Chapter 44

Self versus Others/Self-Regulation

ANNE C. KRENDEL AND TODD F. HEATHERTON

Humans are fundamentally social beings. Collectively, we crave social interactions and actively seek them out whenever possible (Baumeister & Leary, 1995). In fact, social interactions are so vital to our physical and mental well-being that being without them has grave consequences. For instance, social isolation can lead to severe depression and loneliness (Rubin & Mills, 1988) and is a major risk factor for mortality (Cacioppo, Hughes, Waite, Hawkley, & Thisted, 2006; House, Landis, & Umberson, 1988).

As humans, we expend a great deal of energy on maintaining and promoting our social groups. Many of the most popular modern technological advances serve the sole purpose of increasing the quantity of our social interactions. For instance, there has been an unprecedented increase in chat rooms, online gaming, and virtual reality worlds that center around promoting social interactions on the Internet. Meanwhile, increasing obsessions with personal digital assistants, e-mail, text messaging, and cell phones allow us to maintain constant social contact at all times, no matter where we are physically located.

Arguably, a primary cause for the amount of effort we exert on maintaining social interactions stems from the fact that human social interactions are remarkably complex. Determining when someone might have a crush on you, when your friends are no longer your friends, and the best time to ask your boss for a raise each involves highly developed levels of social awareness. Understanding the social rules required to successfully navigate through these myriad social interactions requires complex levels of cognitive processing. It is therefore not surprising that the brain has developed an intricate system of neural networks to support and facilitate social interactions. In this chapter, we describe the fundamental psychological components necessary for social behavior and we review what is currently known regarding known neural correlates. Before we begin our review, it is important to emphasize that the social brain is not located in one discrete location in the brain. Rather, it is comprised of multiple systems throughout the brain. Although we describe each of these systems separately, we do so with the following two caveats: First, each of the subcomponents of the social brain serve multiple functions, many of them overlapping; second, these subcomponents work together to give rise to the social brain as a whole. As we argue in this chapter, damage to any one system leads to profound social deficits. We begin by describing the theoretical basis of our model and then discuss relevant research.

THE SOCIAL BRAIN

Our overall approach follows a social brain sciences perspective that merges evolutionary theory, experimental social cognition, and neuroscience to elucidate the neural mechanisms that support social behavior (Adolphs, 2003; Heatherton, Macrae, & Kelley, 2004). Initial findings using neuroimaging have shown that unique neural substrates are associated with processing social information as compared to general semantic knowledge. For instance, Mitchell, Heatherton, and Macrae (2002) showed that when participants make semantic judgments about words that could either describe a person (e.g., assertive, fickle) or an object such as fruit (e.g., sundried, seedless), two separate networks were engaged. One system was activated when participants made judgments about objects (e.g., left inferior prefrontal cortex and left posterior parietal cortex), and a separate network was engaged when they made judgments about people (e.g., left temporal sulcus, medial prefrontal cortex, and fusiform gyrus; Figure 44.1). Similarly, Mason, Banfield, and Macrae (2004) found that when participants made judgments about whether an action (e.g., running,
determining discrete mental activities that come together to support the social brain (Turk, Heatherton, Macrae, Kelley, & Gazzaniga, 2003). Various cognitive, sensory, motor, somatosensory, and affective processes are essential to the successful navigation of social interactions, and these processes reflect the contribution of several cortical and subcortical regions.

Much of our approach relies on the idea that humans have evolved a fundamental need to belong, which encourages behavior that prevents people from being evicted from their social groups (Baumeister & Leary, 1995; Bowlby, 1969). According to the need-to-belong theory, the need for interpersonal attachments is a fundamental motive that has evolved for adaptive purposes. Over the course of human evolution, those who lived with others were more likely to survive and pass along their genes. Adults who were capable of developing long-term committed relationships were more likely to reproduce and also more likely to have their offspring survive to reproduce. Effective groups shared food, provided mates, and helped care for offspring. As such, human survival has long depended on living within groups; banishment from the group was effectively a death sentence. Baumeister and Leary (1995) argued that the need to belong is a basic motive that activates behavior and influences cognition and emotion, and that it leads to ill effects when not satisfied. Even today not belonging to a group increases a person’s risk for a number of adverse consequences, such as illnesses and premature death (see Cacioppo et al., 2000, 2006).

In essence, the social brain allows individuals to operate as effective group members, which allows them to maintain their status within the group as well as cooperate with other group members in service of the group’s survival needs. Such a system requires four components, each of which is likely to have a discrete neural signature. First, people need self-awareness—to be aware of their behavior so as to gauge it against societal or group norms. Both the psychologist William James and the sociologist George Herbert Mead differentiated between the self as the knower (“I”) and the self as the object that is known (“me”). In the sense of the knower, the self is the subject doing the thinking, feeling, and acting, which we will consider later as part of the executive self. In the sense of the objectified self, the self consists of the knowledge that people hold about themselves, as when they contemplate their best and worst qualities. The experience of self as the object of attention is the psychological state known as self-awareness, which allows people to reflect on their actions and understand the extent to which those actions match both personal values and beliefs as well as group standards. For instance, people who violate societal rules (i.e., by using more than their

Figure 44.1 Activation maps show brain areas to be more active during person trials than during object trials.

Note: Regions of modulation included A: the left temporal sulcus B: the dorsal and ventral mPFC C: the right fusiform gyrus, and D: the right parietal temporal occipital junction.

sitting, or biting) could be performed by a person or a dog, they showed dissociative patterns of neural activation. When making judgments about people, participants revealed a distinct pattern of activation in the prefrontal cortex (e.g., right middle frontal gyrus and medial prefrontal cortex) as compared to when they made judgments about dogs (e.g., occipital and parahippocampal gyri). There are countless other examples of discrete neural networks being recruited to process social as compared to nonsocial information, several of which we explore in more detail throughout this chapter.

From an evolutionary perspective, the brain is an organ that has evolved over millions of years to solve problems related to survival and reproduction. Those ancestors who were able to solve survival problems and adapt to their environments were most likely to reproduce and pass along their genes. Whether certain aspects of the social brain (i.e., a sense of self) truly are adaptive is open to some debate (Leary, 2005), although there is ample evidence that the ability to engage in social interactions provided considerable advantages over the course of evolution, such as facilitating communication and cooperation with group members (Sedikides & Skowronski, 1997). From the social brain sciences perspective, just as there are dedicated brain mechanisms for breathing, walking, and talking, the brain has evolved specialized mechanisms for processing information about the social world. It is important to emphasize here that we are not suggesting that there is a specific “social” module or region of the brain. Rather, psychological processes are distributed throughout the brain, with contributions from multiple subcomponents
fair share of resources by cheating their group mates) tend to feel ashamed of their behavior.

In order to experience social emotions, such as shame, people need to understand how people are reacting to their behavior so as to predict how others will respond to them (Heatherton & Krendl, 2007). In other words, they need the capacity to attribute mental states to others so that they are able to accurately interpret the beliefs and emotional states of those individuals. For instance, to feel guilty about hurting a loved one, people need to understand that other people have feelings. All social emotions related to empathy require the capacity to attribute specific mental states to others (Heatherton & Krendl, 2007).

Similarly, feeling shy is the belief that one is being evaluated by others (thereby giving rise to emotions such as embarrassment), which at its core means recognizing that other people make evaluative judgments. The ability to infer the mental states of others is commonly referred to as mentalizing, or having the capacity for theory of mind (ToM). ToM enables the ability to empathize and cooperate with others, interpret other people’s behavior, and even deceive others when necessary. This capacity is vital to ensure that people understand how others are viewing them in the group.

