

The Neuroscience of Social Cognition

David M. Amodio and Kyle G. Ratner

New York University

Keywords: social neuroscience, social cognition, social psychology, attitudes, intergroup, cognitive neuroscience

Contact information:

David M. Amodio
Department of Psychology and Center for Neural Science
New York University
6 Washington Place, rm. 760
New York, NY 10003
david.amodio@nyu.edu

The Neuroscience of Social Cognition

In his foreword to the 2nd edition of the Handbook of Social Cognition, Thomas Ostrom (1994) described the challenges faced by social cognition researchers when this new approach first emerged in the field of social psychology. He noted that when the first edition of the Handbook was published, “many of the skeptics felt that social cognition would prove to be only a minor (and slightly offensive) digression in the history of social psychology.” The publication of the present Handbook is a testament to the lasting impact that social cognition has had on the field. In many ways, the field of social neuroscience appears to be developing along a similar path. Much like the emergence of social cognition two decades earlier, the emergence of social neuroscience is characterized as an integrative, interdisciplinary approach to questions about psychological mechanism. In this way, social neuroscience is a natural extension of social cognition, in that it addresses similar questions but with an expanded approach that incorporates models of biological pathways and new techniques of physiological measurement. Although researchers operating under the rubric of “social neuroscience” study a wide range of processes that extend far beyond the traditional concerns of social psychology and social cognition, the core contributions of social neuroscience have focused on classic socio-cognitive issues.

The term “social neuroscience” was coined by Cacioppo and Berntson (1992; see also Carlston, 1994) to describe the broad enterprise of examining the interplay of social and physiological levels of analysis. The social neuroscience approach began to appear with increasing frequency in the laboratories of social psychologists, cognitive neuroscientists, developmentalists, and neurologists during the 1990s. More recent reviews by Ochsner and Lieberman (2001) and Klein and Kihlstrom (1998) incorporated contemporary ideas from cognitive neuroscience and neuropsychological patient literatures, respectively, prompting additional perspectives on social neuroscience sometimes referred to as “social cognitive neuroscience” and “social

neuropsychology” (heretofore, we will use the term *social neuroscience* to encompass the broad range of approaches). Over the past decade, social neuroscience has been the subject of several dedicated research conferences, culminating in the formation of the *Social and Affective Neuroscience Society* in 2008 and the *Society for Social Neuroscience* in 2010. Whereas social neuroscience was seen as a novelty at social psychology meetings merely a decade ago, it is now approaching mainstream integration.

This review article focuses on neuroscience approaches to classic and contemporary topics in social cognition. It is notable that social cognition has been studied from several perspectives, in the context of different fields of study. Here, we focus on topics within the field of social cognition, which emerged primarily from researchers working in the social psychology and cognitive psychology traditions. In what follows, we begin with an overview of the methodological approaches and measures typically used in social neuroscience. We then describe social neuroscience research across major areas of social cognition research, with a focus on how neuroscience and physiological approaches have suggested advances in social cognition theories. Finally, we discuss some challenges and limitations of social neuroscience research and then comment on its future directions.

The social neuroscience approach

We begin our review by describing the two main types of questions asked in human social neuroscience and their corresponding methodological approaches, with a discussion of how each may contribute to psychological ideas related to social cognition. We also describe the critical role of reverse inference in drawing conclusions from social neuroscience findings.

Brain mapping approach

Brain mapping studies ask “*where in the brain is _____?*” For example, where in the brain is fear? Where is episodic memory? Where is love? Where is the self? Human brain mapping is a

cornerstone of modern cognitive neuroscience. It concerns the mapping of basic psychological processes to particular structures or networks within the brain.

Early forms of brain mapping involved the probing of exposed brain tissue by a neurosurgeon while a patient reported his or her experience. Today, relatively non-invasive neuroimaging measures, such as functional magnetic resonance imaging (fMRI), are often used for a similar purpose. In cognitive neuroscience, this approach is used to map relatively low-level psychological processes such as basic forms of sensation, perception, and specific aspects of learning and memory. As a general rule, lower-level cognitive processes can be mapped more directly than complex high-level processes onto specific physiological responses.

In social neuroscience, researchers often attempt to map very high-level psychological processes, such as social emotions, the self-concept, trait impressions, and political attitudes, onto the brain. This is where things get trickier. For example, to study the neural basis of romantic love, researchers have scanned participants' brains while they viewed pictures of strangers versus their significant others (e.g., Aron, Fisher, Mashek, Strong, Li, & Brown, 2005). Similarly, to study the neural basis of the self, researchers have scanned the brain while subjects judged whether trait adjectives described them vs. another person (Kelley et al., 2002; Mitchell, Banaji, & Macrae, 2005). A potential problem with this approach is that high-level social psychological phenomena often reflect the emergent properties of several lower-level mechanisms. When patterns of brain activity are contrasted, the resulting patterns might reflect the abstract construct or else the lower-level components. Most likely, they represent the lower-level component processes. Therefore, although high-level psychological ascriptions of brain activity may have heuristic value, they may risk obscuring the important low-level mechanisms that the observed brain activations likely represent.

A defining feature of the brain mapping approach is that it seeks to create a valid mapping of psychological processes onto a pattern of neurophysiological responses. Pure brain-mapping studies are undertaken with few prior assumptions about the psychological function of a brain region. Indeed, the point of such studies is to establish ideas about function through the process of induction across multiple studies using a variety of conceptually-similar tasks and manipulations. This approach is potentially useful for generating new ideas about commonalities in the cognitive processes that may underlie two otherwise distinct psychological functions. For example, some researchers have observed that social exclusion and physical pain activate a common region of the anterior cingulate cortex (Fig 1; among many non-overlapping areas) and concluded that social and physical pain share some common neurocognitive features (Eisenberger, Lieberman, & Williams, 2003; but see Somerville, Heatherton, & Kelley, 2006). Although this approach does not tell us exactly how or why they might be related, simply because the true function of the neural activity is difficult to discern, it nevertheless provokes new ideas about potential relationships between psychological processes. Importantly, brain mapping studies are usually not appropriate for testing hypotheses about the relationship between two psychological variables or the effects of an experimental manipulation on a psychological variable.

Hypothesis testing approach

The hypothesis testing approach in social neuroscience is used to test relationships between psychological variables. This approach begins with the assumption that a particular brain region reflects a specific psychological process. In this regard, it does not concern brain mapping, but instead relies on past brain-mapping studies to have already established the validity of neural indicators. For example, a social psychologist who studies intergroup prejudice might hypothesize that implicit racial bias is rooted in mechanisms of classical fear conditioning (Amodio, Harmon-Jones, & Devine, 2003). To test this hypothesis, one might measure brain activity in the amygdala

(Fig 2) – a structure implicated in fear conditioning in many studies – while a participant completes a behavioral measure of implicit racial bias. In this case, the construct validity of the neural measure of fear conditioning (amygdala activity) is already reasonably established (but see Amodio & Ratner, in press), and the question concerns not the meaning of brain activations, but experimental effects among psychological variables. It is the hypothesis testing approach of social neuroscience that is of primary interest to social cognition researchers. Whereas brain-mapping studies typically inform our understanding of the brain, hypothesis-testing studies typically inform psychological theories.

Critically, brain mapping and psychological hypothesis testing approaches should not be combined within a single analysis; major inferential problems occur as a result (Amodio, in press). This is because a test of a psychological hypothesis assumes that the mapping of a psychological variable to a neural structure is already established (e.g., that the neural measure has construct validity). The brain mapping approach is used to establish the mapping between a psychological variable and neural structure (i.e., to establish construct validity of the neural measure). When these approaches are combined, there is a risk of defining the neural operationalization of a psychological construct on the basis of whether it supports one's theoretical hypothesis – an example of tautological inference.

To illustrate, imagine that a researcher wants to test the hypothesis that empathy involves self-reflection. She wants to see whether higher scores on an empathy questionnaire correlate with greater activity in the medial prefrontal cortex (mPFC, see Fig 1), a region previously linked to self-related judgments, while viewing images of people in pain during an fMRI scan. However, the region of mPFC that has been linked to the self is rather large and idiographic across individuals, and so the researcher cannot be sure exactly which specific regions will represent the self. Hence, the construct validity of the “self” measure is very difficult to establish. To deal with

this problem, the researcher might simply examine a correlation between mPFC activity and empathy scores to see which, if any, parts of the mPFC might relate to empathy. If a region is found to correlate, she would simultaneously infer that it must be the “self” region (establishing the construct) and that “self” activity is indeed associated with greater empathy (testing the psychological hypothesis). This blurs the important independent steps of establishing construct validity (brain mapping) and internal validity (hypothesis testing). Because the construct is validated on the basis of the hypothesis-testing correlation analysis, the logic of the test is circular (Amodio, in press; Barrett, 2009). This analytical approach is fairly common in social neuroscience research, primarily because the social psychological processes of interest are complex and difficult to localize. Nevertheless, this approach is problematic, and consumers of social neuroscience should be aware of such practices and be cautious of their use.

Reverse inference

When considering the two general approaches described above, the issue of *reverse inference* is often a concern. Reverse inference refers to a form of reasoning used heavily in social and cognitive neuroscience to infer the psychological meaning of a brain activation based on previous findings (Poldrack, 2006). In brain mapping studies, a psychological process is manipulated and the resulting pattern of brain activity is observed. The inference that the psychological manipulation produced the brain activity may be described as a forward inference, in that the brain activity clearly follows from the manipulation. The inference is based on the known validity of the manipulation. By contrast, the inference of a psychological process from an observed pattern of brain activity is a *reverse inference*. In this case, the precise meaning of the brain activation is ambiguous and inferred from other studies that have used a particular manipulation to activate the same area. The practice of reverse inference becomes increasingly problematic to the extent that the source of inference – in this case, a brain activation – could

reflect different psychological processes (Cacioppo et al., 2003; Poldrack, 2006). In studies of low-level vision, for example, reverse inference may be a comparatively lesser problem (but still a concern). But as psychological variables become more complex, as they do with cognitive and social processes, the mapping between a particular brain region and a psychological process becomes less certain. In these cases, reverse inference can be a serious problem.

All cognitive and social neuroscience studies rely on reverse inference. That is, to the extent that a neural activation is interpreted as reflecting a psychological process, the use of reverse inference is unavoidable. However, researchers can take steps to bolster the strength of a reverse psychological inference by enhancing the construct validity of a neural indicator and the strength of their experimental designs, such as through the careful use of theory, converging evidence from other studies (including animal research), and the use of behavioral tasks that provide valid manipulations of a construct and interpretable behavioral data.

What types of social cognition questions are amenable to a neuroscience analysis?

First and foremost, the brain is a *mechanism*, and an extremely complex one at that. Hence, neuroscience models and methods are useful for the study of psychological mechanisms, such as those involved in action control, perception, and attention. In this regard, the social neuroscience approach is especially well-suited to the study of social cognition. However, psychological processes that refer to phenomena, rather than to mechanisms, are less amenable to a neuroscience level of analysis. For a socio-cognitive psychologist who is considering the potential benefit of a neuroscience approach, the most critical issue is whether one's question concerns basic psychological mechanism. Can the components of one's mechanistic model be described in terms of low-level functions, such as perception, sensation, low-level cognition, and low-level motivation? If so, then neuroscience models may be particularly useful. If a psychological phenomenon of interest cannot be conceptualized at a low level of analysis, but instead is most

meaningful at a high level of construal, then it may be more difficult to make valid inferential connections between psychological theory and the brain.

Methods of social neuroscience

Contemporary social neuroscience makes use of a wide range of methods that are often used in combination with the more traditional tools of social cognition. In addition to new technologies for measurement, social neuroscience methodology relies on the use of careful experimental designs, valid manipulations of psychological states and processes, and careful inference and interpretation. Here, we describe the most prominent methods currently used in the field and briefly discuss their relative advantages as they relate to experimental designs, issues of construct validity, and psychological inference. A more detailed description of methods in neuroscience approaches to social psychology is provided by Harmon-Jones and Beer (2009).

