

Meeting of minds: the medial frontal cortex and social cognition

David M. Amodio* and Chris D. Frith†

Abstract | Social interaction is a cornerstone of human life, yet the neural mechanisms underlying social cognition are poorly understood. Recently, research that integrates approaches from neuroscience and social psychology has begun to shed light on these processes, and converging evidence from neuroimaging studies suggests a unique role for the medial frontal cortex. We review the emerging literature that relates social cognition to the medial frontal cortex and, on the basis of anatomical and functional characteristics of this brain region, propose a theoretical model of medial frontal cortical function relevant to different aspects of social cognitive processing.

For humans, like many animal species, survival depends on effective social functioning. Social skills facilitate our access to sustenance, protection and mates, and socially adept individuals tend to be healthier and live longer^{1,2}. However, social interaction in humans is exceedingly complex compared with that in other animal species: representations of internal somatic states, knowledge about the self, perceptions of others and interpersonal motivations are carefully orchestrated to support skilled social functioning. This complex set of processes, which is broadly referred to as social cognition³, has recently been associated with activity in a network of brain regions, including the medial frontal cortex (MFC, in which, for convenience, we include the anterior cingulate cortex, ACC), the temporoparietal junction, the superior temporal sulcus and the temporal poles. This research suggests that the MFC has a special role in social cognition, whereas other regions in the network serve more general functions. However, so far, the functional significance of this activity is not well understood.

Social cognition has been studied from various theoretical and methodological perspectives. In the behavioural sciences, social psychologists have investigated how the self interacts dynamically with the social environment, and how knowledge structures of social groups (such as stereotypes) might influence behaviour through both conscious and unconscious mechanisms⁴⁻⁶. Although social psychologists have developed a rich theoretical and methodological framework for examining and understanding social cognition, they have only recently begun to consider its neural substrates. Neuroscientists, meanwhile, have investigated how underlying neural structures support unique yet

coordinated roles in various aspects of social cognition. Initially, neuroscientific explorations of social cognition arose from neuropsychological studies of patients^{7,8}. More recently, non-invasive neuroimaging methods such as functional MRI (fMRI) have permitted neuroscientists to explore the neural correlates of social cognitive phenomena in normally-functioning humans. As a result of these evolving fields, social psychologists and cognitive neuroscientists have begun to cross paths in the domain of social cognitive neuroscience⁹. Although they have arrived from different theoretical and methodological origins and often speak different scientific languages, they share a common goal: to understand the relationship between the brain and the social mind. In this review, we seek to integrate theory and research from neuroscience and social psychology in order to place this work in a broader conceptual framework and promote synergy across fields.

In proposing a theoretical framework for understanding the role of the MFC in social cognition, we consider evidence from three broad categories of tasks suggested by recent studies of functional divisions in the MFC¹⁰. The first category concerns control and monitoring of action, which is typically associated with activity in the dorsal ACC, although some have also proposed that the pre-supplementary motor area (preSMA) has a role¹¹. The second category concerns the monitoring of outcomes that relate to punishments and rewards, which is linked to activity in the orbital cortex. Finally, we focus on the category of primary interest: social cognition. Social cognitive processes, such as self-reflection, person perception, and making inferences about others' thoughts, have been associated with activity extending

*Department of Psychology,
New York University,
6 Washington Place,
New York 10003, USA.

†Wellcome Department of
Imaging Neuroscience,
Institute of Neurology,
University College London,
12 Queen Square, London
WC1N 3BG, UK.

