
- Castelli, F., Frith, C. D., Happe, F., & Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain*, *125*, 1839–1849.
- Chaminade, T., & Decety, J. (2002). Leader or follower? Involvement of the inferior parietal lobule in agency. *Neuroreport*, *13*, 1975–1978.
- Christoff, K., & Gabrieli, D. E. (2000). The frontopolar cortex and human cognition:

- Evidence for a rostrocaudal hierarchical organization within the human prefrontal cortex. *Psychobiology*, 28, 168–186.
- Cools, R., Clark, L., Owen, A. M., & Robbins, T. W. (2002). Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *Journal of Neuroscience*, 22, 4563–4567.
- Craik, F. I. M., Moroz, T. M., Moscovitch, M., Stuss, D. T., Winocur, G., Tulving, E., et al. (1999). In search of the self: A positron emission tomography study. *Psychological Science*, 10, 26–34.
- Cunningham, W. A., Johnson, M. K., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2001). *An fMRI study on the conscious and unconscious evaluations of social groups*. Paper presented at the UCLA Conference on Social Cognitive Neuroscience, Los Angeles, CA.
- Damasio, A. R. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York: Putnam.
- Damasio, A. R., Tranel, D., & Damasio, H. (1990). Face agnosia and the neural substrates of memory. *Annual Review of Neuroscience*, 13, 89–109.
- Davies, P. G., Spencer, S. J., Quinn, D. M., & Gerhardstein, R. (2002). Consuming images: How television commercials that elicit stereotype threat can restrain women academically and professionally. *Personality and Social Psychology Bulletin*, 28, 1615–1628.
- Deacon, T. W. (1997). *The symbolic species: The co-evolution of language and the brain*. New York: Norton.
- de Oliveira-Souza, R., & Moll, J. (2000). The moral brain: A functional MRI study of moral judgement. *Neurology*, 54 (7 Suppl 3), A104.
- Depue, R. A., & Collins, P. F. (1999). Neurobiology of the structure of personality: Dopamine, facilitation of incentive motivation, and extraversion. *Behavioral and Brain Sciences*, 22, 491–569.
- Dewey, J. (1910). *How we think*. Boston: D. C. Heath.
- Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, 115, 107–117.
- Dunning, D. (2003). The zealous self-affirmer: How and why the self lurks so pervasively behind social judgement. In S. Spencer & S. Fein (Eds.), *Motivated social cognition: The 9th Ontario symposium on social cognition* (pp. 45–72). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Duval, S., & Wicklund, R. A. (1973). Effects of objective self-awareness on attribution of causality. *Journal of Experimental Social Psychology*, 9, 17–31.
- Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. Bristol, PA: Taylor & Francis.
- Ebbinghaus, H. (1885/1964). *Memory: A contribution to experimental psychology*. New York: Dover.
- Edelstyn, N. M. J., & Oyeboode, F. (1999). A review of the phenomenology and cognitive neuropsychological origins of the Capgras syndrome. *International Journal of Geriatric Psychiatry*, 14, 48–59.
- Eisenberger, N. I., Satpute, A. B., & Lieberman, M. D. (2003). Self-awareness of arousal states and the anterior cingulate cortex: An fMRI study of personality. *Journal of Cognitive Neuroscience*, 15, s64.
- Elliott, R., Dolan, R. J., & Frith, C. D. (2000). Dissociable functions in the medial and lateral orbitofrontal cortex: Evidence from human neuroimaging studies. *Cerebral Cortex*, 10, 308–317.

- Ellis, A. W., Burton, A. M., Young, A., & Flude, B. M. (1997). Repetition priming between parts and wholes: Tests of a computational model of familiar face recognition. *British Journal of Psychology*, *88*, 579–608.
- Ellis, H. D., & Lewis, M. B. (2001). Capgras delusion: A window on face recognition. *Trends in Cognitive Sciences*, *5*, 149–156.