One value of having ToM is that it supports a third mechanism—threat detection—especially in complex situations. If humans have a fundamental need to belong, then there ought to be mechanisms for detecting inclusionary status (Leary, Tambor, Terdal, & Downs, 1995; Macdonald & Leary, 2005). Put another way, given the importance of group inclusion, humans need to be sensitive to signs that the group might exclude them. There is evidence that people feel anxious when they face exclusion from their social groups. Thus, feeling socially anxious or worrying about potential rejection should lead to heightened social sensitivity. Research has demonstrated that people who worry most about social evaluation (i.e., the shy and lonely) show enhanced memory for social information, are more empathetically accurate, and show heightened abilities to decode social information (Gardner, Pickett, & Brewer, 2000; Gardner, Pickett, Jefferies, & Knowles, 2005; Pickett, Gardner, & Knowles, 2004).

Not all threats, however, are related to social exclusion. Just as people naturally fear dangerous animals (i.e., poisonous snakes and spiders, tigers, and wolves), they also fear (and encounter) harm from other humans. Other group members can transmit disease, act carelessly to place bystanders at risk, waste or steal vital group resources, or poach one’s mate. Similarly, people from other groups also can be dangerous; over the course of human evolution, competition between small groups over scarce resources led to intergroup violence. Hence, there is also a need for mechanisms that detect threats from people from other groups. We argue that threatening people are accorded a special status that identifies them as potentially burdensome or dangerous—they are stigmatized. Thus, as research in social psychology has documented, humans quickly and efficiently categorize others based on information that is evolutionarily meaningful (Macrae & Bodenhausen, 2000).

Finally, there needs to be a mechanism for resolving discrepancies between self-knowledge and social expectations or norms, thereby motivating behavior to resolve any conflict that exists. This executive aspect of self (the "I" as knower) is responsible for ensuring that behaviors that might lead one to be expelled from the group are discouraged and, conversely, that behaviors that promote harmonious social relations within the group are encouraged. The control of one's own behavior is known as self-regulation, which is the process by which people change themselves, including their thoughts, their emotions and moods, their impulsive acts, and their performance at school or work. Although people can delay gratification, control appetites and impulses, and persevere in order to attain goals, many people have difficulties with self-regulation from time to time. Failures of self-regulation are implicated in most of the major problems of contemporary society, including addiction, obesity, risky sex, drunk driving, alcohol abuse, excessive gambling, spiraling consumer debt, marital infidelity, impulsive crimes, and school violence. Many of these behaviors threaten group inclusion; accordingly, understanding their neural basis is of considerable importance.

We do not contend that other psychological processes are unimportant for social functioning. Capacities such as language, memory, vision, along with motivational and basic emotional states, are generally important for functioning within the social group. However, they are not necessary for a person to be a good group member; the blind and deaf can contribute substantially to their groups. By contrast, people with fundamental disturbances in the primary components of self, ToM, threat detection, or self-regulation have fundamental and often specific impairments in social function. The literature is replete with case studies of individuals with brain injuries who suffer social impairments while having other capacities intact (e.g., Phineas Gage). Likewise, individuals with a disturbed sense of self often have interpersonal problems (Stuss & Alexander, 1994): Those who have difficulties with ToM (i.e., autistics) or impoverished empathy (i.e., psychopaths) are socially impaired, and those who fail to regulate their behavior are often ostracized and even imprisoned. Although space limitations preclude a theory discussion of social impairments, our contention is that many of them are due to fundamental problems with one of the core processes we have identified.
Thus, according to our model, the brain has evolved distinct mechanisms for knowing ourselves (self-awareness), knowing how others respond to us (which requires theory of mind), detecting threats from within our social group and beyond, and regulating our actions in order to avoid being ejected from our social groups. Together, these abilities form the foundation of the social brain. In the next few sections, we consider how neuroimaging can elucidate each of these discrete processes.

Self-Reflection and Awareness

The concept of self forms the foundation of the social brain. The self-concept consists of all that we know about ourselves, including things such as name, race, likes, dislikes, beliefs, values, and even whether we possess certain personality traits. According to Baumeister (1998), "the capacity of the human organism to be conscious of itself is a distinguishing feature and is vital to selfhood" (p. 683). Given that self-knowledge plays a critical role in understanding who we are, researchers have long debated whether the brain gives privileged status to information that is self-relevant or alternatively if information processed about the self is treated in the same manner as any other type of information (Bower & Logan, 1979; Klein & Kihlstrom, 1986; Klein & Loftus, 1988; Maki & McCaul, 1985; Markus, 1977; Rogers, Kuiper, & Kirker, 1977). This is the key issue underlying the question of whether self is "special" in any meaningful way. Gillihan and Farah (2005) argue that the majority of the patient and neuroimaging research does not provide sufficient support to conclude that self-relevant information is processed in any "special" way. In this section, we explore several studies that argue that self-relevant information is given a unique and "special" status in the brain.

Perhaps one of the most striking examples of the uniqueness of self is reflected in the self-relevant memory enhancement effect. A seminal study by Timothy Rogers and colleagues (1977) found a memory advantage for information encoded with reference to self. They found that asking people to make personal judgments on trait adjectives (e.g., "Are you mean?") produced significantly improved memory for the words than if the participants were asked to make semantic judgments (e.g., "Define the word mean"). This self-reference memory enhancement effect has been observed in many contexts, such as when people remember information processed with reference to self better than information processed with reference to other people (Symons & Johnson, 1997). The overall picture that emerges is that self-relevant information is especially memorable. Even people who can remember very little can often remember information that is self-relevant. For instance, patients who suffer from severe amnesia (resulting from brain injury, developmental disorders, or Alzheimer's disease) retain the ability to accurately describe whether specific traits are true of the self (Klein, 2004). Klein provides the example of patient K.C., who showed a preserved ability to accurately identify his "new" personality traits after becoming profoundly amnesic and undergoing a radical personality change following a motorcycle accident (Tulving, 1993). Even patients with Alzheimer's disease who suffer severe temporal disorientation and have difficulty recognizing their own family have shown evidence of self-knowledge. Patient K.R., for instance, suffered from profound Alzheimer's disease, yet she was still able to identify self-relevant personality traits accurately (Klein, Cosmides, & Costabile, 2003). But why is information about the self particularly memorable?

During the 1980s, research in social cognition examined the self-reference superiority effect in memory. Although it was undeniable that memory was better for self-relevant information, it was widely debated whether self-relevant information was supported by discrete cognitive systems (Rogers, Kuiper, & Kirker, 1977), or if superior memory could occur simply because people had greater knowledge about the self, which in turn could produce more elaborate encoding and, hence, better memory (Greenwald & Banaji, 1989; Klein & Kihlstrom, 1986). Neuroimaging techniques are exceptionally positioned to resolve the debate regarding the self-reference effect in memory. The first group to use brain imaging to examine this question was Fergus Craik and his colleagues at the University of Toronto. Using positron emission tomography (PET), Craik et al. (1999) examined brain activity while participants rated personality traits for the self or for a familiar other (in this case, the Canadian Prime Minister). These researchers did not replicate the self-relevant memory enhancement effect, but they did observe distinct activations for self-referential material in frontal regions, notably medial prefrontal cortex (mPFC) and areas of the right prefrontal cortex.