Early studies taking a social neuroscience approach primarily used peripheral physiological methods, such as electrocardiogram (e.g., heart rate), galvanic skin response (i.e., skin conductance, a measure of sympathetic activation (*vis-à-vis* palm sweating), and electromyography (e.g., measures of facial muscle activity related to emotional expressions). More recently, neuroimaging methods have become particularly important and, as such, tend to dominate in the literature. These techniques include functional magnetic resonance imaging (fMRI), which measures the flow of oxygenated blood in the brain, and electroencephalography (EEG), which measures electrical activity produced from the firing of neuron populations. EEG activity measured in response to a discrete event, such as a stimulus or subject response, is called an event-related potential (ERP). These methods differ in important and complementary ways. MRI yields high spatial resolution and thus is optimal for determining the location of activity within the brain. But because it assesses slow-moving blood flow, its temporal resolution is slow. By contrast, EEG/ERP yields high temporal resolution and is thus optimal for assessing the timing

of a neural process, but its spatial resolution is comparatively poor. Given their relative strengths, researchers may select fMRI and EEG methods to suit their question, or use both approaches in complementary studies within a program of research. Neuroimaging and psychophysiological approaches may also be combined with measures of hormones, immune factors, and genetics, for example, to provide convergent evidence for a physiological process of interest. However, as with well-established methods in social cognition, the utility of these measures depends on the quality of the question, the experimental paradigm, and careful interpretation.

Major content areas of social neuroscience

The core themes of social cognition concern issues of mechanism. These range from broad mechanisms that cut across areas of social cognition, such as questions concerning automaticity and control, to mechanisms that underlie more specific processes such as person perception, social evaluation, and mentalizing. Attention to these themes is mirrored in the social neuroscience literature, and this section, we provide a review of how the social neuroscience approach has been used to address them.

Automatic and controlled processing

Theories of automatic and controlled processes form a cornerstone of modern social cognition research. Mechanisms of automaticity and control also constitute a central topic in cognitive psychology, and a large body of cognitive neuroscience research has been devoted to their elucidation.

Automaticity

In the social cognition literature, automaticity is often demonstrated on sequential priming tasks, when a prime word is shown to facilitate the categorization of an associated target word without a participants' awareness or intention (Gaertner & McLaughlin, 1983; Dovidio, Evans, & Tyler, 1986). This idea of automatic semantic priming was originally adapted from cognitive

psychology research examining basic semantic associations between words, such as “bread-butter” or “doctor-nurse” (Meyer & Schaneveldt, 1971, 1976). Research on the neural processes associated with semantic priming has generally found evidence of activations in left posterior PFC (Fig 3; e.g., Blaxton et al., 1996; Demb et al., 1995; Raichle et al., 1994, Wagner, Gabrieli, & Verfaellie, 1997) and temporal cortex (Fig 3, Rissman, Eliassen, & Blumstein, 2003; Schacter & Buckner, 1998; Squire, 1992), and *deactivations* in regions linked to attention (e.g., in the parietal cortex, see Fig 3; Gabrieli, 1998). Given other research implicating the left PFC in approach-related motivation and action tendencies (Harmon-Jones, 2003), this pattern of neural correlates suggests a link between automatic semantic processes and goal-driven behavior (Amodio, 2008), consistent with the idea that “thinking is for doing” (Fiske, 1992).

Much of the neuroscience research relevant to issues of automaticity has been conducted in the area of learning and memory. Classic models of memory distinguish between two general categories of memory on the basis of awareness: explicit and implicit. Explicit (declarative) systems typically include episodic and explicit semantic forms of memory, whereas implicit (non-declarative) systems include classical conditioning, semantic priming, procedural memory (i.e., habit learning), and reflex modification (Squire & Zola, 1996). These different forms of memory have been linked to different underlying neural mechanisms, suggesting the existence of multiple memory “systems.” Explicit forms of memory, such as episodic memory and explicit aspects of semantic memory, have been associated primarily with activity in the hippocampus. Classical fear conditioning has been linked to the amygdala, procedural memory and reward conditioning has been shown to involve the striatum (see Fig 2), and implicit semantic memory and conceptual priming effects have been associated with regions of the PFC. Evidence for the separation of these memory systems comes from studies of brain-lesion patients, in which damage to a particular system results in selective impairments in the associated type of learning and memory. These

findings have been corroborated in many fMRI studies. Social cognition researchers have begun to use findings from this literature to refine their ideas about automatic (vs. controlled) processes (Amodio, 2008; Amodio & Ratner, under review).

Relatively few studies have used social neuroscience approaches to examine automatic social cognition. But of these studies, several have examined the automatic activation of racial bias, with a focus on the role of the amygdala as a substrate of implicit fear associations. These studies have observed greater amygdala activity in response to Black vs. White faces among White participants using measures that can assess rapid amygdala responses following the presentation of a face (Amodio, Harmon-Jones, & Devine, 2003) or in response to near-subliminal presentation of faces (Cunningham et al., 2004). In research on self-schemas, Lieberman, Jarcho, & Satpute (2004) measured brain activity while subjects judged themselves on traits related to domains in which they were highly schematic (i.e., had expertise) or were aschematic (low expertise). Lieberman et al. (2004) reasoned that self-judgments related to highly schematic domains would involve automatic systems to a greater extent than judgments concerning aschematic domains. They found that judgments of familiar domain terms were associated with greater activity in the amygdala, ventral mPFC, basal ganglia regions of the brain (e.g., the striatum) – areas that have been associated with automatic and/or implicit processes in other research.

Automaticity refers to modes of cognitive processing and behavioral responses that are characterized by the properties of unintentionality, unawareness, uncontrollability, and efficiency, or some subset of these (Bargh, 1994). However, it is often difficult to distinguish among these properties using traditional behavioral social cognition methods, as they tend to covary highly in observable behavior. Neuroscience research has been useful in parsing the mechanisms associated with the different properties of automaticity. For the present purposes, it is notable that the degree

of awareness associated with these systems does not correspond with the degree of intentionality (i.e., the extent to which it is goal-driven). For example, classical fear conditioning describes a learned response to a threatening stimulus that is largely reflexive and thus not intentional.

Procedural memory, by contrast, is an instrumental process that is goal driven by definition, and thus it represents an intention processes that may unfold without awareness and may be difficult to inhibit. As another example, the implicit activation of semantic associations may occur without intention, but their activation may instigate goal-driven responses (e.g., ideomotor responses), thereby linking them to intentional responses. Considered together, the cognitive neuroscience research on implicit memory systems suggests that the concept of automaticity is very complex and multi-faceted. These findings support Bargh's (1994) position that automaticity is not one construct, but rather a broad category of different processes that may differ in multiple ways. This literature also suggests that the important differences in the functions of implicit processes are not usefully captured by the implicit vs. explicit distinction (Amodio & Ratner, under review; Henke, 2010).

Control

Within both social psychology and cognitive psychology, control refers to the process of responding in an appropriate and intentional manner despite the existence of countervailing forces. Social and cognitive neuroscience research on the process of control has made major advances in elucidating the mechanisms involved in control. Whereas traditional socio-cognitive models typically assume a single mode of controlled processing, neuroscience research has suggested important distinctions among sub-components of control linked to the detection that control is needed, the implementation of an intended response, and the inhibition of unintended responses. Research is also beginning to distinguish mechanisms for proactive vs. corrective forms of control.

Here, we describe the cognitive neuroscience research on these components of control and describe how these findings have been applied to questions of social cognition.

Conflict monitoring. Corrective forms of control are engaged when an intended response begins to be derailed by countervailing forces, such as an automatic tendency that conflicts with one's intention. In order for corrective control to be engaged, this conflict must be detected as the response unfolds. Cognitive neuroscience research on response conflict tasks, such as the color-naming Stroop task, have examined brain activity that occurs during high-conflict trials, such as when the text of a color-identifying word interferes with one's goal to name the ink color in which the word appears (Carter et al., 1998; MacDonald et al., 2000). In these studies, researchers noted that activity in the anterior cingulate cortex (ACC) was particularly strong during response conflict trials. Botvinick, Braver, Barch, Carter, and Cohen (2001) proposed that the ACC provides a *conflict monitoring* function, such that it is involved in detecting conflict between alternative response tendencies and, when conflict arises, it signals regions of the PFC involved in implementing one's intended response over other tendencies. An important feature of the conflict monitoring theory is that the monitoring process is non-homuncular and operates implicitly, without the need for deliberation or awareness (Berns, Cohen, & Mintun, 1997; Nieuwenhuis et al., 2001). ACC activity linked to conflict monitoring is also associated with passive behavioral inhibition (Amodio et al., 2008). That is, when conflict is detected, a person's responding typically slows as they attempt to gain control (Gehring et al., 1993). This process may result in behavioral inhibition associated with attention and vigilance, which may cause a pause in ongoing behavior. This type of passive inhibition is distinguishable from active forms of inhibitory control that involve the intentional stopping of a response.

The findings of cognitive neuroscience research on conflict monitoring and control have been applied to test theories of self-regulation in social cognitive contexts such as stereotyping and

prejudice (Amodio et al., 2004, 2006, 2008; Bartholow, Dickter, & Sestir, 2006), behavioral inhibition processes (Amodio, Master, Yee, & Taylor, 2008), political orientation (Amodio, Jost, Master, & Yee, 2007), religiosity (Inzlicht, McGregor, Hirsh, & Nash, 2009), and social exclusion (Eisenberger et al., 2003). It is notable that some researchers have suggested that ACC activity on control tasks may represent a distress signal or social pain (Eisenberger et al., 2003; Inzlicht, 2009). However, the fact that the ACC is consistently activated by cognitive tasks that do not involve any type of distress or pain, such as the Stroop task, is a problem for these interpretations. Rather, anatomical research on monkeys has revealed that the ACC is strongly interconnected with motor structures as well as PFC regions associated with high-level representations of goals and actions, consistent with the conflict monitoring hypothesis (Miller & Cohen, 2001).

Implementation of control. Once the need for control is detected and signaled by the ACC, a set of controlled processes become engaged, and these are largely associated with activity in the regions of the PFC. At the neuroanatomical level, PFC regions associated with control have the most extensive connections with structures associated with goal-driven action (Fuster, 2001; Passingham, 1993). A general pattern of connectivity also involves inputs from the thalamus and sensory regions to the medial and orbital frontal cortices, associated with the selection of action-relevant information, and output from lateral PFC areas to structures associated with the planning and implementation of action, such as the basal ganglia and motor cortices (Miller & Cohen, 2001). This evidence is relevant for understanding how socio-cognitive processes are regulated, as it suggests that mechanisms of control operate primarily on motor and perceptual processes.

Cognitive neuroscience research has revealed at least three different forms of motor control that have been linked to separate underlying neural mechanisms. First, goal-directed action refers to motor responses that reflect an intended response. The implementation of a goal-directed behavioral response involves bidirectional connections between the PFC and the striatum (i.e., the

fronto-striatal loop), which operates in concert with thalamic and midbrain processes (Middleton & Strick, 2000; Yin & Knowlton, 2006). Control, in the form of goal-directed action, tends to involve left-lateralized PFC activity in right-handed individuals. Research has shown that stronger left PFC activity is associated with the active control of behavior in the context of regulating intergroup behaviors (Amodio, 2010; Amodio et al., 2007), obtaining rewards (Pizzagalli, Sherwood, Henriques, & Davidson, 2005), and instrumental aggression (Harmon-Jones & Sigelman, 2001).