- Ellis, H. D., Lewis, M. B., Moselhy, H. F., & Young, A. W. (2000). Automatic without autonomic responses to familiar faces: Differential components of covert face recognition in a case of Capgras delusion. *Cognitive Neuropsychiatry*, *5*, 255–269.
- Ellis, H. D., & Young, A. W. (1990). Accounting for delusional misidentifications. *British Journal of Psychiatry*, *157*, 239–248.
- Farah, M. J. (1993). Dissociated overt and covert recognition as an emergent property of a lesioned neural network. *Psychological Review*, *100*, 571–588.
- Farah, M. J. (1994). Neuropsychological inference with an interactive brain: A critique of the “locality” assumption. *Behavioral and Brain Sciences*, *17*, 43–104.
- Farah, M. J. (1996). Is face recognition “special”? Evidence from neuropsychology. *Behavioural Brain Research*, *76*(1–2), 181–189.
- Farah, M. J., Wilson, K. D., Drain, H. M., & Tanaka, J. R. (1995). The inverted face inversion effect in prosopagnosia: Evidence for mandatory, face-specific perceptual mechanisms. *Vision Research*, *35*, 2089–2093.
- Farrer, C., & Frith, C. D. (2002). Experiencing oneself vs another person as being the cause of an action: The neural correlates of the experience of agency. *NeuroImage*, *15*, 596–603.
- Farrow, T. F., Zheng, Y., Wilkinson, I. D., Spence, S. A., Deakin, J. F., TARRIER, N., et al. (2001). Investigating the functional anatomy of empathy and forgiveness. *Neuroreport*, *12*, 2433–2438.
- Fink, G. R., Markowitsch, H. J., Reinkemeier, M., Bruckbauer, T., Kessler, J., & Heiss, W. (1996). Cerebral representation of one’s own past: Neural networks involved in autobiographical memory. *Journal of Neuroscience*, *16*, 4275–4282.
- Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 1–74). New York: Academic Press.
- Fletcher, P. C., Frith, C. D., Baker, S. C., Shallice, T., Frackowiak, R. S., & Dolan, R. J. (1995). The mind’s eye—precuneus activation in memory-related imagery. *NeuroImage*, *2*, 195–200.
- Fredrickson, B. L. (2000). Extracting meaning from past affective experiences: The importance of peaks, ends, and specific emotions. *Cognition and Emotion*, *14*, 577–606.
- Funder, D. C. (1987). Errors and mistakes: Evaluating the accuracy of social judgement. *Psychological Bulletin*, *101*, 75–90.
- Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of “theory of mind”. *Trends in Cognitive Sciences*, *7*, 77–83.
- Garavan, H., Ross, T. J., & Stein, E. A. (1999). Right hemispheric dominance of inhibitory control: An event-related functional MRI study. *Proceedings of the National Academy of Sciences of the USA*, *96*, 8301–8306.
- Garrard, P., & Hodges, J. R. (2000). Semantic dementia: Clinical, radiological and pathological perspectives. *Journal of Neurology*, *247*, 409–422.
- Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. (2000). Expertise for cars

- and birds recruit brain areas involved in face recognition. *Nature Neuroscience*, *3*, 191–197.
- Gauthier, I., & Tarr, M. J. (1997). Becoming a “greeble” expert: Exploring mechanisms for face recognition. *Vision Research*, *37*, 1673–1682.
- Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform “face area” increases with expertise in recognizing novel objects. *Nature Neuroscience*, *2*, 568–573.
- Gauthier, I., Tarr, M. J., Moylan, J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (2000). Does visual subordinate-level categorisation engage the functionally defined fusiform face area? *Cognitive Neuropsychology*. Special Issue: *The cognitive neuroscience of face processing*, *17*, 143–163.
- Gilbert, D. T. (1989). Thinking lightly about others: Automatic components of the social inference process. In J. S. Uleman & J. A. Bargh (eds.), *Unintended thought* (pp. 189–211). New York: Guilford Press.
- Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin*, *117*, 21–38.