Using rapid event-related functional magnetic resonance imaging (fMRI), we asked participants to judge 270 trait adjectives in one of three ways: self ("Does the trait describe you?"); other ("Does the trait describe George Bush?"); and case ("Is the trait presented in uppercase letters?"); Kelley et al., 2002). The expected self-relevant enhancement effect was observed in this study, and a direct comparison of "self" trials to "other" trials revealed heightened activation in a number of different brain regions, most notably the mPFC (Figure 44.2). We conducted a subsequent study that showed that mPFC activity during the encoding of self-relevant words later predicted memory for these words (i.e., the greater the mPFC activity during the encoding of each item, the more likely that item was to be remembered;
Macrae, Moran, Heatherton, Banfield, & Kelley, 2004). Thus, we were able to demonstrate that mPFC contributes to the formation of self-relevant memories. This supports our contention that self-referential processing is functionally dissociable from general semantic processing. In other words, the brain has discrete neural substrates that give rise to the self.

The extent to which we include others in our self-concept has been a topic of particular interest for social psychologists. Theories of intimacy and personal relationships might suggest that the self-reference effect is affected by the closeness of a relationship with the other used as a target. Aron and colleagues define closeness as the extension of self into other and suggest that one’s cognitive processes about a close other develop in a way so as to include that person as part of the self (Aron & Aron, 1996; Aron, Aron, Tudor, & Nelson, 1991; Aron & Fraley, 1999). Consistent with this idea, the memorial advantage afforded to self-referenced material can be diminished or eliminated when the comparison target is an intimate other such as a parent, friend, or spouse (Bower & Gilligan, 1979; Keenan & Baillet, 1980).

To address this question, we conducted a study nearly identical to the Kelley et al. (2002) method, but this time we had people make judgments for their best friend rather than for George Bush (Heatherton et al., 2006). Although differences in recognition memory performance for self and intimate other judgments were modest, neural response differences in the mPFC were robust, with self showing much greater activity in mPFC than for best friend or case judgments, which did not differ from one another. These results indicate an mPFC response that is self-specific; that is, in the brain, judgments pertaining to the self were distinct from those made for friends.

Our findings diverge from others that have been reported in which mPFC activity was similar for self and intimate others (Ochsner et al., 2005; Schmitz, Kawahara-Baccus, & Johnson, 2004; Seger, Stone, & Keenan, 2004). Two methodological issues may account for this discrepancy. First, the three previous studies used block designs with fairly long intertrial intervals, whereas our study used an event-related design; the former may obscure single-trial events because brain activity is averaged across the entire block. It is possible that participants engaged in self-reflection between trials within the blocks, thereby mixing self-referential processing with judgments about the intimate others (e.g., such as recalling episodes in which the person acted in accordance with the trait during an interaction with the subject). We also used an unusually large number of research participants (N = 30) and therefore we had the power to detect differences between self and other; the finding of no difference between self and other in the previous studies might be due to power issues. Further research is necessary to resolve the importance of these methodological factors.

Considered together, the findings from our three studies support the idea that mPFC is involved in self-referential processing, and that the actions of this region support greater memory for material encoded with reference to self. These findings are also consistent with those obtained by other researchers (Gusnard, 2005). For instance, Gusnard, Akbudak, Shulman, and Raichle (2001) used a blocked-design fMRI
paradigm to examine judgments about affectively normed pictures and observed mPFC activity that was preferentially associated with self-referential judgments. Johnson et al. (2002) asked participants to respond to a series of questions that demanded access to either personal knowledge (e.g., “I have a quick temper”) or general semantic knowledge (e.g., “Ten seconds is more than a minute”). Their results revealed that self-reflective thought was accompanied by activity in anterior regions of mPFC. Cabeza et al. (2004) found heightened mPFC activation for episodic memory retrieval of autobiographical events. In the study, participants were presented with photographs that either they had taken around campus, or that someone else had taken. The participants showed heightened mPFC activity for photographs they themselves had taken. More recently, Mitchell, Banaji, and Macrae (2005) showed participants a series of faces and asked them to judge physical (i.e., how symmetrical the face appeared) or mental features (i.e., how pleased the people were to have their photographs taken). They found that the activity in mPFC was correlated with the extent to which participants judged the people in the photographs to be similar to them, but only for the mentalizing task.

The mPFC has also been implicated in autobiographical memory, an important component of self-awareness. Knowing yourself requires remembering events unique to your own past experiences. These memories play a large role in your understanding of who you are. Steinworth, Corkin, and Halgren (2006) asked participants to recall past autobiographical memories (e.g., “A birthday party: Who spilled wine on your pants?”) as well as semantic memories (e.g., “A cartoon figure: What is the color of the fur on Garfield?”). They found that participants engaged mPFC when recalling autobiographical memories, but not for semantic memories. Further, Addis, Wong, and Schacter (2006) found that ruminating on future biographical events (i.e., “Imagine your future child”) also elicited activation in the mPFC.

Thus, the imaging literature is quite clear regarding tasks that involve self-reflection; they activate mPFC (Gusnard, 2005). The view that mPFC plays a prominent role in self-referential processing is also supported by neuropsychological evidence of patients with frontal lobe injuries (Feinberg & Keenan, 2005; Stuss & Benson, 1984; Wheeler, Stuss, & Tulving, 1997). However, this is not to suggest that the mPFC is the only neural region that selectively responds to self-relevant information. For instance, Northoff and Bermehl (2004) suggest that the parietal cortex is vital to understanding the location of self in space (Vogeley & Fink, 2003) and the orbitomedial prefrontal cortex tags incoming information as self-relevant so it can be processed by the appropriate system (see also Schore, 2003). The mPFC, however, plays an important role in self-awareness and extensive literature has consistently implicated this region as being particularly important to processing self-relevant information.

Emerging evidence suggests that the same areas implicated in processing self-relevant information also appear to be tonically active when the brain is “at rest”; that is, not performing a cognitive task (Raichle et al., 2001). This so-called “default state” includes the same network of regions observed when participants perform self-relevant processing tasks—mPFC, precuneus, and posterior cingulate (Gusnard et al., 2001). This finding has led to the supposition that when people are “doing nothing,” the default state of the brain is to self-reflect (Gusnard et al., 2001; Kelley et al., 2002; Mason, Norton, Van Horn, Wegner, Grafion, & Macrae, 2007).

The converging evidence the studies described in this section suggests that mPFC plays a critical role in the social brain by giving rise to the execution and storage of self-relevant information. Self-awareness provides the ability us with the necessary information to understand our own social goals in the world. This necessitates being able to take the perspective of another. In the next section, we review the findings from emerging neuroimaging research attempting to isolate the neural mechanisms engaged in theory of mind.

Theory of Mind

The social brain requires more than just being aware of our own mental states and feelings. A vital component of the social brain is the ability to recognize the mental states of others so we can engage in deception, empathize and cooperate with others, and accurately interpret other people’s behavior (Gallagher & Frith, 2003). Our ability to infer the mental states of others is commonly referred to as theory of mind (ToM). The extent to which ToM is a uniquely human trait is highly contentious, and evidence on this point is mixed.

Primitive forms of apparent mentalizing have been recorded in the animal literature, but it is widely debated whether these studies demonstrate ToM or just learned behavior (for review, see Seyfarth & Cheney, 2003). For instance, research with baboons has shown that if a dominant female grunts to a subordinate following aggression, the subordinate’s behavior immediately changes (Cheney & Seyfarth, 1997). Based on this observation, the authors posit two possible explanations: First, the subordinate has recognized a change in the dominant’s attitude toward her (e.g., the dominant is trying to make the subordinate feel less anxious), so the subordinate baboon changes her behavior accordingly (an explanation that necessitates ToM). The second possible explanation for this behavioral
change is that it simply reflects a learned behavior—the subordinate baboon has learned through experience that this type of grunt typically leads to reduced aggression (an explanation that does not require ToM).