A second form of control is active inhibition – the intentional stopping of a response. Active inhibition has been linked to right PFC activity, particularly in the right inferior frontal cortex (Aron, Robbins, & Poldrack, 2004). Social neuroscience research examining participants' responses to outgroup vs. ingroup faces has observed activity in the right inferior PFC, which might reflect the participants' attempt to withhold a potentially race-biased response (Cunningham et al., 2004; Lieberman, Hariri, Jarcho, Eisenberger, & Bookheimer, 2005). A third form of motor control pertains to eye-movements. Oculomotor networks constitute an important interface between action and perception in the context of control, and the control of eye movements is associated with activity in dorsal regions of the PFC (Brodmann's area 8) referred to as the "frontal eye field." Although few, if any, studies of social cognition have examined the role of eye-movement control, some research has shown that the regulation of emotional response may be accomplished by intentionally looking away from aversive stimuli (van Reekum et al., 2007), an idea consistent with classic social psychological work on impulse inhibition among children (Mischel, 1974). Given the field's renewed interest in the role of attention and perception in mechanisms of control (e.g., Amodio, 2010), we expect that this form of control will receive greater attention from socio-cognitive theorists in the near future. Importantly, these three forms of control work in coordination, as suggested by their integrated neural connections.

Models of PFC anatomy and function also highlight the effects of control on sensory and perceptual processing. Through dense connections to the thalamus and other sensory structures (Barbas & Zikopoulos, 2007), the PFC is believed to play a role in selecting motivationally-relevant sensory signals while suppressing irrelevant information, in the service of task goals. The PFC continues to modulate the perception of sensory inputs through connections to visual and auditory association cortices (Medalla, Lera, Feinberg, & Barbas, 2007). In an fMRI study of visual processing, efforts to ignore (vs. remember) a visual stimulus were associated with lower activity in the visual association cortex, and this effect was modulated by activity in the left middle frontal gyrus (Gazzaley et al., 2007). Additionally, studies of pain regulation found that PFC activity was associated with changes in the perception of pain (Salomons, Johnstone, Backonja, Shackman, & Davidson, 2007; Wager et al., 2004). These findings suggest that control regions of the PFC function to regulate sensory input as well as higher-level perceptual processes, presumably in the service of facilitating action. These aspects of control have not yet been integrated into sociocognitive theories of control, but they represent promising avenues for future research.

Thus, social and cognitive neuroscience research has helped to expand and refine socio-cognitive models of automaticity and control. In particular, this research has delineated different mechanisms within the broad categories of automatic and controlled processes. These new models make more specific predictions for how automatic and controlled processes interact and are expressed in behavior. Furthermore, these new models have shifted attention away from phenomenological properties of automaticity and control, such as their degree of implicitness vs. explicitness, and toward functional accounts of these processes that are more useful for understanding social cognitive mechanism and behavior (Amodio & Ratner, under review).

The Self

The Self is one of social psychology's oldest and most enduring constructs, and, not surprisingly, it was among the first constructs to be examined in neuroscientific studies of social psychological processes (e.g., Craik et al., 1999; Klein, Loftus & Kihlstrom, 1996). Most of this research has examined brain activity associated with self-reflection and judgments about the self in comparison to judgments of others. Using PET, Craik et al. (1999) found that judgments of trait words as relating to the self vs. others were associated with large activations in regions of mPFC activity. In a similar study that used fMRI, Kelley et al. (2002) found that when comparing self-judgments of trait words with other-judgments (of George W. Bush), a region of ventral mPFC was more strongly activated. The association between self-related processes and activity in ventral regions of the mPFC has since been replicated in several studies (e.g., Gutchess et al., 2007; Heatherton et al., 2006; Kircher et al., 2002; Pfeifer et al., 2007; Saxe, Moran, Scholz, & Gabrieli, 2006; Schmitz, Kawahara-Baccus, & Johnson, 2004; Turner, Simons, Gilbert, Frith, & Burgess, 2008; Zhu, Zhang, Fan, & Han, 2007). Research on other aspects of the self, such as agency and self-discrepancies, have observed regions of the brain typically involved in more general aspects of visual perception, conflict monitoring, and cognitive control (Blakemore, Oakley, & Frith, 2003; Farrer et al., 2008). Thus, brain activations during self-related judgments may not reflect regions dedicated to the "self" per se, but rather regions supporting domain-general processes that are recruited during self-judgment tasks.

Other research has examined the process of self-monitoring in social situations. In this area of work, self monitoring refers to the continuous process by which individuals evaluate their behavior as it relates to the expectations of and experiences with others. Studies have observed that individuals with damage to the orbital frontal cortex (OFC, see Fig 1) are impaired in several aspects of self monitoring. For example, it has been shown that OFC patients have an impaired ability to prioritize solutions to interpersonal problems (Saver & Damasio, 1991), a tendency to

greet strangers in an overly familiar manner (Rolls, Hornak, Wade, & McGrath, 1994), and to behave in disruptive manners in hospital settings (Blair & Cipolotti, 2000). They also tease strangers inappropriately and are more likely to include unnecessary personal information when answering questions (Beer, Heerey, Keltner, Scabini, & Knight, 2003; Kaczmarek, 1984). These findings are consistent with research suggesting that the OFC supports the monitoring of external social signals (Amodio & Frith, 2006)

Although neuroimaging research on the self is an important area of research in social cognitive neuroscience, the findings of this area of research must be considered in light of some complicated interpretational problems. For example, in Kelley et al., (2002), mPFC activity during judgments of the self and other were both lower compared with baseline; that is, self-related activity is associated with a deactivation in mPFC activity. This implies that during baseline periods, when participants simply viewed a fixation cross during an intertrial interval, self-related brain activity was higher than when participants were explicitly thinking about the self. This observation has led some researchers to suggest that research participants spontaneously focus on the self when at rest (to a greater extent than when instructed to think about the self), and this observation has prompted theories about a baseline “default” network of brain activity that may be related to the self (e.g., Gusnard, Akbudak, Shulman, & Raichle, 2001; Gusnard & Raichle, 2001). Interestingly, the idea that humans reflect on the self by default is inconsistent with studies showing that, when beeped at random points during the day and asked to report on what one was doing at the moment, participants rarely (8% of 4700 responses) reported that they were engaged in some form of self-reflection (Csikszentmihalyi & Figurski, 1982). Thus, the notion that people naturally think about themselves during unconstrained periods is not supported by research. Ultimately, a construct like “the self” may be too broad and complex to be localized to a fairly

simple, circumscribed set of neural structures. Rather, it would be more fruitful to consider the more basic-level psychological mechanisms from which the phenomenon of “the self” emerges.

Person Perception

Person perception refers to a set of processes involved in encoding an object as a person, inferring their attributes and mental states, and preparing to interact. Social neuroscience research has been especially active in examining these processes.

Visual perception of faces. Information about conspecifics and social relationships is eminent in perception and cognition, and the initial stage of social processes often begins with face perception. Research on visual perception suggests that some components of the visual system are specialized for seeing faces, and that this component is localized to the fusiform gyrus in fMRI studies (Fig 3; Kanwisher, McDermot, & Chun, 1997). Although the idea of a specialized face area has been debated, with some arguing that fusiform responses to faces reflects expertise rather than a “face module” (Gauthier, Skudlarski, Gore, & Anderson, 2000) or that face perception involves a more distributed network of neural structures (Haxby, Hoffman, & Gobbini, 2000), the finding that this region responds to faces more than to other objects is consistent.

Faces are also known to elicit a characteristic ERP component that peaks 170 ms after the presentation of a face. This “N170” component is consistently larger to faces than non-face stimuli matched on other visual dimensions (Bentin, Allison, Puce, Perez, & McCarthy, 1996; Carmel & Bentin, 2002), and it is believed to represent the structural encoding of face features as a coherent “face” (e.g., Eimer, 2000). Thus, the N170 is a valuable neural marker of the engagement of early face-specific perception processes. The N170 arises from activity in multiple temporal-occipital structures linked to face processing (Deffke et al., 2007), including the fusiform (see Fig. 3; e.g., Haxby et al., 2000; Allison et al., 1994) and other temporal regions (e.g., Desimone, 1991; Perrett, Rolls, & Caan, 1982).

Research has recently begun to examine whether ingroup and outgroup faces are processed differently at the level of the N170 response. This literature has been mixed: in studies comparing responses to racial ingroup vs. outgroup faces, some have observed larger N170 responses to ingroup faces (Ito & Urland, 2005), others have observed larger N170s to outgroup faces (Walker, Silvert, Hewstone, & Nobre, 2008), and several have reported no difference (Caldara, Rossion, Bovet, & Hauert, 2004; Caldara et al., 2003; Wiese, Stahl, & Schweinberger, 2009; He, Johnson, Dovidio, & McCarthy, 2009). However, the experimental tasks used in these studies have varied widely, and few studies have controlled for important visual differences between the faces of people from different racial groups (e.g. luminance, contrast). Therefore, we suspect that the mixed findings reflect particular task differences. To control for extraneous variables when examining the visual perception of ingroup vs. outgroup members, Ratner and Amodio (2010) used a minimal group categorization task to separate faces into the participants' ingroup and outgroup. Participants viewed and classified faces of White males who were purported to belong to one's ingroup or outgroup while their EEG activity was recorded, from which the N170 ERP component was derived. Ratner and Amodio (2010) found that the N170 response was larger to ingroup faces than out group faces. They interpreted this effect as suggesting that greater visual perception processes are allocated to the viewing of ingroup faces in this task.

Ofan, Rubin, and Amodio (2010) conducted a study to test whether N170 differences to racial ingroup vs. outgroup faces might be related to automatic racial attitudes. In their study, White participants completed a sequential evaluative priming task in which White and Black face primes were followed by positive and negative words. The participant's task was to quickly categorize words as pleasant or unpleasant. Overall, participants showed a pattern of implicit bias in which they were more accurate in identifying negative words following Black faces than White faces, but more accurate in identifying positive words following White faces than Black faces.

Importantly, Ofan et al. (2010) observed that stronger pro-White bias in task behavior was associated with larger N170 responses to Black vs. White faces. This finding suggested that participants' existing implicit racial associations may have affected the way they visually perceived ingroup vs. outgroup faces.

Affective responses to faces. Early face perception has also been shown to involve activity of the amygdala, which plays a role in the rapid evaluation of a face (Whalen et al., 1998; Todorov, Said, Engell, & Oosterhof, 2008). The amygdala receives inputs from sensory areas, either directly or via the thalamus, and thus it is responsive to the earliest stages of perception. Given its broad function to coordinate freezing, fight, or flight responses, as well as instrumental approach, the rapid amygdala response to faces appears to provide an extremely adaptive mechanism for responding to social cues. Indeed, in several studies, Todorov and colleagues (2008; Engell, Haxby, & Todorov, 2007) have shown that the amygdala may be involved in assessing signs of threat or untrustworthiness in early stages of face processing. Other research has observed that patients with amygdala damage have impaired responses to faces expressing fear (Adolphs, Tranel, Damasio, & Damasio, 1994; Vuilleumier, Richardson, Armony, Driver, & Dolan, 2004). Through diverse interconnections throughout the brain, the amygdala serves as an important hub for integrating multiple person perception processes, such as attention, emotion, and memory.