- Gilbert, D. T., Pelham, B. W., & Krull, D. S. (1988). On cognitive busyness: When person perceivers meet persons perceived. *Journal of Personality and Social Psychology*, *54*, 733–740.
- Gilbert, D. T., & Wilson, T. D. (2000). Miswanting: Some problems in the forecasting of future affective states. In J. Forgas (Ed.), *Feeling and thinking: The role of affect in social cognition* (pp. 178–197). Cambridge: Cambridge University Press.
- Golby, A. J., Gabrieli, J. D. E., Chiao, J. Y., & Eberhardt, J. L. (2001). Differential responses in the fusiform region to same-race and other-race faces. *Nature Neuroscience*, *4*, 845–850.
- Goldman-Rakic, P. S. (1988). Topography of cognition: Parallel distributed networks in primate association cortex. *Annual Review of Neuroscience*, *11*, 137–156.
- Greene, J., & Haidt, J. (2002). How (and where) does moral judgement work? *Trends in Cognitive Sciences*, *6*, 517–523.
- Griffin, D. W., & Ross, L. (1991). Subject construal, social inference, and human misunderstanding. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 24, pp. 319–359). San Diego, CA: Academic Press.
- Gross, J. J. (1999). Emotion regulation: Past, present, and future. *Cognition and Emotion*, *13*, 551–573.
- Hariri, A. R., Bookheimer, S. Y., & Mazziotta, J. C. (2000). Modulating emotional response: Effects of a neocortical network on the limbic system. *Neuroreport*, *11*, 43–48.
- 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 out-group vs in-group face stimuli. *Neuroreport*, *11*, 2351–2355.
- 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. *Journal of Cognitive Neuroscience*, *13*, 744–753.
- Hawkins, S. A., & Hastie, R. (1990). Hindsight: Biased judgements of past events after the outcomes are known. *Psychological Bulletin*, *107*, 311–327.
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*, 223–233.
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry*, *51*, 59–67.

- Haxby, J. V., Ungerleider, L. G., Clark, V. P., Schouten, J. L., Hoffman, E. A., & Martin, A. (1999). The effect of face inversion on activity in human neural systems for face and object perception. *Neuron*, *22*, 189–199.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.
- Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology*, *57*, 243–259.
- Hirstein, W., & Ramachandran, V. S. (1997). Capgras syndrome: A novel probe for understanding the neural representation of the identity and familiarity of persons. *Proceedings of the Royal Society of London: Series B*, *264*, 437–444.
- Hodges, J. R., & Graham, K. S. (1998). A reversal of the temporal gradient for famous person knowledge in semantic dementia: Implications for the neural organisation of long-term memory. *Neuropsychologia*, *36*, 803–825.
- Holyoak, K. J., & Hummel, J. E. (2000). The proper treatment of symbols in a connectionist architecture. In E. Dietrich & A. B. Markman (Eds.), *Cognitive dynamics: Conceptual and representational change in humans and machines* (pp. 229–263). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Iacoboni, M., Lieberman, M. D., Knowlton, B. J., Molnar-Szakacs, I., Moritz, M., Throop, C. J., et al. (2004). Watching social interactions produces dorsomedial prefrontal cortex and medial parietal BOLD fMRI signal increases compared to a resting baseline. *NeuroImage*, *21*, 1167–1173.
- Ichheiser, G. (1943). Misinterpretations of personality in everyday life and the psychologist's frame of reference. *Character and Personality*, *12*, 145–160.
- Iversen, S. D., & Mishkin, M. (1970). Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity. *Experimental Brain Research*, *11*, 376–386.
- Izquierda, I., Quillfield, J. A., Zanatta, M. S., Quevedo, J., Schaeffer, E., Schmitz, P. K., et al. (1997). Sequential role of hippocampus and amygdala, entorhinal cortex and parietal cortex in formation and retrieval of memory for inhibitory avoidance in rats. *European Journal of Neuroscience*, *9*, 786–793.
- James, W. (1890/1950). *The principles of psychology*. Dover: New York.