Of the extant animal literature, the most compelling evidence for ToM has been observed in chimpanzees (for review, see Seyfarth & Cheney, 2003). Unlike other non-human primates, chimpanzees are able to recognize intentional gestures such as pointing. Povinelli, Nelson, and Boysen (1990) found that when chimpanzees were given clues by two different experimenters as to where food was hidden, they would follow the hints of the experimenter whom they had observed hide the food, instead of the clues provided by the second experimenter who had waited outside the room while the food was hidden. The authors argued that the chimpanzees’ ability to correctly determine that the person they had seen hide the food would know its true location is clear evidence that the chimpanzees possess ToM. Additionally, chimpanzees follow the gaze of a human or other group member to a specific location, an action that, when observed in children, is believed to be evidence of ToM (Tomasello, Hare, & Agnetta, 1999). However, not all research with chimpanzees points to evidence that they have acquired ToM. Chimpanzees exhibit no preference between begging an experimenter who could plainly see them for food, or begging another experimenter whose face or eyes were covered, and therefore could not see them (Povinelli & Eddy, 1996).

Emerging research on ToM in humans has sought to identify the neural correlates that are selectively engaged during mentalizing tasks. These studies have consistently implicated a network of brain areas, including mPFC, posterior cingulate, and tempo-parietal junction, as the central components of ToM. Of central interest in this work has been the role of the mPFC in ToM tasks. As discussed in the previous section, the mPFC plays a central role in processing self-relevant information. It is therefore not surprising that the same region that supports our ability to determine our own mental states would be involved in our ability to infer the mental states of others.

Compelling evidence has emerged in the patient literature to support the assertion that the mPFC plays a central role in ToM. For instance, research with people who are either autistic or suffer from Asperger’s syndrome (which both have impairments in the ability to mentalize) has revealed that the deficits may result, at least in part, from deficiencies in the mPFC (Gallagher & Frith, 2003). Further, Stuss, Gallup, and Alexander (2001) found that patients with frontal lobe lesions (particularly to the right mPFC) were unable to detect deception, a task requiring ToM.

The advent of PET and fMRI has allowed researchers to examine the neural correlates engaged in ToM tasks in healthy participants. In one of the first attempts to identify brain regions involved in ToM, Fletcher and colleagues (1995) measured neural activity while participants made judgments (based on ToM reasoning) about the motivations of an actor’s behavior. For example, participants had to work out that an actor’s behavior (giving himself up to the police) was based on his assumption about the policeman’s beliefs (the policeman knew he had robbed a shop). Because the policeman’s beliefs were not made explicit in the story, mental state attribution (i.e., ToM reasoning) was required to perform the task. These researchers found that these mental state attributions were accompanied by a relative increase in activation in mPFC. Activation in the mPFC has also been observed in ToM paradigms that use pictures that require mental state attribution to be understood (Gallagher et al., 2000).1

mPFC activation has been observed even in contrived tasks in which participants are led believe they must infer the mental states of others. Gallagher, Jack, Roepstorff, and Frith (2002) had participants play a game of “rock, paper, scissors” while in the PET scanner. Participants were told during some blocks that they were playing against the experimenter, and in others that they were playing against the computer. In truth, they received randomly generated stimuli from the computer throughout the experiment. However, on trials during which participants believed they were playing the experimenter, they showed robust activation in the mPFC (suggesting they were using ToM to determine how the experimenter might play the next hand), whereas they did not show this activation when they believed they were playing the computer. Similarly, mPFC activation was observed on tasks in which cooperation was required among team members. McCabe, Houser, Ryan, Smith, and Trouard (2001) had participants compete in a trust game with either human or computer partners. The authors found that participants showed heightened activation in mPFC when they were cooperating with a human partner. They argue that such cooperation requires ToM because participants must be able to infer the mental states

---

1In several studies described in this chapter, changes in mPFC activity often appear to be decreases in activity from an arbitrary baseline. As discussed earlier in this chapter, mPFC is tonically active when the brain is at rest. In other words, when the brain is engaged in cognitive tasks, mPFC activity appears to decrease relative to resting state (which is measured by the arbitrary baseline) because its most active state occurs at rest (for a more detailed discussion of the default state and mPFC, please see Raichle et al., 2001). Thus, for the sake of clarity, our use the term "activations" in this chapter refers to changes in neural activity that are significantly greater in the experimental than control conditions.
of their partners in order to form mutual expectations of making cooperative choices.

Thus, the mPFC can be engaged when perceivers either are interacting with another human, or simply believe that they are. Further, the extent to which the perceiver can identify with an individual with whom he is interacting may also modulate the extent to which mPFC is engaged. Mitchell, Macrae, and Banaji (2006) had participants who were self-proclaimed liberals and conservatives make judgments about themselves, other liberals, or other conservatives (e.g., “Would this individual enjoy having a roommate from a different country?” “Would this individual drive a small car entirely for environmental reasons?”). The authors found that the activation in the ventral mPFC was greater for judgments made about individuals that participants perceived to be similar to themselves (i.e., participants who were identified as liberals by a measure of implicit political bias showed heightened vmPFC when making judgments about other liberals than when evaluating conservatives), whereas activation was greater in the dorsal mPFC for individuals that participants perceived as being dissimilar to themselves (e.g., liberals evaluating conservatives).

Although the extant literature on ToM has clearly emphasized the role of the mPFC in mentalizing tasks, emerging research has also implicated the tempo-parietal junction (TPJ) in inferring mental states. Saxe and Kanwisher (2003) observed heightened activation in this region when participants read stories that uniquely described the goals or beliefs of an individual (a task that requires ToM), as opposed to when participants read stories that simply described people in physical detail (a task that does not require ToM). In a later study, Saxe and Wexler (2005) had participants make judgments about the mental states of others who either came from similar (familiar) or dissimilar (foreign) backgrounds. In both cases, participants were given a short story about an individual with either a familiar (e.g., “Your friend is happily married”) or foreign (e.g., “Your friend belongs to a cult that promotes extra-marital affairs”) background. Each story then had a “normal” (e.g., “Your friend confided that he hoped his wife never cheated on him”) or “norm-violating” (e.g., “Your friend confided that he hoped his wife would have a relationship outside of marriage”) desire. The authors found that the right TPJ was recruited when perceivers were trying to reconcile incongruent information (e.g., a protagonist from a foreign background who expressed a “normal” mental state), which they argued requires greater mentalizing. The mPFC, however, did not discriminate between stories that would require differing levels of mentalizing. Thus, the authors argue that the right TPJ, not the mPFC, is uniquely engaged in the attribution of mental states.

To further explore the possibly diverging roles of the TPJ and mPFC, Saxe and Powell (2006) gave participants stories that described either the appearance of a person (e.g., “Joe was a heavy-set man, with a gut that fell over his belt. He was balding and combed his blonde hair over the top of his head. His face was pleasant, with large brown eyes”), a bodily sensation (e.g., “Sheila skipped breakfast because she was late for the train to her mother’s. By time she got off the train, she was starving. Her stomach was rumbling, and she could smell food everywhere”), or thoughts (e.g., “Nick knew that his sister’s flight from San Francisco was delayed 10 hours. Only one flight was delayed so much that night, so when he got to the airport, he knew that flight was hers”). The authors found that only the TPJ (bilaterally) and posterior cingulate were selectively recruited when participants read stories about a protagonist’s thoughts or beliefs (the only task that would require ToM), but not in the other two conditions. They observed similar patterns of mPFC activation in all three conditions. In other words, participants recruited mPFC equally when they were reading any story describing an individual, and not just in conditions in which they were making mental state inferences. The authors therefore argue that the TPJ is uniquely engaged in mentalizing tasks, whereas the mPFC may play a broader role in social processing.