Research on implicit forms of intergroup bias has also examined the role of the amygdala in the rapid perception of outgroup members. Early research on the neural correlates of implicit racial bias found that greater anti-Black implicit attitudes, as measured using the Implicit Association Test, was associated with greater amygdala activity to Black vs. White faces, as measured using fMRI (Phelps et al., 2000). In order to make stronger claims about the automaticity of this effect, and to connect it to a classical fear conditioning mechanism of learning

and memory, Amodio et al. (2003) used the startle eyeblink method of assessing amygdala activity. To assess rapid changes in amygdala activity, the authors examined amygdala responses to Black, White, and Asian faces occurring at just 400 ms after a face was presented, as well as at 4000 ms after face presentation. Amodio et al. (2003) observed effects on amygdala activity to Black vs. White faces at 400 ms that were related to individual differences, as well as greater average amygdala responses to Black vs. White faces at 4000 ms. The startle eyeblink measure also allowed us to make stronger inferences about the specific role of the amygdala in implicit intergroup perception. That is, the startle eyeblink effect is modulated by the central nucleus of the amygdala (Fig 4) – the structure specifically involved in the fear response, which includes freezing behavior and vigilance to potential threat, ambiguity, or novelty in one’s environment. By comparison, fMRI measures of amygdala activity cannot easily discern activity in different amygdala nuclei, and thus observed amygdala activity may not relate to fear and vigilance, but rather to the engagement of instrumental behaviors (i.e., via the basal nucleus; see Fig 4) or to other functions. Amodio et al. (2003) used their findings to make a theoretical connection between affective forms of implicit racial bias and mechanisms of fear conditioning linked to the amygdala. Subsequent research replicated this basic effect by showing greater amygdala activity to Black vs. White faces when faces were presented very briefly (~30 ms) using fMRI (Cunningham et al., 2004).

Recent fMRI research on responses to novel ingroup and outgroup faces found that the amygdala was more responsive to novel ingroup faces regardless of the race of ingroup members (Van Bavel, Packer, & Cunningham, 2008). This finding suggests that amygdala activity to faces may track the motivational significance of a social cue, and not simply threat. This broader view is consistent with evidence that the amygdala is important for coordinating attentional and behavioral responses to stimuli that are novel or otherwise of special significance (Whalen, 1998). Although

these findings appear to contradict earlier research suggesting that the amygdala response to outgroup members might reflect a fear-conditioned response, they likely reflect activity of a different region of the amygdala. That is, whereas fear-related responses are associated with the central nucleus and its connections to peripheral autonomic systems, motivated (i.e., instrumental) responses are associated primarily with the basal nucleus and its connections to regions of the brain associated with planning and action (see Fig 4; LeDoux, 2000). Both likely play a role in initial person perception processes.

Categorical responses to faces. How rapidly are people categorized into groups? Above, we reviewed research suggesting that faces are identified as representing persons as quickly as 170 ms following their appearance. Several ERP studies have reported multiple ERP components following the N170 component that reliably distinguish faces on the basis of social group. In particular, the P2 component, named as such because it is the second major positive-going component to follow from a visual stimulus, has been found to be larger to racial outgroups than ingroups (Amodio, 2010; Bartholow & Dickter, 2007; Ito, Willadsen-Jensen, & Correll, 2007; Ito & Urland, 2003). The larger P2 effect to outgroups has been interpreted as reflecting greater perceptual attention to outgroup (vs. ingroup) faces, which could be due to several different reasons, such as concerns about showing prejudice to Black faces, the use of the Black face as a cue to enact a non-prejudiced response, or to threat associated with the outgroup. Other research has used ERP methods to examine the speed with which people's group memberships are identified. It is notable that very early neural responses to social categories likely reflect a combination of bottom-up perceptual processes and top-down processes, whereby the perceiver was actively vigilant for category distinctions. Furthermore, some research has shown that when a task requires classification on one particular social dimension (e.g., gender), early ERP components appear to "implicitly" differentiate race (e.g., Ito & Urland, 2003). This type of

finding may reflect the fact that White participants usually realize that an experimental task presenting pictures of White and Black individuals concerns responses to race, and thus regardless of the classification task, race is salient, and participants may remain vigilant to such cues. In sum, ERP research has shown that the categorical processing of persons occurs very rapidly, but that this rapid processing may reflect top-down anticipatory processes, in addition to bottom-up stimulus-driven processes (Amodio, 2010).

Stereotyping. Whereas much research has examined the neural correlates of evaluation in person perception, relatively little has investigated social stereotypes. Stereotypes are believed to represent cognitive structures stored in memory that represent a set of attributes associated with a social group (Devine, 1989; Hamilton, 1981). Amodio and Devine (2006; see also Amodio, 2008; Amodio & Mendoza, 2010) proposed that stereotypes are rooted in mechanisms of semantic memory and selection, which are associated with neural activity in the temporal lobes and lateral posterior PFC (e.g., Brodmann areas 45 and 47). Behavioral and neuroscience research on semantic learning systems has explored the dynamics of how such associations are learned and expressed in behavior. By linking stereotypes to this literature, researchers can apply findings from the neuroscience and memory literature to understand stereotyping processes (Amodio, 2008). For example, whereas affective associations are learned quickly and are relatively indelible, semantic associations may be learned and unlearned through a process of repeated pairings and non-pairings. Furthermore, semantic learning systems are more likely to be expressed in trait impressions, goal representations, and goal-driven behaviors, and thus they are more likely to emerge in verbal responses (Amodio & Devine, 2006).

Some fMRI studies have examined neural activity associated with the completion of stereotyping tasks (Knutson, Mah, Manly, & Grafman, J, 2007; Mitchell, Ames, Jenkins, & Banaji, 2009; Quadflieg et al., 2009), although they have not explored the mechanisms of stereotyping per

se. That is, these studies have explored brain activity associated with more general processes recruited during the completion of a stereotyping task, such as cognitive conflict, response inhibition, or face perception. For example, in a brain lesion study, patients with mPFC damage did not show bias on a male vs. female IAT (Milne & Grafman, 2001). However, it is likely that the mPFC damage interfered with the general process of response conflict that drives the IAT effect, rather than representing stereotype knowledge (a function typically ascribed to the PFC and temporal lobes). Thus, the neural mechanisms of stereotyping remain largely untested, although researchers have already applied findings from the broader cognitive neuroscience literature on semantic selection and representation to help understand the representation and functions of stereotypes (Amodio, 2008).

Mentalizing and Theory of Mind

The process of inferring another person's unique motives and perspectives is referred to as *mentalizing*, and this process underlies one's *Theory of Mind* (Frith & Frith, 1999). Theory of Mind is best characterized by tasks that involve false belief or deceptive intent – tasks on which successful performance depends on one's ability to take another person's perspective. In an early study, Fletcher et al. (1995) examined brain activity while normal subjects read a set of short stories. These mentalizing stories involved jokes or lies as a literary device – that is, they made sense to the extent that the reader understood that a character was the victim of a lie or joke. Hence, the stories required an understanding of a character's false belief. Control stories did not rely on such devices, but rather involved straightforward physical descriptions. Although several brain regions were activated by these stories, only the mPFC was uniquely more active during the mentalizing stories. A similar set of mentalizing activations was observed in another study when subjects viewed movies of people showing deceptive intent (Grezes et al., 2004).

A subsequent PET experiment by Castelli, Happe, Frith, and Frith (2000) connected the previous findings with the attribution literature in social psychology by measuring brain activity while participants viewed a set of videos inspired by the famous Heider and Simmel (1944) animations, in which three shapes moved in an anthropomorphic fashion that implied social interactions. The authors found that the viewing of this type of animation also elicited mPFC activity, compared with control videos in which the movement of the shapes was not interpreted anthropomorphically. Research by Harris, Todorov, & Fiske (2005) explicitly linked this line of research to social psychological theories of attribution. In their study, participants read descriptions of people's behaviors that varied in consistency distinctiveness, and consensus – the three factors believed to determine dispositional attribution (Jones & Davis, 1976; Kelley, 1972). Harris et al. (2005) observed the strongest degree of mPFC activity in conditions with the highest degree of dispositional attribution.

Over the past decade, a large body of research has associated activity of the mPFC with a range of tasks involving mentalizing and complex aspects of person perception (Amodio & Frith, 2006; Frith & Frith, 1999; Saxe, Carey, & Kanwisher, 2004). These tasks often also elicit activity in regions of the superior temporal lobe (or temporal-parietal junction), which have been linked to the general aspects of attention and visualization of biological motion, as well as activity in the temporal poles, believed to support conceptual representations of social information (Frith & Frith, 1999; but see Scholz, Triantafyllous, Whitfield-Gabrieli, Brown, & Saxe, 2009). The literature on mentalizing makes contact with research on the development of Theory of Mind in children, such that the mPFC is relatively slow to develop compared with other regions (Bunge et al., 2002), adding converging evidence to the idea that the mPFC is an important substrate of these socio-cognitive processes.

Since the original findings linking mentalizing to regions of mPFC, researchers have asked whether other forms of person perception are associated with activity in similar brain regions. A series of studies by Mitchell, Macrae, and colleagues proposed that social-cognitive aspects of person perception, such as the ascription of trait attributes to a person, might also activate areas of mPFC (even if they do not necessarily require mentalizing). For example, when subjects judged noun-adjective word pairs that described a person, compared with those describing an inanimate object, activity was found in regions of interest within the mPFC, as well as areas of the temporal cortex and the temporal-parietal junction (see Fig 3; Mitchell et al., 2002). This pattern of activity has been seen across several studies using similar tasks (e.g., Mitchell et al. 2005, 2006). Other researchers have observed activity in similar regions when simply viewing faces in an easy memory task (Gobbini et al., 2004), demonstrating that activity in this region to faces does not necessarily imply the inference of traits. Some research has found that viewing and making trait judgments of unfamiliar faces or dissimilar people is associated with activity in more dorsal regions of the mPFC, whereas more familiar and/or similar faces are associated with activity in more ventral regions (Gobbini et al., 2004; Mitchell et al., 2006). It is notable, however, that the mPFC is a large region of cortex, and the specific locus of person-related activity varies considerably across studies (Amodio & Frith, 2006; Gilbert et al., 2007).

As with neuroimaging studies of the self, activity associated with forming impressions of both people and inanimate objects is typically lower than baseline mPFC activity (e.g., Mitchell et al., 2004). If the process of person perception is truly located in the mPFC, then the data imply that subjects engage more strongly in person perception during baseline periods (i.e., viewing a fixation cross) than when they are explicitly engaged in the person perception process. This explanation assumes that people naturally reflect upon others when at rest (presumably while also thinking about the self), an assumption that has yet to be supported by data and appears

inconsistent with research on self-reflection (e.g., Csikszentmihalyi & Figurski, 1982). By contrast, other researchers have noted that this region serves a domain-general process of coordinating one's responses with complex (e.g., externally-guided) plans (Amodio & Frith, 2006; Amodio et al., 2006), such as when a research subject prepares for an upcoming trial during intertrial intervals (when mPFC activity is usually highest).

Although the potential contribution of brain-mapping abstract constructs like “the self” and “social cognition” to socio-cognitive theory remains to be established, fMRI research on mentalizing, person perception, and the self has inspired interesting debates about the processes through which a person judges another's thoughts or intentions. “Simulation” theory posits that people consider how the self would respond in the other person's situation and then respond accordingly. “Theory” theory posits that people have an implicit theory of how another person would respond in a particular situation and, rather than reflecting on the self, form their perceptions based on this theory. As evidence for “theory” theory, some researchers have noted that brain activity associated with self and other judgments is related to different regions of the mPFC (Saxe, 2005). However, based on the observations that judgments of similar or familiar others activate a region close to areas activated by self-reflection, other researchers have argued in favor of simulation theory (Mitchell, 2005). Although a lively debate, these interpretations are tentative because they have relied primarily on reverse inferences about the function of brain areas (cf. Poldrack, 2006) and the notion that these regions of the mPFC truly represent the “self” and “social cognition.”