e-mails:
cfrith@fil.ion.ucl.ac.uk;
david.amodio@nyu.edu
doi:10.1038/nrn1884

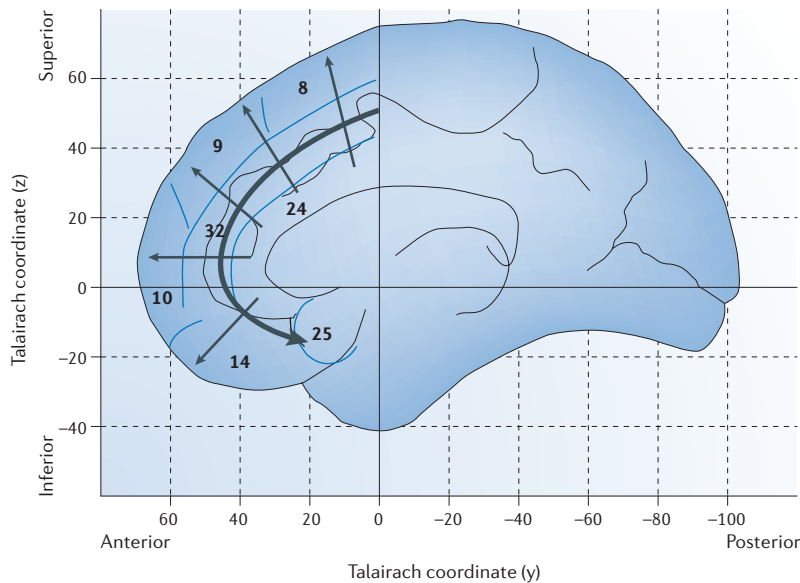


Figure 1 | Anatomical studies of the medial frontal cortex reveal two major axes of connectivity. The first axis, illustrated by the thick arrow, bends around the genu of the corpus callosum. Along this axis, the more superior regions (Brodmann areas 8 and 9) and the more superior parts of the anterior cingulate cortex have strong connections with lateral premotor, supplemental motor and cingulate motor areas, whereas the most inferior areas have strong connections with the rhinal cortex. The second axis of connectivity runs at right angles to the first, as illustrated by the thin arrows. Along this axis, cingulate regions have stronger connections with the amygdala than do more dorsal and frontopolar regions. The approximate demarcation of the Brodmann areas is taken from REF. 18.

from the ACC to the anterior frontal poles, most typically located in the transitional area between these two regions, the paracingulate cortex. Throughout this review we use the MFC as a designation that subsumes all these regions. We begin by outlining the anatomical subregions of the human MFC and their respective connections with other brain regions, primarily on the basis of anatomical studies of the monkey brain. Next, we review research reporting selective MFC activation in tasks associated with action monitoring, outcome monitoring, self-knowledge, person-knowledge and mentalizing — restricting included research to studies of normal adults using common neuroimaging methods such as fMRI, positron emission tomography (PET) and electroencephalography (EEG). We then propose a theoretical framework to account for the observed activation patterns, whereby the MFC supports a general mechanism for the integration of complex representations of possible actions and anticipated outcomes, and suggest that such integration is particularly relevant to the domain of social cognition.

Connectivity of the MFC

The MFC consists of Brodmann areas (BAs) 9 and 10 (medial regions), 24, 25 and 32, with 11 and 14 in the medial orbital cortex (FIG. 1). Most MFC projections are intrinsic or involve neighbouring prefrontal areas¹². With regard to more distal connections, the medial and lateral regions of orbitofrontal cortex (OFC) are part of

distinct networks. Medial regions of the OFC receive few direct sensory-related inputs, in contrast to lateral regions¹³. Rather, major afferents to the medial regions come from the dorsolateral prefrontal cortex, temporal pole, anterior superior temporal gyrus, parietotemporal cortex and posterior cingulate cortex^{12,13}.

There are two distinct axes within the medial prefrontal region, along which patterns of connectivity vary (FIG. 1). The first axis bends around the genu of the corpus callosum (thick arrow in FIG. 1), along which the most inferior caudal areas (BAs 25, 24 and 32) have strong connections with the rhinal cortex, compared with the adjacent, more rostral areas (BAs 14, 32 and 10). The most superior part of the medial prefrontal cortex (BA 9) has few, if any, connections with the rhinal cortex, but instead has robust connections with the lateral premotor cortex, the supplementary motor area and the cingulate motor area¹². The more superior and caudal parts of the ACC (BAs 24 and 32) are also connected with the premotor cortex¹⁴. A second axis runs at right angles to the first, distinguishing cingulate regions from frontopolar regions (thin arrows in FIG. 1). The amygdala has strong inputs to cingulate regions (BAs 24, 25 and 32), but only weak connections with frontopolar regions (BAs 9 and 10)^{15,16}. Evidence for these distinct axes of connectivity is also reflected in thalamic connections with the medial prefrontal cortex¹⁷.