Although the findings suggest that TPJ is involved in mentalizing, its precise role in theory of mind is widely debated. Gobbinii, Konarack, Bryan, Montgomery, and Haxby (2007) presented participants with false belief stories and pictures of geometric shapes moving in a socially relevant manner: two tasks that both require mentalizing. They only observed TPJ activation in the story condition, not in the social animation condition. Based on these findings, the authors argue that TPJ may play a role in interpreting “covert mental states” to help predict future behavior, but that it is not involved in interpreting beliefs and actions based on perceived actions. In an attempt to dissociate the recruitment of TPJ from mPFC in theory of mind tasks, Ciaramidaro and colleagues (2007) presented participants with cartoons depicting unique actions with disparate goals: private intentions (i.e., fixing a broken light bulb to read), prospective social intentions (i.e., observing a single person prepare dinner for someone else, reflecting a social intention to engage in a social interaction), and communicative intention (i.e., observing person A obtain a glass of water for person B after being asked to do so). The authors found that the right TPJ was active for all three conditions relative to control scenarios, whereas mPFC was only active for the social intention conditions, and not the private intentions. Importantly, the authors found a functional dissociation within the TPJ in that right TPJ in engaged in processing all intentions, but left TPJ is only engaged in processing discrete social intentions (communicative intentions). Together, these findings suggest that the neural network engaged during theory of mind tasks may be more complex than previously thought.
Although the studies discussed in this section reveal conflicting evidence for identifying the central mechanisms that give rise to ToM, one point remains clear: The brain has a specialized network engaged in mentalizing tasks. This key ingredient to the social brain allows us to engage in complex social interactions involving cooperation, deception, and empathy. ToM provides us a specialized tool by which to detect threats from within our social group. For instance, the ability to understand when you have committed a social error, and then determine how to overcome that social error, requires ToM. Successfully detecting threats from the environment—both from within the social group and beyond—is an important aspect of the social brain. We explore the unique mechanisms dedicated to threat detection in the next section.

**Threat Detection**

An important aspect of the social brain is the ability to detect threats in the environment. These threats can be threats to our physical self (i.e., quickly detecting when a predator is pursuing us or recognizing individuals who may threaten our social group or resources) or they may be threats to our psychological self (i.e., threats that affect our status within our social group via social rejection or social exclusion). It is apparent that the brain has developed an efficient system to respond to threats from the environment. For instance, the superior temporal sulcus is uniquely sensitive to detecting biological motion (i.e., movement of the eyes, mouth, hands, and body; Allison, Puce, & McCarthy, 2000; Grezes et al., 2001). However, the amygdala is central to this threat detection system (for review, see Whalen, 1998). The amygdala automatically detects potentially aversive stimuli in the environment, sometimes causing us to jump away from an object that resembles a threat (i.e., a curved branch in the forest that we mistake for a snake) even without knowing what it is.

Research with nonhuman primates has shown that amygdala lesions impair appropriate fear responses to novel stimuli (Amaral, 2002). For instance, Amaral showed that primates with amygdala lesions approach and play with a toy snake, while primates with intact amygdala waver from the toy. Evidence from patient research with humans suggests that the amygdala damage impairs patients’ ability to accurately identify the arousal associated with negative stimuli. Intriguingly, however, patients with amygdala damage are able to accurately identify the valence of positive and negative stimuli (Bernston, Bechara, Damasio, Tranel, & Cacioppo, 2007).

The amygdala may also play an important role in social threats. The amygdala shows heightened activation in response to fearful faces (Whalen, 1998). Research has shown that people are remarkably adept at recognizing fearful faces (Schubo, Gendolla, Meinecke, & Abele, 2006), and this ability is pervasive over the life span (Mather & Knight, 2006). Chiu, Ambady, and Deldin (2004) demonstrated that high-prejudiced White individuals more quickly evaluated angry Black faces as compared to happy Black faces. In other words, high-prejudice participants in this study could more quickly detect and evaluate an out-group member when that individual conveyed threat (i.e., via an angry facial expression). Norris, Chen, Zhu, Small, and Cacioppo (2004) demonstrated that images with social and emotional content have an additive effect on the amygdala. They showed participants images that contained (a) social information only, (b) emotional information only, and (c) social and emotional information together along with neutral controls. They found the amygdala activation was significantly stronger in response to the pictures that conveyed both social and emotional information as compared to all other conditions. A similar finding emerged from a study by Ito and Cacioppo (2000) using event-related potentials. They found that negative images with social information (e.g., negative pictures that included people) received heightened processing as compared to images that were void of social information. Together these findings suggest that negative social information elicits heightened neural activation as compared to negative nonsocial information, particularly in the amygdala.

Given the important role of the amygdala in detecting threats in the environment, it is not surprising that the amygdala also is largely involved in our ability to detect social threats. Emerging research on threat detection from outgroup members has focused primarily on the role of the amygdala in automatically detecting threats from the environment (for review, see Eberhardt, 2005). We next explore these findings in more detail.

**Threats from Outgroups**

There is a ubiquitous tendency among humans to stigmatize outgroup members (Kurzban & Leary, 2001). Stigma refers to an attribute that renders individuals “devalued, spoiled or flawed in the eyes of others” (Crocker, Major, & Steele, 1998). Broadly defined, common stigmas include people of different races, people who are physically disabled (e.g., paraplegics), or people with mental disabilities (i.e., schizophrenics; Goffman, 1963). Extensive research has revealed that stigmas automatically elicit powerful and often negative responses from perceivers (for review, see Crocker et al., 1998).

Kurzban and Leary (2001) argued that outgroup members are stigmatized as a way of helping social groups protect themselves from outside threats. Emerging neuroscience research has focused primarily on the stigma of race and has revealed that the amygdala plays a central role...
in perceiving other races (Cunningham et al., 2004; Hart et al., 2000; Lieberman, Harriri, Jarche, Eisenberger, & Bookheimer, 2005; Phelps et al., 2000; Richeson et al., 2003). However, it remains an open question what role the amygdala is playing in perceiving outgroup members. Specifically, is the amygdala responding because we experience negative emotions upon seeing outgroup members, or is the amygdala simply responding to the novelty of the outgroup member?

Phelps and colleagues (2000) found a strong positive correlation between heightened amygdala activation of White participants to their anti-Black bias (as measured by the Implicit Association Test; Greenwald, McGhee, & Schwartz, 1998), which may suggest that the amygdala’s role in perceiving race is more emotion based. Conversely, Hart and colleagues (2000) demonstrated the amygdala activation of White participants habituated to the presentation of White, but not Black, faces over time, suggesting that the amygdala simply responds to the novelty of the social stimulus. Further suggesting a more subtle role of the amygdala in processing outgroup members, Wheeler and Fiske (2005) found that the types of judgments that participants make about faces affects amygdala activity. For instance, when White participants were asked to evaluate Black faces, amygdala activity was observed only when the target was socially categorized (e.g., “Is this individual over 21 years old?”), and not when participants were asked to individuate the target (“Would this individual like this vegetable?”).

The amygdala is only one of several neural areas engaged during the evaluation of an outgroup member. Emerging research from neuroimaging has revealed that areas of the prefrontal cortex involved in cognitive control are also engaged in these tasks. For instance, Cunningham and colleagues (2004) showed that the amygdala responded to pictures of Black faces when presented very quickly (30 ms). However, when the faces were presented longer (525 ms), the amygdala response was dampened, and instead increased activation was observed in the prefrontal cortex. The authors argue that the heightened activation in the prefrontal cortex may have been inhibiting the automatic response elicited by the amygdala.