Empathy. Empathy is broadly defined as concern for another's welfare (Batson, 1991), which may involve the process of experiencing another's perspective and affective response (Lamm, Batson, & Decety, 2007). As with mentalizing and Theory of Mind, empathy is complex and involves a broad set of neural and psychological processes associated with affect, perception,

social cognition, self-regulation, mimicry, and action (Decety, in press). Building on neuroscience studies of mentalizing, research on the neural substrates of empathy has focused primarily on the role of the mPFC (Decety, in press; Rameson & Lieberman, 2009). Many studies have examined empathy by measuring brain activity while a subject views another person experiencing pain. For example, Singer et al. (2004) used fMRI to measure brain activity while participants experienced a painful stimulus or viewed a loved one receiving the same stimulus. A set of structures, including the rostral ACC and anterior insula (see Fig 2), were active in both conditions, relative to baseline, and activity in these areas to a loved one's pain was greater among subjects with higher scores on a trait empathy scale. Other research suggests that similar brain regions were more active while watching racial ingroups experiencing pain than racial outgroups (Xu, Zuo, Wang, & Han, 2009). There are also suggestive findings from lesion patient studies, in which damage to the ventral mPFC and ACC are associated with impaired empathy (Shamay-Tsoory, Tomer, Berger, & Aharon-Peretz, 2003). Given that the ACC is involved in a wide range of processes involving expectancy violation, these findings may reflect some aspect of expectancy violation or concern when either the self or another person is subjected to pain, rather than suggesting that empathy is related specifically to the experience of pain. Overall, this body of research has focused primarily on the brain mapping of empathic processes.

Related to work on empathy, a "mirror neuron" system has been proposed as a brain network devoted to understanding other people through their actions (Iacoboni & Dapretto, 2006). The mirror neuron idea originated from single-unit recording in the macaque premotor cortex, in which the same neuron fired when the monkey moved its arm toward a food reward and when it watched an experimenter move its arm toward the reward (Gallese et al., 1996). "Mirror neuron" is not a literal term, in the sense that no single neuron can be described as providing a mirroring function. Rather, "mirror neurons" refer loosely to areas of the brain that are activated both when

an individual observes the behavior of another person, and when one performs the same behavior. Brain regions that have been implicated in “mirror neuron” networks include premotor cortex, inferior frontal cortex, superior temporal sulcus, anterior insula, and the amygdala (Iacoboni & Dapretto, 2006; Rizzolatti & Sinigaglia, 2010), although the patterns and locations of activity in these regions vary considerably from study to study. Although the notion that we relate to other people by representing their actions and mental states in the same way we represent our own actions and states has intuitive appeal, more recent theoretical analyses have questioned the plausibility of mirror neurons as a mechanism of action understanding (Decety, in press; Hickock, 2009; Niedenthal, 2007; Saxe, 2005). Aside from questions about the neural substrates, the fact that so many social interactions often require complementary responses (e.g., when conversing or dancing), rather than mimicry, calls into question the idea that human social behavior is rooted in a mirroring system. Hence, more research will be needed to assess the utility of the mirror neuron idea.

Humanization. Humanization refers to the process of seeing another person as possessing the characteristics unique to the human mind and the rights associated with being a member of society. Hence, dehumanization refers to the denial of such qualities to certain persons (Haslam, 2006). Members of one’s own social group are typically perceived as possessing these qualities, whereas members of a low-status outgroup are often seen as lacking many of these qualities (Leyens et al., 2001, 2003). The process of “humanization” is associated with empathy and mentalizing and typically refers to these processes as they relate to people. By comparison, empathy and mentalizing may also relate to non-humans and inanimate objects, as a form of anthropomorphism (Epley, Waytz, & Cacioppo, 2007). In a study by Harris and Fiske (2006), activity in a region of mPFC was greater when participants viewed pictures of valued others (e.g., members of the ingroup, people of high social status) compared with “dehumanized” individuals,

such as drug abusers and homeless people. Research on dehumanization has connected the neuroscience work on mentalizing and the mPFC to important themes within social psychology concerning intergroup relations (Harris & Fiske, 2009).

Limitations and Challenges

The social neuroscience approach has produced many exciting advances in research on social cognition, ranging from the way we perceive people and infer their motives and beliefs to how we regulate social responses. Nevertheless, as a recently-developed approach to these issues, the field continues to grapple with some important limitations and challenges. We note some of the major issues in this section.

Psychological inference problems. The enterprise of social neuroscience research is built on the assumption that socio-cognitive functions can be inferred from patterns of neural activity. As described in the first section of this article, this assumption is often difficult to meet. In particular, the mPFC region of the frontal cortex is often described as a key substrate of several socio-cognitive processes. However, the specific role of this region in social cognition remains poorly understood. What exactly is the theoretical significance of the mPFC as it relates to the self, person perception, and mentalizing? Mitchell has argued (2009) that the fact that these “social” processes all activate the same regions indicates that social psychology is a “natural kind,” meaning that social psychological processes have a unique and privileged place in neural activity. Although provocative, the “natural kinds” argument is complicated by the consistent finding that the mPFC is more strongly activated at rest than when participants actively engage in social cognition.

Taking a different approach to this issue, Amodio and Frith (2006) considered the neuroanatomical properties of the mPFC in its relation to social cognition. They noted that the mPFC is a highly interconnected region of brain uniquely situated to integrate information about

internal processes (e.g., motor responses, visceral states) and higher-level representations of goals, reward contingencies, and complex expectancies linked to more anterior regions of the frontal cortex. In their analysis, the mPFC is thought to be involved in processes that involve a complex interplay of internal states and the tracking of abstract external contingencies for one's response. Among humans, social cognition is the most important and most complex form of cognition, highly interconnected with emotional, perceptual, motivational, and behavioral processes. According to this view, the mPFC is not the neural instantiation of the self or social cognition, and "social psychology" is not a natural kind in the brain. More research will be needed to understand the significance of the mPFC as it relates to socio-cognitive processes.

Evolving understanding of neural function. Neuroscience is a large and rapidly advancing field, and changes in our understanding of neural function are inevitable. Social neuroscience researchers will need to keep themselves apprised of developments in neuroscience and update interpretations of past findings accordingly. For example, as noted above, interpretations of the amygdala have changed as its function has become better understood. In early social neuroscience work, the amygdala was often interpreted as the "fear" center or as the general locus of emotion in the brain. However, it is now understood to represent a diverse set of processes involved in attention, vigilance, and memory, and the coordination of autonomic reactions as well as instrumental responses (Whalen, 1998; Killcross, Robbins, & Everitt, 1997). These functions also reflect the different functions of subnuclei in the amygdala. Thus, the notion that the amygdala represents "negative emotion" is too simplistic and no longer tenable. The same issue applies to nearly every other neural structure. Therefore, as our understanding of the brain advances, prior interpretations of research findings will likely require revision, and researchers will need to remain open to this possibility.

Limitations in measurement. In addition to the methodological concerns discussed above, physiological and neuroimaging measures sometimes introduce important limitations to the scope of experimental methods normally employed by social psychologists. Beyond issues of cost and training, the recording equipment is sometimes invasive or otherwise constraining, and these factors have direct implications for the manipulation and measurement of psychological variables. For example, fMRI recording requires that a participant lie very still on a narrow scanner bed with his or her head and upper body ensconced in the narrow scanner bore (i.e., a plastic tube). A “bite bar” or other means of immobilizing the participants’ head is often used. During scans, the room is usually darkened, and the participant wears earplugs to attenuate the loud buzzing and whirring noises from the pulsing scanner. This environment places important limitations on the type of research that may be conducted. Experimenters must contend with the participant’s anxiety and distractibility during the study, which may interfere with experimental manipulations. Experimenters must also design tasks that can be implemented with stimulus presentation through LCD goggles (or a back-reflected LCD monitor) and/or responses made on a button box that is usually held in the subject’s right hand.

Beyond these obvious limitations, a recent study found that immobilization in the supine position may significantly reduce the psychological engagement in approach motivation (Harmon-Jones & Peterson, 2010). In line with recent work on embodiment and situated cognition (Smith & Semin, 2004), this research suggests that constraints on a participant’s body (as in the fMRI scanner) have important effects on emotion and cognition, especially as they pertain to action. Other common social neuroscience methods, such as EEG/ERP, are less constrictive, but still present limitations that require special considerations concerning technical and psychological issues.

Future Directions

Social neuroscience is a natural extension of the classic social cognition approach. In both cases, the core questions concern mechanism, and this may be why neuroscience approaches to social psychology have flourished in the domain of social cognition. Just as social cognition transitioned quickly from the fringe to the mainstream in the field of social psychology during the 1980's, social neuroscience ideas and methods are being integrated rapidly into contemporary approaches to social cognition and social psychology more broadly. In addition to new journals and academic societies dedicated to the development of social neuroscience, the broader field of social psychology has made efforts to integrate social neuroscience approaches, through special issues in journals and special training workshops and research symposia at conferences. At the same time, doctoral programs in social psychology increasingly involve training in cognitive neuroscience theories and methods. As social neuroscience becomes integrated into mainstream social cognition research, we expect that scientists will begin to use the approach more effectively to probe social cognitive mechanism, addressing enduring conundrums while generating novel questions and ideas. The research described in this chapter represents the vanguard of the social neuroscience approach. When the second edition of this Handbook of Social Cognition was published, Ostrom commented on the enormous advances achieved in the field since the publication of the first edition. If history is to repeat itself, then we will look forward to the advances in the neuroscience of social cognition that will be described in the next edition.

References

- Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. R. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature*, *372*, 669-72.
- Allison, T., Ginter, H., McCarthy, G., Nobre, A. C., Puce, A., Luby, M. and Spencer, D. D. (1994). Face Recognition in Human Extrastriate Cortex. *Journal of Neurophysiology*, *71*, 821-825.
- Amodio, D. M. (in press). Can neuroscience advance social psychological theory? Social neuroscience for the behavioral social psychologist. *Social Cognition*.
- Amodio, D. M. (2008). The social neuroscience of intergroup relations. *European Review of Social Psychology*, *19*, 1-54.
- Amodio, D. M. (2010). Coordinated roles of motivation and perception in the regulation of intergroup responses: Frontal cortical asymmetry effects on the P2 event-related potential and behavior. *Journal of Cognitive Neuroscience*, *22*, 2609-2617.
- Amodio, D. M. & Devine, P. G. (2006). Stereotyping and evaluation in implicit race bias: Evidence for independent constructs and unique effects on behavior. *Journal of Personality and Social Psychology*, *91*, 652-661.
- Amodio, D. M., Devine, P. G., & Harmon-Jones, E. (2007). A dynamic model of guilt: Implications for motivation and self-regulation in the context of prejudice. *Psychological Science*, *18*, 524-530.
- Amodio, D. M., Devine, P. G., & Harmon-Jones, E. (2008). Individual differences in the regulation of intergroup bias: The role of conflict monitoring and neural signals for control. *Journal of Personality and Social Psychology*, *94*, 60-74.
- Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*, 268-277.
- Amodio, D. M., Harmon-Jones, E., & Devine, P. G. (2003). Individual differences in the activation and control of affective race bias as assessed by startle eyeblink responses and self-report. *Journal of Personality and Social Psychology*, *84*, 738-753.
- 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.
- Amodio, D. M., Jost, J. T., Master, S. L., & Yee, C. M. (2007). Neurocognitive correlates of liberalism & conservatism. *Nature Neuroscience*, *10*, 1246-1247.
- Amodio, D. M., Kubota, J. T., Harmon-Jones, E., & Devine, P. G. (2006). Alternative mechanisms for regulating racial responses according to internal vs. external cues. *Social Cognitive and Affective Neuroscience*, *1*, 26-36.
- Amodio, D. M., Master, S. L. Yee, C. M., & Taylor, S. E. (2008). Neurocognitive components of behavioral inhibition and activation systems: Implications for theories of self-regulation. *Psychophysiology*, *45*, 11-19.
- Amodio, D. M., & Mendoza, S. A. (2010). Implicit intergroup bias: Cognitive, affective, and motivational underpinnings. In B. Gawronski and B. K. Payne (Eds.) *Handbook of Implicit Social Cognition* (pp. 353-374). New York: Guilford.
- Amodio, D. M., & Ratner, K.G. (in press). A social neuroscience analysis of the regulation of intergroup responses. To appear in J. Decety and J. T. Cacioppo, *Handbook of Social Neuroscience*. Oxford University Press.
- Amodio, D. M., & Ratner, K. G. (under review). *A memory systems model of implicit social cognition*.

- Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences*, *8*, 170-177.
- Aron, A., Fisher, F., Mashek, D. A., Strong, G., Li, H., & Brown, L. L. (2005). Reward, motivation, and emotion systems associated with early-stage romantic love. *Journal of Neurophysiology*, *94*, 327-337.
- Barbas, H., & Zikopoulos, B. (2007). The prefrontal cortex and flexible behavior. *The Neuroscientist*, *13*, 532-545.
- Bargh, J. A. (1994). The Four Horsemen of automaticity: Awareness, efficiency, intention, and control in social cognition. In R. S. Wyer, Jr., & T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed., pp. 1-40). Hillsdale, NJ: Erlbaum.
- Barrett, L. F. (2009). Understanding the mind by measuring the brain: Lessons from measuring behavior (Commentary on Vul et al., 2009). *Perspectives on Psychological Science*, *4*, 314-318.
- Bartholow, B. D., & Dickter, C. L. (2007). Social cognitive neuroscience of person perception: A selective review focused on the event-related brain potential. In E. Harmon-Jones & P. Winkielman (Eds.), *Social neuroscience: Integrating biological and psychological explanations of social behavior* (pp. 376-400). New York: Guilford Press.
- Bartholow, B. D., Dickter, C. L., & Sestir, M. A. (2006). Stereotype activation and control of race bias: Cognitive control of inhibition and its impairment by alcohol. *Journal of Personality and Social Psychology*, *90*, 272-287.
- Batson, C. D. (1991). *The Altruism question: toward a social-psychological answer*. Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.
- Beer, J. S., Heerey, E. H., Keltner, D., Scabini, D., & Knight, R. T. (2003). The regulatory function of self-conscious emotion: Insights from patients with orbitofrontal damage. *Journal of Personality and Social Psychology*, *85*, 594-604.
- Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological Studies of Face Perception in Humans. *Journal of Cognitive Neuroscience*, *8*, 551-565.
- Berns, G. S., Cohen, J. D., & Mintun, M. A. (1997). Brain regions responsive to novelty in the absence of awareness. *Science*, *276*, 1272-1275.
- Blair, R. J. R., & Cipolotti, L. (2000). Impaired social response reversal: A case of "acquired sociopathy." *Brain*, *123*, 1122-1141.
- Blakemore, S.-J., Oakley, D. A., & Frith, C. D. (2003). Delusions of alien control in the normal brain. *Neuropsychologia*, *41*, 1058-1067.
- Blaxton, T. A., Bookheimer, S. Y., Zeffiro, T. A., Figlozzi, C. M., William, D. D., & Theodore, W. H. (1996). Functional mapping of human memory using PET: Comparisons of conceptual and perceptual tasks. *Canadian Journal of Experimental Psychology*, *50*, 42-56.
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*, 624-652.
- Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Valdya, C. J., & Gabrieli, J. D. E. (2002). Immature frontal lobe contributions to cognitive control in children: Evidence from fMRI. *Neuron*, *33*, 301-311.
- Cacioppo, J. T., & Berntson, G. G. (1992). Social psychological contributions to the decade of the brain: Doctrine of multilevel analysis. *American Psychologist*, *47*, 1019-1028.
- Cacioppo, J. T., Berntson, G. G., Lorig, T. S., Norris, C. J., Rickett, E., & Nusbaum, H. (2003). Just because you're imaging the brain doesn't mean you can stop using your head: A primer and set of first principles. *Journal of Personality and Social Psychology*, *85*, 650-661.

- Caldara, R., Rossion, B., Bovet, P., & Hauert, C. (2004). Event-related potentials and time course of the 'other-race' face classification advantage. *Neuroreport*, *15*, 905.
- Caldara, R., Thut, G., Servoir, P., Michel, C., Bovet, P., & Renault, B. (2003). Face versus non-face object perception and the 'other-race' effect: a spatio-temporal event-related potential study. *Clinical Neurophysiology*, *114*, 515-528.
- Carlston, D. E. (1994). Associated Systems Theory: A systematic approach to the cognitive representation of persons and events. In R. S. Wyer (Ed.), *Associated systems theory: Advances in social cognition* (Vol. 7, pp. 1-78). Hillsdale, NJ: Erlbaum.
- Carmel, D., & Bentin, S. (2002). Domain specificity versus expertise: factors influencing distinct processing of faces. *Cognition*, *83*, 1-29.
- 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.
- Castelli, F., Happe, F., Frith, U., & Frith, C. (2000). Movement and mind: A functional imaging study of perception and interpretation of complex intentional movement patterns. *Neuroimage*, *12*, 314-325.
- Craik, F. I. M., Moroz, T. M., Moscovitch, M., Stuss, D. T., Winocur, G., Tulving, E., & Kapur, S. (1999). In search of the self: A positron emission tomography study. *Psychological Science*, *10*, 26-34.
- Csikszentmihalyi, M., & Figurski, T. J. (1982). Self-awareness and aversive experience in everyday life. *Journal of Personality*, *50*, 15-28.
- Cunningham, W. A., Johnson, M. K., Raye, C. L., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2004). Separable neural components in the processing of Black and White Faces. *Psychological Science*, *15*, 806-813.
- Decety, J. (in press). To what extent is the experience of empathy mediated by shared neural circuits? *Emotion Review*.
- Deffke, I., Sander, T., Heidenreich, J., Sommer, W., Curio, G., Trahms, L., et al. (2007). MEG/EEG sources of the 170-ms response to faces are co-localized in the fusiform gyrus. *Neuroimage*, *35*, 1495-1501.
- Demb, J. B., Desmond, J. E., Wagner, A. D., Vaidya, Chandan J., Glover, G. H., & Gabrieli, J. D. (1995). Semantic encoding and retrieval in the left inferior prefrontal cortex: A functional MRI study of task difficulty and process specificity. *The Journal of Neuroscience*, *15*, 5870-5878.
- Desimone, R. (1991). Face-selective cells in the temporal cortex of monkeys. *Journal of Cognitive Neuroscience*, *3*, 1-8.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5-18.
- Dovidio, J. F., Evans, N., & Tyler, R. B. (1986). Racial stereotypes: The contents of their cognitive representations. *Journal of Experimental Social Psychology*, *22*, 22-37.
- Eimer, M. (2000). Event-related brain potentials distinguish processing stages involved in face perception and recognition. *Clinical Neurophysiology*, *111*, 694-705.
- Eisenberger, N. I., Lieberman, M. D., & Williams, K. D. (2003). Does rejection hurt? An fMRI study of social exclusion. *Science*, *302*, 290-292.
- Engell, A. D., Haxby, J. V., & Todorov, A. (2007). Implicit trustworthiness decisions: Automatic coding of face properties in human amygdala. *Journal of Cognitive Neuroscience*, *19*, 1508-1519.
- Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, *114*, 864-886.

- Farrer, C., Frey, S. H., Van Horn, J. D., Tunik, E., Turk, D., Inati, S., et al. (2008). The angular gyrus computes action awareness representations. *Cerebral Cortex*, *18*, 254-261.
- Fiske, S. T. (1992). Thinking is for doing: Portraits of social cognition from daguerreotype to laserphoto. *Journal of Personality and Social Psychology*, *63*, 877-889.
- Fletcher, P.C., Happe, F., Frith, U., Baker, S.C., Dolan, R.J., Frackowiak, R.S., et al. (1995). Other minds in the brain: A functional imaging study of "theory of mind" in story comprehension. *Cognition*, *57*, 109-128.
- Frith, C. D., & Frith, U. (1999). Interacting minds—A biological basis. *Science*, *286*, 1692-1695.
- Fuster, J.M. (2001). The prefrontal cortex-An update: Time is of the essence. *Neuron*, *2*, 319-333.
- Gaertner, S. L., & McLaughlin, J. P. (1983). Racial stereotypes: Associations and ascriptions of positive and negative characteristics. *Social Psychology Quarterly*, *46*, 23-30.
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*, 593-609.
- Gauthier, I., Skudlarski, P., Gore, J.C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, *3*, 191-197.
- Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science*, *4*, 385-390.
- Gilbert, S. J., Spengler, S., Simons, J. S., Steele, J. D., Lawrie, S. M., Frith, C. D., & Burgess, P. W. (2006). Functional Specialization within Rostral Prefrontal Cortex (Area 10): A Meta-analysis. *Journal of Cognitive Neuroscience*, *18*, 932-948.
- Gobbini, M. I., Leibenluft, E., Santiago, N., & Haxby, J. V. (2004). Social and emotional attachment in the neural representation of faces. *NeuroImage*, *22*, 1628-1635.
- Gazzaley, A., Rissman, J., Cooney, J.W., Rutman, A., Seibert, T., Clapp, W., & D'Esposito, M. (2007). Functional interactions between prefrontal and visual association cortex contribute to top-down modulation of visual processing. *Cerebral Cortex*, *17*, 125-35
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4-27.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. K. L. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, *74*, 1464-1480.
- Grezes, J., Frith, C., & Passingham, R. F. (2004). Brain mechanisms for inferring deceit in the actions of others. *The Journal of Neuroscience*, *24*, 5500-5505.
- Gross, J. J., & Levenson, R. W. (1993). Emotional suppression: Physiology, self-report, and expressive behavior. *Journal of Personality and Social Psychology*, *64*, 970-986.
- Gross, J. J., & Levenson, R. W. (1997). Hiding feelings: The acute effects of inhibiting negative and positive emotion. *Journal of Abnormal Psychology*, *106*, 95-103.
- Güntürkün, O., Diekamp, B., Manns, M., Nottelmann, F., Prior, H., Schwarz, A., & Skiba, M. (2000). Asymmetry pays: Visual lateralization improves discrimination success in pigeons. *Current Biology*, *10*, 1079-1081.
- Gusnard, D.A., Akbudak, E., Shulman, G.L., Raichle, M.E. (2001). Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *Proceedings of the National Academy of Sciences USA*, *98*, 4259-64.
- Gusnard, D. A., & Raichle, M. E. (2001). Searching for a baseline: Functional imaging and the resting human brain. *Nature Reviews Neuroscience*, *2*, 685-694.
- Gutchess, A. H., Kensinger, E. A., & Schacter, D. L. (2007). Aging, self-referencing, and medial prefrontal cortex. *Social Neuroscience*, *2*, 117-133.