- Johnson, S. C., Baxter, L. C., Wilder, L. S., Pipe, J. G., Heiserman, J. E., & Prigatano, G. P. (2002). Neural correlates of self-reflection. *Brain*, *125*, 1808–1814.
- Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 220–266). New York: Academic Press.
- Jonides, J., Smith, E. E., Marshuetz, C., Koeppe, R. A., & Reuter-Lorenz, P. A. (1998). Inhibition in verbal working memory revealed by brain activation. *Proceedings of the National Academy of Sciences of the USA*, *95*, 8410–8413.
- Kahneman, D., Fredrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When more pain is preferred to less: Adding a better ending. *Psychological Science*, *4*, 401–405.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*, 4302–4311.
- Kanwisher, N., Tong, F., & Nakayama, K. (1998). The effect of face inversion on the human fusiform face area. *Cognition*, *68*, B1–B11.
- Keenan, J. P., McCutcheon, B., Freund, S., Gallup, G. G., Jr., Sanders, G., & Pascual-Leone, A. (1999). Left hand advantage in a self-face recognition task. *Neuropsychologia*, *37*, 1421–1425.

- Keenan, J. P., Nelson, A., O'Connor, M., & Pascual-Leone, A. (2001). Self-recognition and the right hemisphere. *Nature*, *409*, 305.
- Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska Symposium on Motivation* (Vol. 15 pp. 192–240). Lincoln: University of Nebraska Press.
- 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. *Journal of Cognitive Neuroscience*, *14*, 785–794.
- Kircher, T. T. J., Senior, C., Phillips, M. L., Benson, P. J., Bullmore, E. T., Brammer, M., et al. (2000). Towards a functional neuroanatomy of self-processing: Effects of faces and words. *Cognitive Brain Research*, *10*, 133–144.
- Kjaer, K. W., Nowak, M., & Lou, H. C. (2002). Reflective self-awareness and conscious states: PET evidence for a common midline parieto-frontal core. *NeuroImage*, *17*, 1080–1086.
- Kjaer, T. W., Nowak, M., Kjaer, K. W., Lou, A. R., & Lou, H. C. (2001). Precuneus-prefrontal activity during awareness of visual verbal stimuli. *Consciousness and Cognition*, *10*, 356–365.
- Klein, S. B., Cosmides, L., Costabile, K. A., & Mei, L. (2002). Is there something special about the self? A neuropsychological case study. *Journal of Research in Personality*, *36*, 490–506.
- Klein, S. B., & Kihlstrom, J. F. (1998). On bridging the gap between social-personality psychology and neuropsychology. *Personality and Social Psychology Review*, *2*, 228–242.
- Klein, S. B., Loftus, J., & Kihlstrom, J. F. (1996). Self-knowledge of an amnesic patient: Toward a neuropsychology of personality and social psychology. *Journal of Experimental Psychology: General*, *125*, 250–260.
- Klein, S. B., Loftus, J., Trafton, J. G., & Fuhrman, R. W. (1992). Use of exemplars and abstractions in trait judgements: A model of trait knowledge about the self and others. *Journal of Personality and Social Psychology*, *63*, 739–753.
- Knutson, B., Adams, C. M., Hong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *Journal of Neuroscience*, *21*, 1–5.
- Kohler, W. (1947). *Gestalt psychology*. New York: Liveright.
- Konishi, S., Nakajima, K., Uchida, I., Kikyo, H., Kameyama, M., & Miyashita, Y. (1999). Common inhibitory mechanism in human inferior prefrontal cortex revealed by event-related functional MRI. *Brain*, *122*, 981–999.
- Kosslyn, S. M., Thompson, W. L., & Alpert, N. M. (1997). Neural systems shared by visual imagery and visual perception: A positron emission tomography study. *NeuroImage*, *6*, 320–334.
- Kramer, T. H., Buckhout, R., & Eugenio, P. (1990). Weapon focus, arousal, and eyewitness memory: Attention must be paid. *Law and Human Behavior*, *14*, 167–184.