What little is known about the connectivity of the MFC has been derived from studies of monkeys. Although the same architectonic areas can be identified in humans and monkeys^{18,19}, the frontopolar region (including BAs 10 and 32) is greatly expanded in humans relative to monkeys¹⁹. However, whether this region is expanded in humans relative to great apes remains controversial²⁰. Clearly some caution is warranted in applying connectivity findings in monkeys to humans. Nevertheless, a recent meta-analysis of PET-derived functional connectivity in the human brain supports the distinctions between these superior–inferior and caudal–rostral axes in human medial frontal regions²¹. Results from the recently developed technique of diffusion tractography also suggest considerable similarity between connectivity in human and monkey prefrontal cortices²².

Functional divisions of the MFC

In human studies, functional divisions may be determined by the nature of various tasks found to activate medial frontal regions. The most caudal region of the MFC contains one or more cingulate motor areas, which are differentially involved in movements of the hand, eye and mouth^{23,24}, and activity in this region has been related directly to behavioural response rates²⁵. Koski & Paus²¹ suggest that the division between the caudal and rostral ACC can be made at the vertical plane defined by the Talairach coordinate $y = 10$ (line a in FIG. 2). The more posterior region of the rostral ACC (prACC, using the nomenclature of Picard and Strick²⁴ but sometimes called the dorsal ACC) has been associated with ‘cognitive’ tasks (for example, attention and error monitoring), whereas the more anterior region of the rostral ACC

Talairach coordinates

Talairach coordinates provide a standardized method for describing the location of activations in the brain in three-dimensional space. Talairach space comprises x, y and z coordinates (represented as x,y,z): x denotes left versus right, y denotes rostral (anterior) versus caudal (posterior), and z denotes dorsal (superior) versus ventral (inferior). The Montreal Neurological Institute’s system uses the same metric space, but their coordinates are based on a slightly larger and more representative brain.

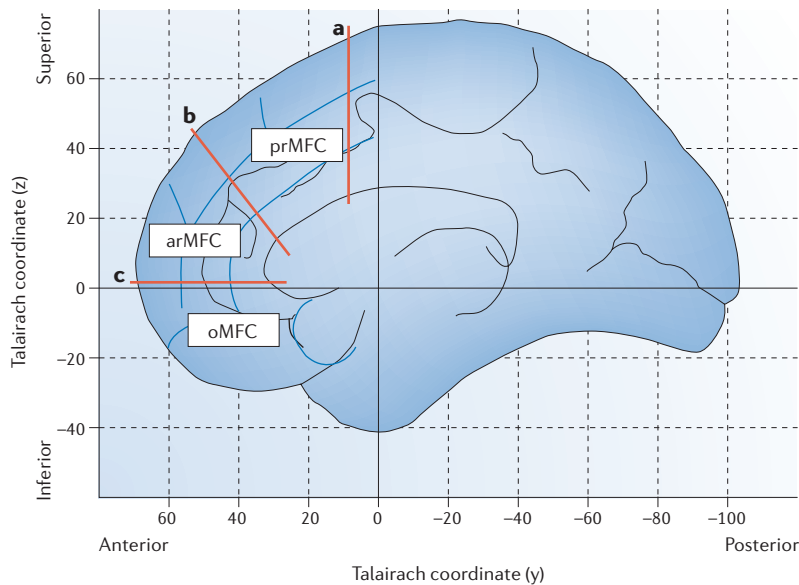


Figure 2 | Functional divisions of the medial prefrontal cortex. Meta-analyses of task-related neural activations observed in the medial frontal cortex (MFC) have revealed functional divisions associated with cognitive versus emotional processes^{10,25,26}. The more posterior region of the rostral MFC (prMFC) is activated by cognitive tasks, such as those designed to engage action monitoring and attention. By contrast, the more anterior region of the rostral MFC (arMFC) is activated by emotional tasks, such as rating one's emotions in response to pictures of varying valence. Line **a** denotes the division between the prMFC and caudal MFC, and line **b** divides the prMFC and arMFC. Line **c** marks the boundary between the arMFC and orbital MFC (oMFC). The oMFC has been linked to the monitoring of task outcomes associated with punishment or reward. For the purposes of this diagram, the MFC includes the anterior cingulate cortex.