Richeson and colleagues (2003) also found that White participants engage prefrontal control mechanisms (i.e., dorsolateral prefrontal cortex and anterior cingulate cortex) in response to viewing Black faces (Figure 44.3). However, they found that the activation of these areas was positively correlated with anti-Black bias. In other words, they found that White individuals with greater anti-Black bias recruit some of these cognitive control areas to a greater extent than White individuals with less anti-Black bias. They argue that this heightened activation results from the higher bias Whites’ attempts to mask their prejudice (see also Richeson & Shelton, 2003).

Figure 44.3 Statistical activation maps of Black faces versus White faces contrast, showing regions in right and left middle frontal gyri, as well as right anterior cingulate cortex.

Neuroscience research on nonrace stigmas is scant, but two studies examine the neural mechanisms engaged in perceiving nonrace stigmas (Harris & Fiske, 2006; Krendl, Macrae, Kelley, Fugelsang, & Heatherton, 2006). Krendl and colleagues (2006) examined the neural mechanisms engaged during explicit (conscious) and implicit (unconscious) evaluations of socially stigmatized individuals (e.g., people who are unattractive or who have numerous facial piercings). We demonstrated that individuals also engaged inhibitory mechanisms in response to viewing socially stigmatized targets, even when the perceivers were unaware that they were evaluating the targets. We also showed that stigmas that were explicitly rated as being more aversive elicited heightened amygdala response (Figure 44.4).

Harris and Fiske (2006) examined nonrace stigmas from the perspective of mentalizing. Specifically, they sought to identify whether activation in the medial prefrontal cortex (mPFC) is modulated by the type of stigma group being evaluated. They found that the mPFC was less active than when participants evaluated “extreme outgroup members” (homeless people, drug addicts) as compared to other stigmatized groups (e.g., older adults, disabled people). However, in a subsequent study, they found that the activity of mPFC was further modulated by the type of judgment individuals made about the stigmatized individuals (Harris & Fiske, 2007). For instance, making judgments about whether individuals would like certain vegetables elicited heightened mPFC activity in response to both extreme outgroup members and other stigmatized individuals as compared to making age judgments about the individuals.

Considered as a whole, these results revealed two important points: (1) outgroup members appear to automatically activate aversive responses in perceivers, particularly when the perceiver is high in prejudice; and (2) perceivers...
must successfully engage cognitive control mechanisms to inhibit aversive responses to stigmatized outgroup members. Together these studies reveal a unique social function for the amygdala and prefrontal cortex in detecting threats from outgroup members. In the next section, we explore the manner in which the social brain can detect threats from the ingroup.

**Threat from Ingroups**

Ingroup threats (i.e., social rejection or isolation) pose arguably the most potential harm to group members because they can result in ejection from one's social group. There are several possible causes for ingroup threat. For instance, Kurzban and Leary (2001) argue that individuals will be ostracized from the social group if they endanger the group (i.e., they have a disease that poses a risk to the group) or if they do not contribute to the group (i.e., they are missing a limb and thereby unable to help gather resources). Such individuals, according to Kurzban and Leary, are stigmatized and socially isolated from the group. Further, they argue that individuals who directly violate the rules of the group are subjected to social isolation. In other words, people who steal from the group or intentionally harm other group members will be ejected from the group. Thus, to maintain group membership, one must adhere to the social norms of the group. It is on this final point that we focus most of our attention in this section.

Violations of social norms are met with harsh punishment or ostracism, even among nonhuman primates. Rhesus macaques, for instance, will unleash significantly more aggression on group members who find food and do not share it with their cohort (Hauser, 1992). For humans, such deviations from social group norms may result in social rejection, or even ejection from one's social group (Kurzban & Leary, 2001), a punishment most individuals want to avoid at all costs. Thus, the social brain has evolved an extensive network to detect violations of social group norms, thereby serving to protect our membership in the group.

The ability to detect ingroup social threats appears to rely, at least in part, on the anterior cingulate cortex (ACC), a region that has been implicated in conflict resolution (Kerns et al., 2004). The ACC has been implicated in responding to social interactions that provide conflicting social feedback (Eisenberger, Lieberman, & Williams, 2003). Somerville, Heatherton, and Kelley (2006) provided subjects with false feedback that was either negative or positive (rejection or not), and also that was incongruent or congruent with their expectations (expectancy violation or not). Results revealed a double dissociation between dorsal and
ventral ACC regions. The dorsal ACC (dACC) was uniquely sensitive to expectancy violations, and ventral (vACC) was uniquely sensitive to social feedback (Figure 44.5), with significantly greater response to negative feedback than positive feedback, irrespective of expectancy violations.

Similarly, Krendl, Richeson, Kelley, and Heatherton (2008) found vACC activation during a social threat task. We conducted an fMRI study in which women were reminded of gender stereotypes about math ability while they were completing difficult math problems. Women showed an increase in vACC activity while performing difficult math problems after a social threat was induced (reminding them of gender stereotypes), whereas in the absence of social threat, women instead showed heightened activation over time in regions associated with math learning, and no change in vACC activation (Figure 44.6). Not surprisingly, women who were threatened exhibited a decrease in math performance over time whereas women who were not threatened improved in performance over time. Based on these findings, we conclude that the vACC is engaged in social and emotional processing.

Perhaps one of the most immediate sources of ingroup threats stems from violating social norms. In order to protect against such threats, the social brain has developed an intricate network of aptly named social emotions to warn us when we have violated social norms. Social emotions are complex emotions (e.g., admiration, jealousy, envy, irritation, flirtatiousness) that promote long-term social relationships and interactions. These emotions are critical in ensuring that we adhere to social norms in social interactions.

Powerful social emotions commonly referred to as moral emotions are engaged to identify when we have acted inappropriately and violated a social norm. These emotions include guilt, pity, embarrassment, shame, pride, awe, contempt, indignation, moral disgust, and gratitude (Moll, de Oliveira-Souza, Zahn, & Grafman, 2008). Their purpose is to elicit negative reactions that will (hopefully) prevent us from committing the violation in the future. Emerging neuroimaging research has begun to identify the unique neural mechanisms that are selectively engaged to process social and moral emotions. Not surprisingly, social and moral emotions engage many of the same structures activated by basic emotions (e.g., anger, fear), but they also selectively engage neural networks involved in assessing affect (e.g., amygdala, orbitofrontal cortex), as well as those regions that are involved in building cognitive schemas about the social world (left prefrontal cortex, right parietal cortex, anterior and posterior cingulate cortex; Adolphs, 2003).

The role of the amygdala in processing social emotions is a recent and novel finding that has emerged from patient and neuroimaging research. For instance, Adolphs, Baron-Cohen, and Tanel (2002) presented facial expressions of social emotions (arrogant, guilt, admiration, flirtatiousness) to patients with amygdala damage. Patients with unilateral or bilateral amygdala damage were impaired when recognizing those specific emotions; moreover, they were
more impaired at recognizing social emotions than basic
emotions.

Ruby and Decety (2004) conducted a PET study in
which participants were asked to choose the appro-
priate reaction (from varying perspectives) to sentences that
represented different social emotions (embarrassment,
pride, shame, guilt, admiration, irritation), or nonsocial
emotions and nonemotional sentences. Results revealed
heightened amygdala activation during the processing of
all social emotions, regardless of the perspective taken
during the task. However, it is important to note that the
authors do not dissociate between types of social emotions
in the task. Thus, it is unclear whether the amygdala activ-
ation observed was driven by a specific emotion.