- Kroger, J. K., Sabb, F. W., Fales, C. L., Bookheimer, S. Y., Cohen, M. S., & Holyoak, K. J. (2002). Recruitment of anterior dorsolateral prefrontal cortex in human reasoning: A parametric study of relational complexity. *Cerebral Cortex*, *12*, 477–485.
- Kruglanski, A. W., & Freund, T. (1983). The freezing and unfreezing of lay inferences: Effects on impression primacy, ethnic stereotyping, and numerical anchoring. *Journal of Experimental Social Psychology*, *19*, 448–468.

- Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to Western thought*. New York: Basic Books.
- Lane, R. D., Fink, G. R., Chau, P. M.-L., & Dolan, R. J. (1997). Neural activation during selective attention to subjective emotional responses. *Neuroreport*, *8*, 3969–3972.
- Lane, R. D., & Nadel, L. (2000). *Cognitive neuroscience of emotion*. New York: Oxford University Press.
- Lazarus, R. S. (1991). *Emotion and adaptation*. New York: Oxford University Press.
- Leung, H. C., Skudlarski, P., Gatenby, J. C., Peterson, B. S., & Gore, J. E. (2000). An event-related functional MRI study of the Stroop color word interference task. *Cerebral Cortex*, *10*, 552–560.
- Leveroni, C. L., Seidenberg, M., Mayer, A. R., Mead, L. A., Binder, J. R., & Rao, S. M. (2000). Neural systems underlying the recognition of familiar and newly learned faces. *Journal of Neuroscience*, *20*, 878–886.
- Levine, B., Freedman, M., Dawson, D., Black, S., & Stuss, D. T. (1999). Ventral frontal contribution to self-regulation: Convergence of episodic memory and inhibition. *Neurocase*, *5*, 263–275.
- Lieberman, M. D. (2000). Intuition: A social cognitive neuroscience approach. *Psychological Bulletin*, *126*, 109–137.
- Lieberman, M. D. (2003). Reflective and reflexive judgement processes: A social cognitive neuroscience approach. In J. P. Forgas, K. R. Williams, & W. von Hippel (Eds.), *Social judgements: Explicit and implicit processes* (pp. 44–67). New York: Cambridge University Press.
- Lieberman, M. D., Gaunt, R., Gilbert, D. T., & Trope, Y. (2002). Reflection and reflexion: A social cognitive neuroscience approach to attributional inference. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 199–249). New York: Academic Press.
- Lieberman, M. D., Hariri, A., Jarcho, J. J., Eisenberger, N. I., & Bookheimer, S. Y. (2004). *An fMRI investigation of the associative and perceptual nature of race-related amygdala activity*. Unpublished manuscript.
- 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., Jarcho, J. M., & Satpute, A. B. (in press). Evidence-based and intuition-based self-knowledge: An fMRI study. *Journal of Personality and Social Psychology*.
- 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. *Psychological Science*, *121*, 135–140.
- Locke, J. (1975). *An essay concerning human understanding*. Oxford: Oxford University Press.
- Loftus, E. F., Loftus, G. R., & Messo, J. (1987). Some facts about “weapon focus”. *Law and Human Behavior*, *11*, 55–62.
- Macmillan, M. (2000). *An odd kind of fame*. Cambridge, MA: MIT Press.
- Macrae, C. N., Milne, A. B., & Bodenhausen, G. V. (1994). Stereotypes as energy-saving devices: A peek inside the cognitive toolbox. *Journal of Personality and Social Psychology*, *66*, 37–47.
- Malloy, P., Cimino, C., & Westlake, R. (1992). Differential diagnosis of primary and

- secondary Capgras delusions. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology*, 5, 83–96.
- Malpass, R. S., & Kravitz, J. (1969). Recognition for faces of own and other race. *Journal of Personality and Social Psychology*, 13, 330–334.
- Markus, H. R. (1977). Self-schemata and processing information about the self. *Journal of Personality and Social Psychology*, 35, 63–78.