(arACC) has been associated with 'emotional' tasks (for example, rating the pleasantness of pictures)¹⁰. On the basis of a detailed meta-analysis of studies in which activation of the MFC was observed, Steele & Lawrie²⁶ confirmed this distinction between cognitive and emotional regions, and defined a boundary providing maximum discrimination (line **b** in FIG. 2). Koski and Paus²¹ suggest that there is a fourth, subcallosal region, which is defined by the horizontal plane at Talairach coordinate $z = 2$, approximating the split in the ACC between supracallosal BAs 24 and 32, and subcallosal BAs 24 and 14. A meta-analysis of PET studies suggests that subcallosal activations are related to autonomic and visceral aspects of emotional responses²¹. In the following sections we examine more closely the wide range of experimental tasks that have been found to activate different divisions of the medial wall of the frontal cortex. On the basis of these findings, we speculate on the processes instantiated in these regions. Before focusing on the uniquely social cognitive functions ascribed to the anterior rostral MFC (arMFC), which includes the paracingulate cortex, we characterize the processes associated with the regions that flank the arMFC along the caudal–rostral axis noted above to provide a theoretical and anatomical context for our final discussion of arMFC function. The location of the activity elicited by the various studies discussed below is shown in FIG. 3.

Stroop colour-naming task

The Stroop task is commonly used to investigate response conflict. Participants view words presented in colours (for example, red and blue) that are either compatible (red written in red) or incompatible (red written in blue) with the word meaning. On incompatible trials, participants must inhibit the prepotent tendency to read the word's text in order to correctly report the colour of the word.

Response inhibition

Response inhibition refers to the process of withholding a habitual response when changing task demands require an alternative response. Response inhibition is a crucial component of behavioural regulation that has been ascribed as a function of the posterior rostral ACC by much research, although it probably involves the coordination of several neural systems.

Posterior region of the rostral MFC

The posterior rostral MFC (prMFC) has been implicated in the continuous internal monitoring of action across several studies²⁷. Humans continuously monitor their actions to ensure that they are consistent with intentions and the current situational context. Action monitoring is particularly important in situations involving response conflicts, as in the Stroop colour-naming task, or requiring response inhibition. Such conflicts typically elicit increased response errors and engage slower, more controlled patterns of response^{28–30}. Neuroimaging and event-related potential (ERP) research has linked the process of action monitoring to MFC activity. Barch *et al.*³¹ report an extensive meta-analysis of functional imaging studies that included data from three different tasks involving action monitoring: those in which prepotent responses must be inhibited; responses are not fully determined by the task context; or errors are committed. Activity observed in these studies generally clusters in the prACC (mean Talairach coordinates: 3,19,35). In a study by Gerhing *et al.*³², participants categorized target stimuli that were sometimes flanked by distractor stimuli associated with an erroneous response, which caused conflict and elicited errors. Response errors on this task, which reflect a conflict between intention and behaviour, evoked an ERP component that has been localized to the ACC^{33–35}. Subsequent research has shown that this component, the error-related negativity (ERN), is larger when stimulus conflict is high versus low³⁶, when errors lead to large versus small monetary losses³⁷, and when errors indicate the undesired application of social stereotypes³⁸.

These findings suggest that conflict monitoring, error monitoring and response selection might depend on a single underlying process instantiated in the prMFC (but for an alternative view, see REFS 10,37). For example, recent studies have emphasized the role of decision making in action selection — that is, a mechanism for choosing one action rather than another (for an example, see REF. 39). Given a choice, we select actions expected to lead to better outcomes. Such selection requires a representation of expected values of different actions, as well as the continuous monitoring of outcomes in order to update these expected values. Several studies have investigated prediction and monitoring processes associated with selection of action. Walton *et al.*³⁹ observed activity in the prMFC (0,18,36) when participants monitored the outcome of actions that were self-selected, but not when they monitored the outcome of externally-guided actions. Knutson *et al.*⁴⁰ reported that the activity in the prMFC (0,22,42) was correlated with trial-by-trial variations in the anticipated probability of monetary gain. In research by Coricelli *et al.*, a similar region of prMFC activity (0,24,33) was associated with regret, that is, discovering that an unselected action would have led to a better outcome⁴¹. Finally, Brown & Braver reported that prMFC activation (8,33,33) was associated with prediction of the probability of error⁴². This set of findings is also consistent with the suggestion that the prMFC is involved in the processing of ambiguous response feedback¹¹.