Berthoz, Grezes, Armony, Passingham, and Dolan
(2006) conducted an fMRI study to examine intentional
violations of social norms. In the study, participants were
presented with stories (e.g., “You are invited for a Japanese
dinner at your friend’s house”), and one of three endings:
one was descriptive of normal social behavior (e.g., “You
have a bite of the first course, like it, and congratulate
your friend for her good cooking”); one that described
an embarrassing conclusion (e.g., “You have a bite of the
first course, you choke and spit out the food while you
are coughing”); or one in which the protagonist violated
a social norm (e.g., “You have a bite of the first course,
but do not like it and spit the food back into your plate”).
Participants evaluated the statements from their own
perspective, or someone else’s. When taking their own
perspective, participants showed greater amygdala activation
in response to intentional violations of social norms.

However, the amygdala is only one part of the neural
network engaged in perceiving social and moral emotions.
Shin and colleagues (2000) used PET to isolate the neural
correlates of guilt, a moral emotion. Prior to the PET scan,
participants were asked to provide written accounts of three
distinct events: one involving the most guilt the participant had ever experienced, and two additional events involving no clear emotion. During the scan, a tape-recording of the autobiographical events was presented aurally to participants, and they were asked to reexperience the event to which they were listening. Direct comparisons between guilt-induction conditions versus neutral elicited heightened activation in anterior cingulate gyrus and left anterior insula.

A recent fMRI study on embarrassment, also a moral emotion, found that the anterior cingulate cortex was activated when it was clear that a social norm had been violated. Berthoz, Armony, Blair, and Dolan (2002) used fMRI to investigate the neural systems supporting embarrassing situations and violations of social norms using the same paradigm described previously. Violation of social norms yielded greater activation in the anterior cingulate than embarrassment. Both violation of social norms and embarrassment led to greater activation in the medial prefrontal cortex (such as observed in self-awareness and ToM tasks) and the left orbitofrontal cortex.

mPFC activation was also observed in an fMRI study by Takahashi and colleagues (2004) when they compared guilt and embarrassment to neutral emotions. In the study, participants were shown sentences that had been previously rated as inducing guilt (e.g., “I shoplifted a dress from the store”), embarrassment (e.g., “I soiled my underwear”), or no emotion (neutral; e.g., “I washed my clothes”). Both guilt and embarrassment elicited mPFC activation, but direct comparisons between the two revealed heightened activation for mPFC in the guilt condition as compared to embarrassment. The involvement of mPFC in social and moral emotions suggests that these emotions may uniquely engage some form of mentalizing in order to be effective.

One possible explanation for why social and moral emotions may engage an extensive network of activation is that simply experiencing social emotions does not make them effective. Instead, social emotions must inform the perceiver of what social norms were followed or violated and provide either a reward or punishment (respectively) to encourage or deter future reoccurrences. However, in order to be effective, the perceiver of the social emotion must possess self-awareness and ToM. For instance, when we commit a social error during a social interaction, we may feel embarrassed. However, to recognize that we are embarrassed because of the social error, we must have self-awareness. Conversely, we may recognize the social error first by realizing that the person to whom our comment was directed is upset by our remark. This recognition may then lead to a feeling of embarrassment. In this case, we would assess violations of social norms by making an inference about the mental state of someone else, which requires ToM.

Possessing self-awareness and ToM only allow us to understand when a social norm has been violated and does little to prevent us from committing the errors in the future. Here is where self-regulation plays a vital role in the social brain. Self-regulation allows us to control our behavior to ensure that we do not violate social norms. We next explore how the social brain gives rise to self-regulation.

Self-Regulation

A unique aspect of human behavior is the ability to regulate and control thoughts and actions, an ability commonly referred to as self-regulation. Self-regulation allows us to make plans, choose from alternatives, focus attention on pursuit of goals, inhibit competing thoughts, and regulate social behavior (Baumeister, Heatherton, & Tice, 1994; Baumeister & Vohs, 2004; Metcalfe & Mischel, 1999; Wegner, 1994). Extensive evidence from neuroimaging and patient research demonstrates that the prefrontal cortex is imperative in successfully engaging self-regulatory processes, as befitting its label as “chief executive” of the brain (Goldberg, 2001). The vital role of the prefrontal cortex in self-regulation was famously observed in the case of Phineas Gage, who suffered profound frontal lobe damage when a railroad spike misfired into his head. Formerly described by friends as dependable, polite, and hardworking, Gage became capricious and volatile after the accident. Gage’s failure to regulate his social behavior after his injury was among the first lines of evidence that the prefrontal cortex supports the inhibitory mechanisms necessary to regulate behavior.

Since Gage’s accident, abundant patient and neuroimaging research has identified discrete brain regions within prefrontal cortex that are critical for self-regulation (for review, see Banfield, Wyland, Macrae, Münte, & Heatherton, 2004), primarily the dorsolateral prefrontal cortex (DLPFC; involved in modulating of cognitive control), the orbitofrontal cortex (OFC; involved in integrating cognitive and affective information), and the anterior cingulate cortex (ACC; involved in conflict resolution).

The DLPFC has been associated with planning, novelty processing, choice, the control of memory, and working memory and language function (see D’Esposito et al., 1995; Dronkers, Redfern, & Knight, 2000; Fuster, Brodner, & Kroger, 2000; Goldman-Rakic, 1987). Damage to this area often results in patients’ experiencing an inability to inhibit certain behaviors (Pandya & Barnes, 1987). Damage to the OFC, which controls our behavioral and emotional output and how we interact with others (Dolan, 1999), often results in striking, and sometimes aggressive, behavioral changes (e.g., Rolls, Hornak, Wade, & McGrath, 1994). Damage to the OFC usually results in personality changes.
such as indifference, impaired social judgment and responsiveness, poor self-regulation, lack of impulse control, and poor judgment and insight (Damasio, 1994; Stone, Baron-Cohen, & Knight, 1998; Stuss & Alexander, 2000). Patients with OFC damage often cannot inhibit desires for instant gratification and thus may commit thefts or sexually aggressive behavior (Blumer & Benson, 1975; Grafman et al., 1996).

The ACC is essential for initiating actions, evaluating conflicts, and also inhibiting prepotent responses, processes heavily involved in self-regulation (Kerns et al., 2004). The ACC is functionally dissociated into the dorsal ACC that evaluates cognitive conflict, and the ventral ACC that evaluates emotional conflict (Bush, Luu, & Posner, 2000). The ACC is often engaged whenever any kind of “supervisory input” is required (Badgaiyan & Posner, 1998). In fact, it is widely accepted that the ACC is somehow involved in evaluating the degree and nature of conflict, whereas other parts of the brain (particularly the PFC) may be involved in resolving the conflict itself (Botvinick, Cohen, & Carter, 2004; Cohen, Botvinick, & Carter, 2000; Kerns et al., 2004).

Emerging neuroimaging research has sought to more clearly identify the neural structures in self-regulation by examining the structures engaged in emotion and cognitive regulation. Ochsner, Bunge, Gross, and Gabrieli (2002) showed participants highly negative pictures and instructed them either to “attend” (study the picture and be aware of, but not try to alter, their feelings toward it) or “reappraise” (reinterpret the picture in such a way that it would no longer elicit a negative response) the photograph. The authors found that reappraising the photographs led to decreased subjective negative affect, and this was reflected in a reduction of activity in the amygdala (a region implicated in processing fear) and OFC, and increased activation in the lateral and medial prefrontal cortex, as well as in the anterior cingulate cortex.

In a later study, Ochsner and colleagues (2004) instructed participants to increase negative affect toward the image (by imagining themselves or a loved one as the central figure in a highly negative photograph) or decrease their negative affect to the photograph (by psychologically distancing themselves from it). Here, the authors found that extensive networks in the prefrontal cortex (left prefrontal cortex, dorsal anterior cingulate, and dorsal mPFC) were engaged when participants were using self-regulatory processes either to increase or decrease their affective response to the photographs. They observed that enhancing negative emotions engaged primarily left-lateralized prefrontal regions, whereas suppressing negative affect engaged bilateral prefrontal regions. Importantly, they also observed that activity in the amygdala decreased when participants actively decreased their negative affect to the picture, and increased when they increased their negative affect.