- Hamilton, D. L. (1981). Stereotyping and intergroup behavior: Some thoughts on the cognitive approach. In D.L. Hamilton (ed.), *Cognitive processes in stereotyping and intergroup behavior* (pp. 333-353). Hillsdale, NJ: Erlbaum.
- Harmon-Jones, E. (2003). Anger and the behavioural approach system. *Personality and Individual Differences, 35*, 995-1005.
- Harmon-Jones, E., & Beer, J. S. (Eds., 2009). *Methods in social neuroscience*. Guilford Publications: New York.
- Harmon-Jones, E., & Peterson, C. K. (2009). Supine body position reduces neural response to anger evocation. *Psychological Science, 20*, 1209-1210.
- Harmon-Jones, E., & Sigelman, J. D. (2001). State anger and prefrontal brain activity: Evidence that insult-related relative left-prefrontal activation is associated with experienced anger and aggression. *Journal of Personality and Social Psychology, 80*, 797-803.
- Harris, L. T., & Fiske, S. T. (2006). Dehumanizing the lowest of the low: Neuro-imaging responses to extreme outgroups. *Psychological Science, 17*, 847-853.
- Harris, L. T., & Fiske, S. T. (2009). Social neuroscience evidence for dehumanised perception. *European Review of Social Psychology, 20*, 192-231.
- Harris, L. T., Todorov, A., & Fiske, S. T. (2005). Attributions on the brain: Neuro-imaging dispositional inferences, beyond Theory of Mind. *NeuroImage, 28*, 763-769.
- Haxby, J., Hoffman, E., & Gobbini, M. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4*, 223-232.
- He, Y., Johnson, M. K., Dovidio, J. F., & McCarthy, G. (2009). The relation between race-related implicit associations and scalp-recorded neural activity evoked by faces from different races. *Social Neuroscience, 4*, 426 – 442.
- Heatherton, T. F., Wyland, C. L., Macrae, C. N., Demos, K. E., Denny, B. T., & Kelley, W. M. (2006). Medial prefrontal activity differentiates self from close others. *Social Cognitive and Affective Neuroscience, 1*, 18-25.
- Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology, 57*, 243-259.
- Henke, K. (2010). A model for memory systems based on processing modes rather than consciousness. *Nature Reviews Neuroscience, 11*, 523-532.
- Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of cognitive neuroscience, 21*, 1229-1243..
- Iacoboni, M., & Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nature Reviews Neuroscience, 7*, 942-951.
- Inzlicht, M., McGregor, I., Hirsh, J. B., & Nash, K. A. (2009). Neural markers of religious conviction. *Psychological Science, 20*, 385-392.
- Ito, T., & Urland, G. (2005). The influence of processing objectives on the perception of faces: An ERP study of race and gender perception. *Cognitive, Affective, & Behavioral Neuroscience, 5*, 21-36.
- Ito, T.A., Willadsen-Jensen, E.C., & Correll, J. (2007). Social neuroscience and social perception: New perspectives on categorization, prejudice, and stereotyping. In E. Harmon-Jones & P. Winkielman (Eds.), *Social Neuroscience: Integrating Biological and Psychological Explanations of Social Behavior* (pp. 401–421). New York: Guilford.
- Jones, E. E., & Davis, K. E. (1965). From Acts to Dispositions: The Attribution Process in Person Perception. In Leonard Berkowitz, ed., *Advances in Experimental Social Psychology, Vol. 2* (pp. 219-266). New York: Academic Press.
- Kaczmarek, B. L. J. (1984). Neurolinguistic analysis of verbal utterances in patients with focal lesions of frontal lobes. *Brain and Language, 21*, 52–58.

- Kanwisher, N. G., 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.
- Kelley, H. H. (1972) Attribution in social interaction, In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins and B. Weiner (eds.), *Attribution: Perceiving the causes of behavior* (pp. 1-26), Morristown, NJ: General Learning 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.
- Killcross, A. S., Robbins, T. W., & Everitt, B. J. (1997). Different types of fear conditioned behavior mediated by separate nuclei in amygdala. *Nature*, *388*, 377-380.
- Kircher, T. T. J., Brammer, M., Bullmore, E., Simmons, A., Bartels, M., & David, A. S. (2002). The neural correlates of intentional and incidental self processing. *Neuropsychologia*, *40*, 683-692.
- 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.
- Knutson, K.M., Mah, L., Manly, C.F., and Grafman, J. (2007). Neural Correlates of Automatic Beliefs about Gender and Race. *Human Brain Mapping*, *28*, 915-930.
- Lamm, C., Batson, C. D., & Decety, J. (2007). The neural substrate of human empathy: effects of perspective-taking and cognitive appraisal. *Journal of Cognitive Neuroscience*, *19*, 42-58.
- LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience*, *23*, 155-184.
- Lieberman, M. D., Jarcho, J. M., & Satpute, A. B. (2004). Evidence-based and intuition-based self-knowledge: An fMRI study. *Journal of Personality and Social Psychology*, *87*, 421-35.
- Lieberman, M. D., Hariri, A., Jarcho, J. M., Eisenberger, N. I., & Bookheimer, S. Y. (2005). An fMRI investigation of race-related amygdala activity in African-American and Caucasian-American individuals. *Nature Neuroscience*, *8*, 720-722.
- MacDonald, A.W., III, Cohen, J.D., Stenger, V.A., & Carter, C.S. (2000). Dissociating the role of the dorsolateral prefrontal cortex and anterior cingulate cortex in cognitive control. *Science*, *288*, 1835-1838.
- Medalla, M., Lera, P., Feinberg, M., & Barbas, H. (2007). Specificity in inhibitory systems associated with prefrontal pathways to temporal cortex in primates. *Cerebral Cortex*, *17*, 136-50.
- Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*, 227-234.
- Meyer, D. E., & Schvaneveldt, R. W. (1976). Meaning, memory, structure, and mental processes. *Science*, *192*, 27-33.
- Middleton, F.A., & Strick, P.L. (2000). Basal ganglia output and cognition: evidence from anatomical, behavioral, and clinical studies. *Brain and Cognition*, *42*, 183-200.
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167-202.
- Milne, E., & Grafman, J. (2001). Ventromedial prefrontal cortex lesions in humans eliminate implicit gender stereotyping. *The Journal of Neuroscience*, *21*, 1-6.

- Mischel, W. (1974). Processes in delay of gratification. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 7, pp. 249-292). New York: Academic Press.
- Mitchell, J. P. (2005). The false dichotomy between simulation and theory-theory: The argument's error. *Trends in Cognitive Sciences*, *9*, 363-364.
- Mitchell, J.P.(2009). Social psychology as a natural kind. *Trends in Cognitive Sciences*, *13*, 246-251.
- Mitchell, J. P., Ames, D. L., Jenkins, A. C., & Banaji, M. R. (2009). Neural correlates of stereotype application. *Journal of Cognitive Neuroscience*, *21*, 594-604.
- Mitchell, J. P., Banaji, M. R., & Macrae, C. N. (2005). General and specific contributions of the medial prefrontal cortex to knowledge about mental states. *NeuroImage*, *28*, 757-762.
- Mitchell, J. P., Heatherton, T. F., & Macrae, C. N. (2002). Distinct neural systems subserve person and object knowledge. *Proceedings of the National Academy of Sciences*, *99*, 15238-15243.
- Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2004). Encoding-specific effects of social cognition on the neural correlates of subsequent memory. *Journal of Cognitive Neuroscience*, *24*, 4912-4917.
- Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron*, *50*, 655-663.
- Niedenthal, P.M. (2007). Embodying emotion. *Science*, *316*, 1002-1005.
- Nieuwenhuis, S., Ridderinkhof, K. R., Blom, J., Band, G. P. H., & Kok, A. (2001). Error-related brain potentials are differently related to awareness of response errors: Evidence from an antisaccade task. *Psychophysiology*, *38*, 752-760.
- Ochsner, K. N., & Lieberman, M. D. (2001). The emergence of social cognitive neuroscience. *American Psychologist*, *56*, 717-734.
- Ostrom T. (1994). Foreword. In R. S. Wyer, Jr., & T. K. Srull (Eds.), *Handbook of social cognition*, 2nd Ed. (pp. vii-xii). Hillsdale, NJ: Erlbaum.
- Passingham, R. (1993). *The Frontal Lobes and Voluntary Action*. Oxford, UK: Oxford University Press.
- Perrett, D., Rolls, E., & Caan, W. (1982). Visual neurons responsive to faces in the monkey temporal cortex. *Experimental Brain Research*, *47*, 329-342.
- Pfeifer, J. H., Lieberman, M. D., & Dapretto, M. (2007). "I know you are but what am I?!": Neural Bases of self- and social knowledge in children and adults. *Journal of Cognitive Neuroscience*, *19*, 1323-1337.
- Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience*, *12*, 729-738.
- Pizzagalli, D. A., Sherwood, R. J., Henriques, J. B., & Davidson, R. J. (2005). Frontal brain asymmetry and reward responsiveness: A source-localization study. *Psychological Science*, *16*, 805-813.
- Poldrack, R.A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, *10*, 59-63.
- Quadflieg, S., Turk, D. J., Waiter, G. D., Mitchell, J. P., Jenkins, A. C., & Macrae, C. N. (2009). Exploring the neural correlates of social stereotyping. *Journal of Cognitive Neuroscience*, *21*, 1560-1570.
- Raichle, M. E., Fiez, J. A., Videen, T. O., MacLeod, A. K., Pardo, J. V., et al. (1994). Practice-related changes in human brain functional anatomy during nonmotor learning. *Cerebral Cortex*, *4*, 8-26.

- Rameson, L. T., & Lieberman, M. D. (2009). Empathy: A social cognitive neuroscience approach. *Social and Personality Psychology Compass*, 3, 94-110.
- Rissman, J., Eliassen, J. C., & Blumstein, S. E. (2003). An event-related fMRI investigation of implicit semantic priming. *Journal of Cognitive Neuroscience*, 15, 1160-1175.
- Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nature Reviews Neuroscience*, 11, 264-274.
- Rolls, E. T., Hornak, J., Wade, D., & McGrath, J. (1994). Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. *Journal of Neurology, Neurosurgery, and Psychiatry*, 57, 1518-1524.
- Salomons TV, Johnstone T, Backonja MM, Shackman AS, Davidson RJ. (2007) Individual Differences in the Effects of Perceived Controllability on Pain Perception: Critical Role of the Prefrontal Cortex. *The Journal of Cognitive Neuroscience*, 19, 993-2003.
- Saver, J. L., & Damasio, A. R. (1991). Preserved access and processing of social knowledge in a patient with acquired sociopathy due to ventromedial frontal damage. *Neuropsychologia*, 29, 1241-1249.
- Saxe, R. (2005). Against simulation: The argument from error. *Trends in Cognitive Sciences*, 9, 174-179.
- Saxe, R., Carey, S., & Kanwisher, N. (2004). Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annual Review of Psychology*, 55, 87-124.
- Saxe, R., Moran, J. M., Scholz, J., & Gabrieli, J. (2006). Overlapping and non-overlapping brain regions for theory of mind and self reflection in individual subjects. *Social Cognitive and Affective Neuroscience*, 1, 229-234.
- Schacter, D.L., & Buckner, R.L. (1998). Priming and the brain. *Neuron*, 20, 185-195.
- Schmitz, T. W., Kawahara-Baccus, T. N., & Johnson, S. C. (2004). Metacognitive evaluation, self-relevance, and the right prefrontal cortex. *NeuroImage*, 22, 941-947.
- Scholz, J., Triantafyllous, C., Whitfield-Gabrieli, S., Brown, E. N., & Saxe, R. (2009). Distinct regions of the right temporo-parietal junction are selective for theory of mind and exogenous attention. *PLoS ONE*, 4, 1-7
- Shamay-Tsoory, S. G., Tomer, R., Berger, B. D., & Aharon-Peretz, J. (2003). Characterization of empathy deficits following prefrontal brain damage: The role of the right ventromedial prefrontal cortex. *Journal of Cognitive Neuroscience*, 15, 324-337.
- Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, 303, 1157-62.
- Smith, E. R., & Semin, G. R. (2004). Socially situated cognition: Cognition in its social context. *Advances in Experimental Social Psychology*, 36, 53-117.
- Somerville, L. H., Heatherton, T. F., & Kelley, W. M. (2006). Anterior cingulate cortex responds differentially to expectancy violation and social rejection. *Nature Neuroscience*, 9, 1007-1008.
- Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review*, 99, 195-231.
- Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences*, 93, 13515-13522.
- Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences*, 12, 455-460.
- Turner, M. S., Simons, J. S., Gilbert, S. J., Frith, C. D., & Burgess, P. W. (2008). Distinct roles for lateral and medial rostral prefrontal cortex in source monitoring of perceived and imagined events. *Neuropsychologia*, 46, 1442-1453.

- Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2008). The neural substrates of in-group bias: A functional magnetic resonance imaging investigation. *Psychological Science, 11*, 1131-1139.
- van Reekum, C. M., Johnstone, T., Urry, H. L., Thurow, M. E., Schaefer, H. S., Alexander, A. L., & Davidson, R. J. (2007). Gaze fixations predict brain activation during the voluntary regulation of picture-induced negative affect. *Neuroimage, 36*, 1041-1055.
- Vuilleumier, P., Richardson, M. P., Armony, J. L., Driver, J., & Dolan, R. J. (2004). Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nature Neuroscience, 7*, 1271-1278.
- Wager, T. D., Rilling, J., Smith, E. E., Sokolik, A., Casey, K., Kosslyn, S. M., Davidson, R.J., Rose, R. M., Cohen, J. D. (2004) Placebo-induced changes in fMRI in the anticipation and experience of pain, *Science, 303*, 1162-1167.
- Wagner, A.D., Gabrieli, J.D.E., & Verfaellie, M. (1997). Dissociations between familiarity processes in explicit recognition and implicit perceptual memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 305-323.
- Walker, P., Silvert, L., Hewstone, M., & Nobre, A. (2008). Social contact and other-race face processing in the human brain. *Social Cognitive and Affective Neuroscience, 3*, 16-25.
- Whalen, P. J. (1998). Fear, vigilance, and ambiguity: initial neuroimaging studies of the human amygdala. *Current Directions in Psychological Science, 7*, 177-188.
- Wiese, H., Stahl, J., & Schweinberger, S. R. (2009). Configural processing of other-race faces is delayed but not decreased. *Biological Psychology, 81*, 103-109.
- Xu, X., Zuo, X., Wang, X., & Han, S. (2009). Do you feel my pain? Racial group membership modulates empathic neural responses. *The Journal of Neuroscience, 29*, 8525-8529.
- Yin, H. H. & Knowlton, B. J (2006). The role of the basal ganglia in habit formation. *Nature Reviews Neuroscience, 7*(6), 464-476
- Zhu, Y., Zhang, L., Fan, L., & Han, S. (2007). Neural basis of cultural influence on self-representation. *NeuroImage, 34*, 1310-1316.

Acknowledgments

Work on this article was supported by a grant from the National Science Foundation to David Amodio (BCS 0847350) and an NSF Graduate Research Fellowship to Kyle Ratner. We thank Jillian Swencions for her assistance in preparing this chapter.

Figure Captions

1. Medial view of the left hemisphere of the brain. mPFC = medial prefrontal cortex, OFC = orbital frontal cortex.
2. View of coronal slice through brain, with structures on the left side labeled. AMG = amygdala.
3. Lateral surface of the right hemisphere of the brain. dlPFC = dorsolateral prefrontal cortex, vlPFC = ventrolateral prefrontal cortex, pPFC = posterior prefrontal cortex, TPJ = temporo-parietal junction.
4. Schematic of information flow within the amygdala, illustrating connections between nuclei and intercalated cell masses. Sensory information enters via the lateral nucleus; signals out from the amygdala flow from the central nucleus (e.g., the fear response) and the basal nucleus (e.g., instrumental response). LA = lateral nucleus; ITC = intercalated cells; B = basal nucleus; CE = central nucleus.

Figure 1

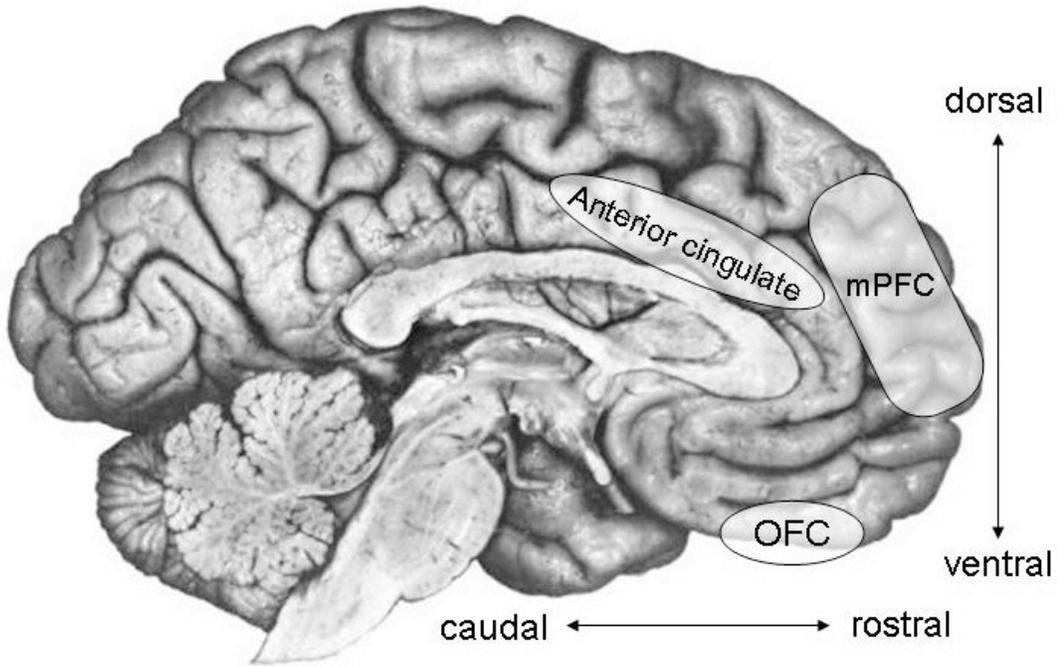


Figure 2

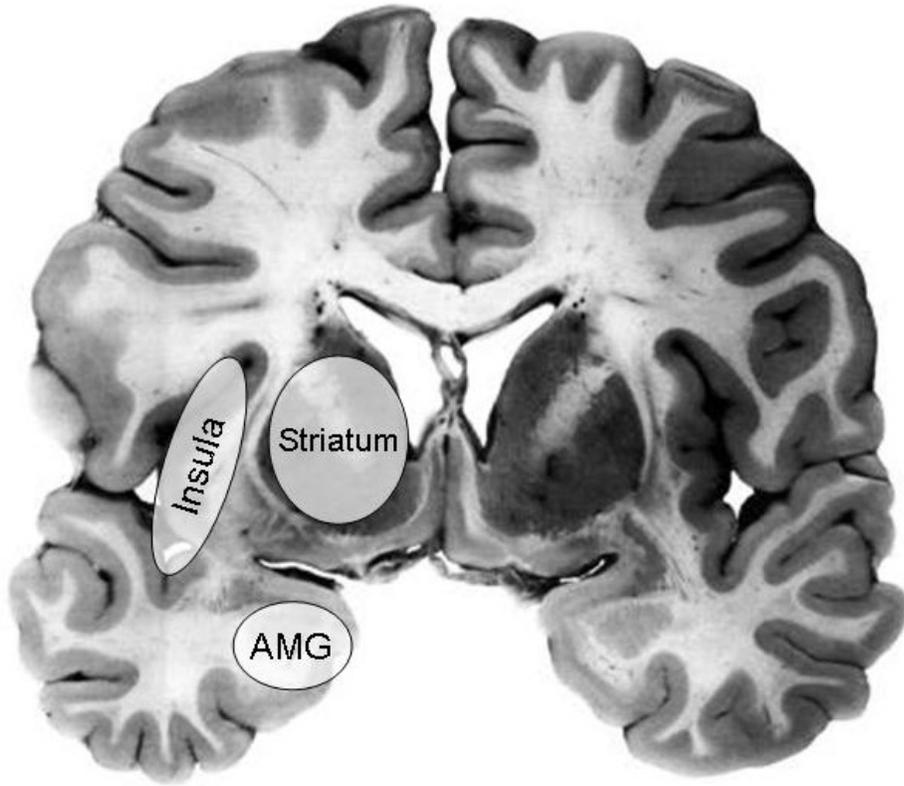


Figure 3

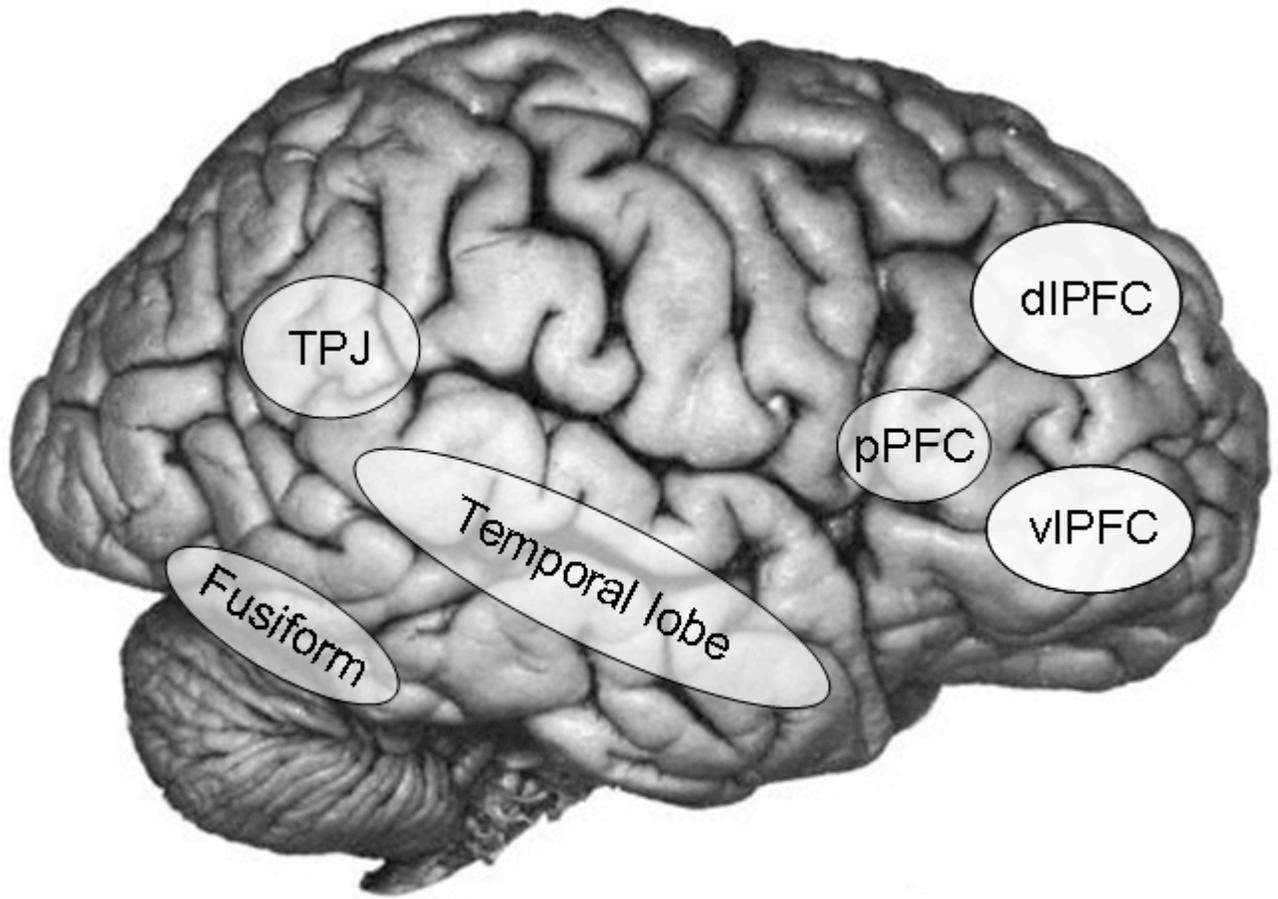


Figure 4