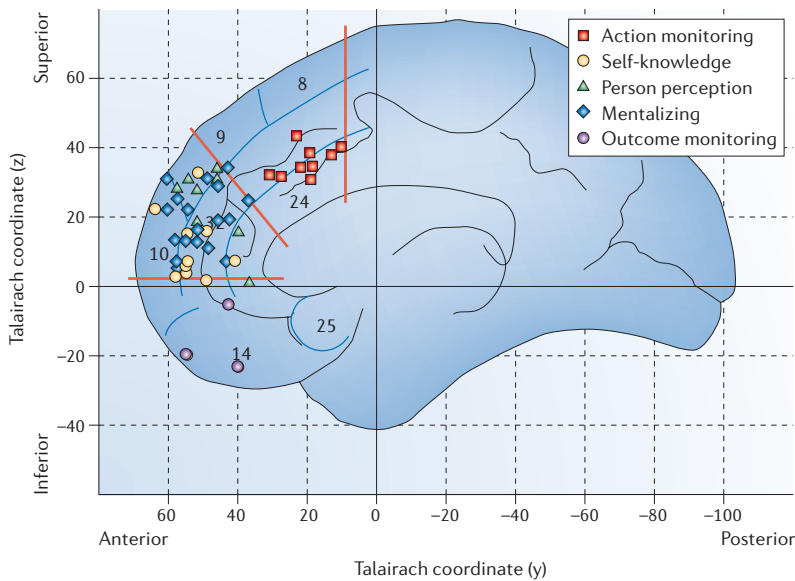


Figure 3 | Mapping of medial frontal cortex activations observed during action monitoring, social cognition and outcome monitoring. A meta-analysis of medial frontal cortex (MFC) activations suggests that social cognition tasks, which involve self-knowledge^{52,53,55–58,96,97,120,121}, person perception^{61–65,69,122,123} and mentalizing^{55,72,73,77–79,109,111,112,118,124–131}, activate areas in the anterior rostral MFC (arMFC). By contrast, activations from action-monitoring tasks^{28,29,83,99,132–136} occur in the posterior rostral region of the MFC (prMFC), and activations from tasks involving the monitoring of outcomes^{40,39,41} occur in the orbital MFC (oMFC).

Considered as a whole, the literature suggests that the prMFC is involved in representing and continuously updating the value of possible future actions in order to regulate behaviour⁴³.

Orbital region of the MFC

Traditionally, the orbitofrontal cortex (OFC) has been implicated in processing information concerning rewards and punishments (for an example, see REF. 44). Elliott *et al.*⁴⁵ proposed distinct roles for the lateral and medial OFC on the basis of neuroimaging studies of learning and gambling tasks. They concluded that the OFC is involved in monitoring the reward value of stimuli and responses, with the lateral OFC having a special role in situations in which responses to previously-rewarded stimuli must be suppressed. This characterization is complementary to our description of the posterior rostral MFC as being involved in monitoring the value of different possible actions. Rather than guiding behaviour in terms of the value of possible actions (as with the prMFC), we propose that the oMFC guides behaviour in terms of the value of possible outcomes. Supporting this characterization, Knutson *et al.* reported that an area of the orbital region of the MFC (oMFC; mean Talairach coordinates $-4,52,-6$), in addition to the prMFC, was associated with anticipated gain probability⁴⁰. Walton *et al.*³⁹ found that the activity in the oMFC ($12,54,-22$), but not the prMFC, was elicited by the need to monitor the outcomes of externally guided actions. They conjectured that, as the actions did not need to be chosen by the participants, the values of these

actions were not relevant and therefore the prMFC was not activated. Furthermore, Coricelli *et al.* found that activity in the oMFC ($-10,40,-24$) correlated with the amount of anticipated regret associated with a decision⁴¹. Studies of patients with lesions to the oMFC support this characterization of the role of this region^{46,47}.

Taken together, these results are consistent with the idea that the oMFC represents and updates the value of possible future outcomes, just as the prMFC represents and updates the value of possible future actions^{39,48}. These characterizations are consistent with the anatomical connectivity of these regions, with the oMFC being primarily connected to sensory association areas, and the prMFC being primarily connected to the motor system. These characterizations are also consistent with similar functional distinctions in the striatum regarding action versus reward, which have topographical connections to posterior versus anterior regions of the frontal cortex^{49,50}.