- McCarthy, G., Puce, A., Gore, J. C., & Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience*, 9, 605–610.
- McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102, 419–457.
- Mead, G. H. (1934). *Mind, self, and society*. Chicago: University of Chicago Press.
- Meadows, J. C. (1974). The anatomical basis of prosopagnosia. *Journal of Neurology, Neurosurgery and Psychiatry*, 37, 489–501.
- Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law*, 7, 3–35.
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24, 167–202.
- Miller, G. A., Galanter, E., & Pribram, K. (1960). *Plans and the structure of behavior*. New York: Holt, Rinehart & Winston.
- Milne, E., & Grafman, J. (2001). Ventromedial prefrontal cortex lesions in humans eliminate implicit gender stereotyping. *Journal of Neuroscience*, 21, 151–156.
- Moll, J., de Oliveira-Souza, R., Bramati, I. E., & Grafman, J. (2002a). Functional networks in emotional moral and nonmoral social judgements. *NeuroImage*, 16, 696–703.
- Moll, J., de Oliveira-Souza, R., & Eslinger, P. J. (2003). Morals and the human brain: A working model. *Neuroreport*, 14, 299–305.
- Moll, J., de Oliveira-Souza, R., Eslinger, P. J., Bramati, I. E., Muourao-Miranda, J., Andreiuolo, P. A., et al. (2002b). The neural correlates of moral sensitivity: A functional magnetic resonance imaging investigation of basic and moral emotions. *Journal of Neuroscience*, 22, 2730–2736.
- Monchi, O., Petrides, M., Petre, V., Worsley, K., & Dagher, A. (2001). Wisconsin card sorting revisited: Distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. *Journal of Neuroscience*, 21, 7733–7741.
- Mummery, C. J., Patterson, K., Hodges, J. R., & Wise, R. J. S. (2000). A voxel based morphometry study of semantic dementia: Relationship between temporal lobe atrophy and semantic memory. *Annals of Neurology*, 47, 36–45.
- Nietzsche, F. (1909/1969). *Ecce homo*. New York: Vintage Books.
- Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, 104, 266–300.
- Ochsner, K. N., Bunge, S. A., Gross, J. J., & Gabrieli, J. D. (2002). Rethinking feelings: An fMRI study of the cognitive regulation of emotion. *Journal of Cognitive Neuroscience*, 14, 1215–1229.
- Ochsner, K. N., & Lieberman, M. D. (2001). The emergence of social cognitive neuroscience. *American Psychologist*, 56, 717–734.
- Ongur, D., & Price, J. L. (2000). The organization of networks within the orbital

- and medial prefrontal cortex of rats, monkeys and humans. *Cerebral Cortex*, 10, 206–219.
- O'Reilly, R. C., Braver, T. S., & Cohen, J. D. (1999). A biologically based computational model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 375–411). New York: Cambridge University Press.
- O'Reilly, R. C., & Norman, K. A. (2003). Hippocampal and neocortical contributions to memory: Advances in the complementary learning systems framework. *Trends in Cognitive Sciences*, 6, 505–510.
- Papageorgiou, C., Lykouras, L., Ventouras, E., Uzunoglu, N., & Christodoulou, G. N. (2002). Abnormal P300 in a case of delusional misidentification with coinciding Capgras and Fregoli symptoms. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 26, 805–810.
- Perrett, D. I., Jellema, T., Frigerio, E., & Burt, M. (2001). *Using "social attention" cues (where others are attending) to interpret actions, intentions and emotions of others*. Paper presented at the UCLA Conference on Social Cognitive Neuroscience, Los Angeles, CA.
- Pettigrew, T. F. (1979). The ultimate attribution error: Extending Allport's cognitive analysis of prejudice. *Personality and Social Psychology Bulletin*, 5, 461–476.
- Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., et al. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience*, 12, 729–738.
- Portas, C. M., Strange, B. A., Friston, K. J., Dolan, R. J., & Frith, C. D. (2000). How does the brain sustain a visual percept? *Proceedings of the Royal Society of London: Series B*, 267, 845–850.