Self-regulation research has not been limited to modulating negative affect. Kim and Hamann (2007) observed increased activation in the left prefrontal cortex when participants were asked to increase either positive or negative affective responses to stimuli, and predominantly bilateral prefrontal activity in response to suppressing positive or negative affect. Importantly, the dorsal mPFC and the OFC were engaged for regulating both positive and negative emotions. Activation of the amygdala increased during the increase condition for both positive and negative pictures, and decreased in the suppress conditions in response to both positive and negative stimuli as well. Beareegard, Levesque, and Bourgouin (2001) had male participants view erotic films clips while undergoing fMRI. Participants were instructed either to allow themselves to become aroused during the clips, or to suppress any arousal they might be feeling. The authors found that suppressing arousal resulted in heightened activation in the right superior prefrontal and right anterior cingulate cortices with no accompanying activation in the limbic areas (e.g., amygdala, hypothalamus) that were active during the arousal condition.

Together, these findings have had important implications in patient populations. For instance, emerging research on patients with severe depression has revealed that their prefrontal cortex is unable to suppress amygdala activation in emotion regulation tasks (Johnstone, van Reekum, Urry, Kalin, & Davidson, 2007). When depressed patients are asked to suppress negative affect to highly aversive pictures, depressed patients show a positive correlation between the amygdala and vmPFC, whereas controls demonstrate a negative correlation between the two. In other words, the more controls engage vmPFC to suppress negative affect, the greater decrease is observed in activity in the amygdala. However, the more depressed patients try to suppress negative affect, the greater the activity in their amygdala.

Another important form of self-regulation that is critical for daily living is mental control. Successfully controlling the contents of consciousness is a difficult task—worries intrude when people least desire them and it is not uncommon for the mind to wander when people should be focused on a particular task or objective. Research by Wegner (1994) demonstrated that trying to suppress a particular thought can paradoxically lead to an increase in the very thought one is attempting to suppress. One open question is whether cognitive control of thoughts and actions involves similar component processes and therefore recruits common brain regions. If so, one might expect to observe ACC activity during attempts to control thoughts. To address this issue, we conducted an fMRI study of attempted
thought suppression (Wyland, Kelley, Macrae, Gordon, & Heatherton, 2003). Participants were each asked to provide a personally relevant thought that was especially salient to them at that moment (e.g., “study for an exam” or “a phone call with a distant girlfriend”). During the scan, they were given visual prompts that instructed them to: (a) suppress the personally relevant thought they had generated before the task (“suppress”), (b) think about nothing (“clear”), or (c) think about anything (“free thought”). Both the “suppress” and “clear” conditions required a form of thought suppression. To dissociate these two processes, no overt behavioral response was collected (e.g., pushing a button to index thought intrusions) as such a requirement contaminates thought suppression with response generation. Moreover, we were not concerned with failures of cognitive control per se, but rather the ongoing process of mental regulation.

The results indicated that the brain regions previously implicated in the suppression of overt behavior were also active during attempts to control the emergence of unwanted thoughts. Specifically, we found greater activation in the ACC for the “suppress” condition than for the “free thought” condition. When the “clear” and “free thought” conditions were compared, a more diverse pattern of neural activation was observed. Specifically, greater activation was observed in the anterior cingulate, left inferior frontal cortex, right insula, right parietal cortex, and right medial frontal cortex in the “clear” as compared to the “free thought” condition. The greater activity for the “clear” condition may have occurred because it is more difficult to suppress all thoughts than to suppress a specific thought. Our participants reported having a great deal of difficulty with this condition (the interested reader should go ahead and try this; it is nearly impossible). As previously demonstrated by Wegner and his colleagues (1989), suppressing a specific thought can be achieved relatively easily by thinking of something else, especially if that other thought captures attention.

Because we are also interested in examining failures of mental control, we conducted a second study in which participants were asked to suppress the specific thought of white bears (Mitchell, Heatherton, Kelley, Wyland, and Macrae, 2006). In this study, participants were scanned while alternatively trying to suppress the thought of a white bear or freely thinking about anything that came to mind; in either case, they pressed a button to indicate a white bear thought (we also had a third condition in between blocks that required participants to press a key when a yellow light appeared in order to control for simple motor movement). We found that right dorsolateral prefrontal cortex showed a sustained response when participants were attempting to suppress thoughts (i.e., the tonic state), whereas the ACC was activated by the intrusion of forbidden thoughts (i.e., transient events). This pattern of results is consistent with neural models of cognitive control that emphasize the interplay between PFC and ACC in cognitive control (Kerns et al., 2004) as well as with the Wegner’s (1994) model of mental control. The use of imaging is well suited to contribute to our ability to examine theoretical models of self-regulation, both in emotion and mental regulation tasks.

SUMMARY

In this chapter, we described how neuroimaging has informed our knowledge of the unique components of the social brain. Our research has identified a number of frontal lobe regions that appear to subserve important human talents, such as the ability to introspect, evaluate ourselves and others, detect social threats around us, and to purposefully modify our thoughts and behaviors in the pursuit of goals. Knowing where in the brain something happens doesn’t by itself tell you very much. But, knowing that there are consistent patterns of brain activation associated with specific psychological tasks provides evidence that the two are connected, and provides an opportunity to identify component processes that might be important for a full understanding of mental constructs. We believe that a social brain sciences approach will be useful for understanding the nature of the social brain.

Now that we have identified various regions of the brain that comprise the social brain, one next logical step is to try to identify the specific role of these regions. Many of the regions discussed in this chapter (e.g., the amygdala, ACC, DLPCF) have been implicated in nonsocial tasks. It would therefore be misleading to suggest that these regions are solely “social brain areas.” More than likely, these regions have a broader role in the brain (e.g., threat detection, conflict resolution) that renders them useful both in certain cognitive and social tasks.

However, the role of the mPFC in the social brain is particularly puzzling because it is robustly activated during self-relevant and theory of mind tasks, but otherwise it is inactive during most cognitive tasks (Raichle et al., 2001). It is thus thought that the mPFC is engaged when the brain is “inactive,” suggesting perhaps that the “default state” of the brain is introspection (Raichle et al., 2001). One possible explanation for this is that the deactivations observed in mPFC during cognitive tasks are due to the fact that available neural resources are required to perform the task at hand, and therefore fewer are available for inward reflection, thus causing a decline in activity in the mPFC (Gusnard, 2005). Another possibility is that mPFC operates by binding together various physical experiences and
cognitive operations that have implications for self. The prefrontal cortex receives input from all sensory modalities, and is therefore the brain region where inputs from internal sources conjoin with information received from the outside world. This region may act in a metacognitive fashion to monitor all stimuli, whether internal or external, so that our conscious sense of self at any particular moment reflects a workspace determined by which brain regions are most active.

Finally, evidence is accumulating that the medial prefrontal cortex is important not only for processing information about the self, but also for inferring mental states in others (Macrae et al., 2004; Mitchell et al., 2005). This raises the possibility that having a self might be adaptive because it allows us to simulate the mental lives of others, thereby allowing us to better know others and predict their behavior. Functionally, having ToM allows us to be good group members because we can predict how others will respond to our actions and ensure that we act in accordance with group norms and values. Such a theory is consistent with the argument that a symbolic self is adaptive (Sedikides & Skowronski, 1997). These and other theories will inspire further research on the social brain.

REFERENCES


References 877


Self versus Others/Self-Regulation