Anterior region of the rostral MFC

The location of the anterior rostral region of the MFC (arMFC) between the two regions discussed so far — the prMFC and oMFC — suggests it has access to information about both actions and outcomes. However, this characterization does not sufficiently explain the wide range of tasks shown to activate the arMFC, which comprise roughly three different categories: self-knowledge, person knowledge and mentalizing (BOX 1).

Self-knowledge. Socrates famously urged his followers to ‘know thyself’. Modern psychology suggests this is much easier said than done. The self is a complex and dynamic phenomenon that is often difficult to operationalize for scientific study⁵¹. At its most basic level, self-knowledge involves the ability to differentiate the self from other objects and to recognize attributes and preferences related to oneself. Initial neuroimaging investigations of the self have asked participants to determine whether a series of trait words apply to themselves. Evaluation of self-related traits has been shown to elicit activity in the arMFC in several studies^{52–56} (but see BOX 2). Extending this basic finding, Macrae *et al.* gave their subjects a surprise recognition test of trait words, some of which subjects previously judged according to their self-relevance⁵⁷. Words associated with increased arMFC activity during initial viewing were more likely to be remembered.

Activation of the arMFC has also been associated with the monitoring of one’s own emotional state. For example, Ochsner *et al.*⁵⁸ monitored neural activity while participants viewed images depicting a person in a positive, negative or neutral scene. On each trial, participants were asked to judge their own affective response, the affective response of the person in the picture, or whether the picture depicted an indoor or outdoor scene. Judgements of one’s own affective response, relative to judging whether the scene was indoors or outdoors, activated the arMFC. On the basis of previous studies^{10,59} and their own research²⁶, Steele and Laurie have suggested that this region

Event-related potential (ERP). An electrical signal produced by summated postsynaptic potentials of cortical neurons in response to a discrete event, such as a stimulus or response in an experimental task. Typically recorded from the scalp in humans, ERPs can be measured with extremely high temporal resolution and can be used to track rapid, real-time changes in neural activity.

Box 1 | Mentalizing tasks

The story of Max and the chocolate

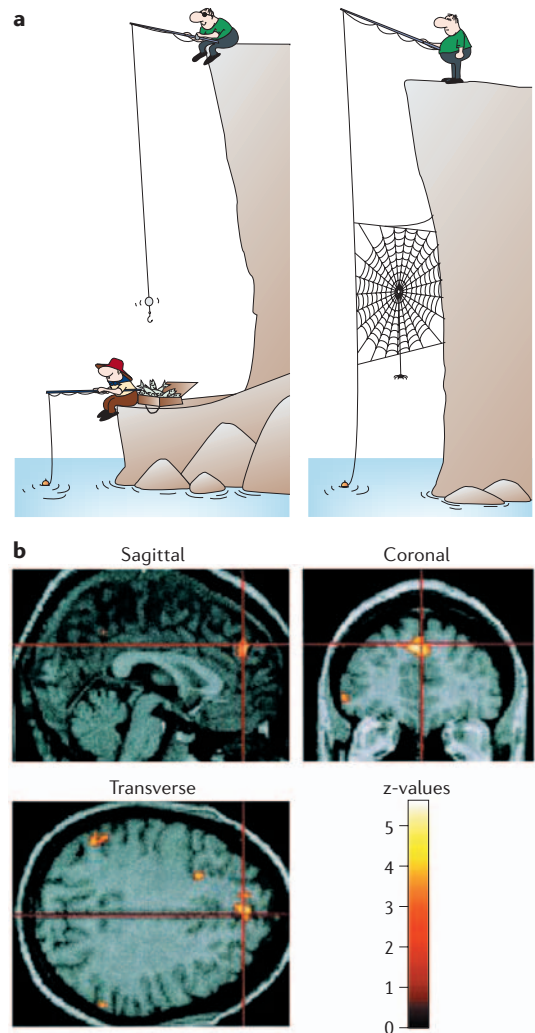
Max eats half his chocolate bar and puts the rest away in the kitchen cupboard. He then goes out to play in the sun. Meanwhile, Max's mother comes into the kitchen, opens the cupboard and sees the chocolate bar. She puts it in the fridge.