- Puce, A., Allison, T., Gore, J. C., & McCarthy, G. (1995). Face-sensitive regions in human extrastriate cortex studied by functional MRI. *Journal of Neurophysiology*, 74, 1192–1199.
- Rainville, P., Duncan, G. H., Price, D. D., Carrier, B., & Bushnell, M. C. (1997). Pain affect encoded in human anterior cingulate but not somatosensory cortex. *Science*, 277, 968–971.
- Ramachandran, V. S. (1995). Anosognosia in parietal lobe syndrome. *Consciousness and Cognition*, 4, 22–51.
- Reeder, G. D. (1993). Trait–behavior relations and dispositional inference. *Personality and Social Psychology Bulletin*, 19, 586–593.
- Reeder, G. D., Kumar, S., Hesson-McInnis, M. S., & Trafimow, D. (2002). Inferences about the morality of an aggressor: The role of perceived motive. *Journal of Personality and Social Psychology*, 83, 789–803.
- Reinkemeier, M., Markowitsch, H. J., Rauch, M., & Kessler, J. (1997). Differential impairments in recalling people's names: A case study in search of neuroanatomical correlates. *Neuropsychologia*, 35, 677–684.
- Richeson, J. A., Baird, A. A., Gordon, H. L., Heatherton, T. F., Wyland, C. L., Trawalter, S. et al. (2003). An fMRI investigation of the impact of interracial contact on executive function. *Nature Neuroscience*, 6, 1323–1328.
- Richeson, J. A., & Shelton, J. N. (2003). When prejudice does not pay: Effects of interracial contact on executive function. *Psychological Science*, 14, 287–290.
- Rolls, E. T. (1999). *The brain and emotion*. New York: Oxford University Press.
- Ross, L. (1977). The intuitive psychologist and his shortcomings. In L. Berkowitz

- (Ed.), *Advances in Experimental Social Psychology* (Vol. 10, pp. 173–220). New York: Academic Press.
- Ross, M. (1989). The relation of implicit theories to the construction of personal histories. *Psychological Review*, *96*, 341–357.
- Ross, M., & Sicoly, F. (1979). Egocentric biases in availability and attribution. *Journal of Personality and Social Psychology*, *37*, 322–336.
- Rubinsztein, J. S., Fletcher, P. C., Rogers, R. D., Ho, L. W., Aigbirhio, F. I., Paykel, E. S., et al. (2001). Decision making in mania: A PET study. *Brain*, *124*, 2550–2563.
- Ruby, P., & Decety, J. (2001). Effect of subjective perspective taking during simulation of action: A PET investigation of agency. *Nature Neuroscience*, *4*, 546–550.
- Sartre, J. P. (1937). *Transcendence of the ego*. New York: Hill & Wang.
- Schachter, S., & Singer, J. E. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, *69*, 379–399.
- Schoenbaum, G., Chiba, A. A., & Gallagher, M. (1999). Neural encoding in orbito-frontal cortex and basolateral amygdala during olfactory discrimination learning. *Journal of Neuroscience*, *19*, 1876–1884.
- Schoenbaum, G., Chiba, A. A., & Gallagher, M. (2000). Changes in functional connectivity in orbitofrontal cortex and basolateral amygdala during learning and reversal training. *Journal of Neuroscience*, *20*, 5179–5189.
- Schultz, R. T., Grelotti, D. J., Klin, A., Kleinman, J., Van der Gaag, C., Marois, R., et al. (2003). The role of the fusiform face area in social cognition: Implications for the pathobiology of autism. *Philosophical Transactions of the Royal Society of London: Series B*, *358*, 415–427.
- Seeck, M., Mainwaring, N., Ives, J., Blume, H., Dubuisson, D., Cosgrove, R., et al. (1993). Differential neural activity in the human temporal lobe evoked by faces of family members and friends. *Annals of Neurology*, *34*, 369–372.
- Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge: Cambridge University Press.
- Shultz, T. R., & Lepper, M. R. (1995). Cognitive dissonance reduction as constraint satisfaction. *Psychological Review*, *103*, 219–240.
- Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*, 3–22.
- Small, D. M., Zatorre, R. J., Dagher, A., Evans, A. C., & Jones-Gotman, M. (2001). Changes in brain activity related to eating chocolate: From pleasure to aversion. *Brain*, *124*, 1720–1733.
- Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, *4*, 108–131.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*, 797–811.
- Storms, M. D. (1973). Videotape and the attribution process: Reversing actors' and observers' points of view. *Journal of Personality and Social Psychology*, *27*, 165–175.
- Stuss, D. T. (1991). Self, awareness, and the frontal lobes: A neuropsychological perspective. In J. Strauss & G. R. Goethals (Eds.), *The self: Interdisciplinary approaches* (pp. 255–277). New York: Springer-Verlag.
- Suomi, S. J. (2003). Gene–environment interactions and the neurobiology of social conflict. *Annals of the New York Academy of Sciences*, *1008*, 132–139.

- Suzuki, W., Saleem, K. S., & Tanaka, K. (2000). Divergent backward projections from the anterior part of the inferotemporal cortex (area TE) in the macaque. *Journal of Comparative Neurology*, *422*, 206–228.
- Swann, W. B. (1984). Quest for accuracy in person perception: A matter of pragmatics. *Psychological Review*, *91*, 457–477.
- Tanaka, J. R., & Curran, T. (2001). A neural basis for expert object recognition. *Psychological Science*, *12*, 43–47.
- Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *46A*, 225–245.
- Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103*, 193–210.
- Tetlock, P. E. (1983). Accountability and complexity of thought. *Journal of Personality and Social Psychology*, *45*, 74–83.
- Tomita, H., Ohbayashi, M., Nakahara, K., Hasegawa, I., & Miyashita, Y. (1999). Top-down signal from prefrontal cortex in executive control of memory retrieval. *Nature*, *401*, 699–701.
- Tong, F., Nakayama, K., Moscovitch, M., Weinrib, O., & Kanwisher, N. (2000). Response properties of the human fusiform face area. *Cognitive Neuropsychology*. Special Issue: *The cognitive neuroscience of face processing*, *17*, 257–279.
- Trope, Y. (1986). Identification and inferential processes in dispositional attribution. *Psychological Review*, *93*, 239–257.
- Turk, D. J., Heatherton, T. F., Kelley, W. M., Funnell, M. G., Gazzaniga, M. S., & Macrae, C. N. (2002). Mike or me? Self-recognition in a split-brain patient. *Nature Neuroscience*, *5*, 841–842.
- Turner, R. H. (1976). The real self: From institution to impulse. *American Journal of Sociology*, *81*, 989–1016.
- Tversky, A., & Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131.
- Vohs, K. D., & Heatherton, T. F. (2000). Self-regulatory failure: A resource-depletion approach. *Psychological Science*, *11*, 249–254.
- Wegener, D. T., & Petty, R. E. (1995). Flexible correction processes in social judgement: The role of naive theories in corrections for perceived bias. *Journal of Personality and Social Psychology*, *68*, 36–51.
- Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review*, *101*, 34–52.
- Whitehead, A. N. (1911). *An introduction to mathematics*. London: Williams and Norgate.
- Wiener, N. (1948). *Cybernetics: Control and communication in the animal and the machine*. Cambridge, MA: MIT Press.
- Wojciulik, E., Kanwisher, N., & Driver, J. (1998). Covert visual attention modulates face-specific activity in the human fusiform gyrus: fMRI study. *Journal of Neurophysiology*, *79*, 1574–1578.
- Word, C. O., Zanna, M. P., & Cooper, J. (1974). The nonverbal mediation of self-fulfilling prophecies in interracial interaction. *Journal of Experimental Social Psychology*, *10*, 109–120.
- Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*, 141–145.