When Max comes back into the kitchen, where does he look for his chocolate bar: in the cupboard, or in the fridge?¹⁰⁸

Mentalizing is the cognitive process that needs to be engaged to answer this question. We have to recognize that Max's behaviour will be determined by the current contents of his mind and what he believes about the world, not by the actual state of the world. We must recognize that he doesn't know his mother has moved the chocolate, and therefore falsely believes that it is still in the cupboard. This story is an example of a first-order false belief. If Max had peeped back into the kitchen when his mother was moving the chocolate, his mother would have had a second-order false belief. She would have falsely believed that Max believed the chocolate was in the cupboard. Mentalizing is relevant for thinking about other peoples' intentions and desires as well as their beliefs.

Brain imaging studies of mentalizing (or theory of mind) have used many different tasks. For example, stories — as in the example above — or strip cartoons illustrating similar stories without words are presented, and the subject is asked to explain the behaviour of the characters or choose the appropriate ending to the narrative (for examples, see REFS 72, 109). In other studies, subjects passively view animations of simple objects that move and interact in a way that automatically elicits attributions of mental states^{110,111}. In yet other studies, the subject engages in a real-time interaction with another person during a competitive or cooperative game (for example, rock-paper-scissors) in which success requires the 'reading' of the intentions of the other person (for example, see REF. 112). It is striking that these different paradigms all elicit activity in the medial frontal cortex when contrasted with appropriate control tasks⁷⁰.

The cartoon (panel **a**) illustrates the idea of theory of mind^{113,114}. The joke on the left depends on the deceit of the man who is stealing the fish. The joke on the right does not involve theory of mind. Comparison of the two types of joke shows activation in the medial prefrontal cortex, which is shown in a single subject in panel **b**. Panels **a** and **b** reproduced, with permission, from REF. 115 © (2003) Macmillan Publishers Ltd.



(mean Talairach coordinates: 5,46,18) is concerned with emotion, in contrast to the adjacent more posterior and superior region that is concerned with cognition. However, this conclusion is based largely on research involving emotion induction, in which participants are asked to report their emotional experience. Such commonly-used 'emotion' tasks overlap significantly with tasks assessing self-knowledge — that is, being asked to report one's emotional response is essentially a question about self-knowledge. Given the observation that many other studies activating the same region did not involve strong emotions, we suggest that its characterization as an emotional sector of the MFC is not appropriate. However, direct comparisons of introspection about 'hot' and 'cold' mental states might well reveal systematic differences.

Person perception. Much research investigating the neural activity associated with the perception and judgments of other people has implicated the arMFC⁶⁰. Participants in a study performed by Mitchell *et al.*⁶¹ judged whether adjectives 'could ever be true' of preceding nouns that referred to people or inanimate objects. Judgments about people activated regions of the arMFC (for example, 3,39,0), whereas judgments about inanimate objects activated regions associated with semantic memory. Similar activations were found when participants decided whether behaviours were appropriate for people versus dogs (10,48,32)⁶², formed impressions about people as opposed to objects (-9,54,36)⁶³, observed social interactions (2,52,26)⁶⁴, and viewed personally familiar faces (-4,53,19)⁶⁵. It is not clear from these studies whether the observed arMFC activity is associated with thinking

Box 2 | The meaning of resting state activation in the MFC

In experimental tasks involving self-judgments, reductions in the blood oxygen level dependent (BOLD) responses from a baseline condition are often observed. For example, Mitchell *et al.*⁶¹ showed that self-related judgments elicited greater activity in the medial frontal cortex (MFC) than other-related judgments. However, self-related judgments did not elicit more activity than a baseline condition involving passive fixation on a cross-hair; rather, other-related judgments were associated with a significant reduction relative to baseline activity. Reduction of the BOLD signal in the medial prefrontal cortex is also observed in comparison to a resting baseline with a number of cognitive tasks, such as working memory¹¹⁶. This curious but commonly observed pattern raises questions about the nature of the baseline condition. What cognitive processes are engaged when subjects are simply instructed to rest? Some theorists have suggested that a resting baseline promotes self-reflective thought^{96,117} or social 'day dreaming'¹¹⁸. D'Argebeau *et al.* provided direct evidence for these speculations¹¹⁹. They confirmed that self-referential thoughts did occur during rest and that activity in the MFC (measured with positron emission tomography) correlated with the amount of self-referential processing.

It seems plausible, therefore, that a baseline of unconstrained rest is likely to elicit some of the same cognitive processes and associated neural activity as are engaged by explicit social tasks. It is also likely that the cognitive activity that occurs during rest might depend on the context in which this condition occurs. In order to interpret the results of studies concerned with activity in the MFC, high level comparison tasks are essential⁶⁷. We suggest that if unconstrained rest is used as an additional baseline condition, then the experimenter should make some attempt to discover what cognitive processes are occurring under this condition.

about people *per se* or with the mental states ascribed to them. Supporting the latter interpretation, Mitchell *et al.* found that activity in the arMFC (9,54,36) was associated with thinking about the mental states of dogs as well as those of people⁶⁶.

Observations from self-knowledge studies raise the possibility that activations elicited during the judgment of self-attributes (discussed above) might actually represent a more general process of thinking about 'social' attributes, regardless of whether they pertain to the self. Although several studies have directly contrasted thinking about one attribute of the self versus another, the results are equivocal. For example, Kelley *et al.* observed more activity in the arMFC (10,52,2) of participants when they were thinking about attributes of the self versus George W. Bush⁵³. However, Schmitz *et al.*⁵⁵ observed activity in a similar region (6,56,4) when participants thought about either the self or a close friend (see also REF. 67). The key difference between these studies might lie in the degree of similarity between the self and the other person. This possibility has been investigated explicitly by Mitchell *et al.*⁶⁸. In their study, participants viewed a series of faces and judged them for similarity to themselves. Thinking about more similar others led to greater arMFC activity (9,57,3) in an area very close to the peak activity associated with self-description reported by Kelley *et al.*⁵³. Indeed, there seems to be some evidence for a spatial separation between activity elicited when thinking about the self and a similar or familiar other versus unknown others (for examples, see REFS 67,69). The region activated by the self and others close to the self is in the most inferior portion of the region we have labelled the arMFC, whereas the region activated by unknown others is in the most superior portion (FIG. 3).

Mentalizing. To engage in successful social interaction, one must recognize that others have independent experiences and intentions, and perhaps even ulterior motives. The ability to represent another person's psychological perspective is referred to as mentalizing⁷⁰ and requires theory of mind⁷¹ (BOX 1). Mentalizing allows us to predict the behaviour of others. Initial investigations into the neural correlates of mentalizing observed a characteristic network of activations when participants read stories about social interaction, which included the temporoparietal junction, the superior temporal sulcus, the temporal poles, the posterior cingulate cortex and the MFC^{72,73}. Various mentalizing tasks have been studied, including story and cartoon comprehension, and the viewing of real-time interactions, which consistently activate the MFC, primarily in the arMFC region^{70,74}.

In the social psychological literature, the process of determining the causes of a person's behaviour (for example, their beliefs, perceptions or goals) is known as attribution^{75,76}. Attribution research asks how we know when a person's behaviour reflects their disposition or their situation. Although attribution is more about the interaction of personality and the situation in determining a person's behaviour, whereas mentalizing is more about inferring another's current knowledge and intentions, a recent fMRI investigation of attribution processes found that judgments made on the basis of clear attributional information are associated with activation of the arMFC (5,50,0), as when typical mentalizing tasks are used⁷⁷.

Walter *et al.*⁷⁸ make an interesting distinction between reading the private intentions of a person (for example, replacing a light bulb in order to read) and the communicative intentions that are involved in social interactions (for example, showing a map to request directions). They claim that only the reading of communicative intentions is associated with activity in the paracingulate cortex, whereas reading private intentions activates the ACC proper. The same distinction was observed by Grèzes *et al.* using a very different paradigm^{79,80}. In these studies, participants observed a video of someone lifting a box. In the first study, the person in the video had sometimes been deceived about the weight of the box. When the subjects judged, from the lifting movements, that the person in the video had a false belief (leading to a private intention about how to lift the box), greater activity was seen in the prMFC (-2,26,56). In the second study, the person in the video sometimes tried to deceive the observer by pretending that the box was heavier or lighter than it really was. When judging that the person in the video was being deceptive (a communicative intention), greater activity was seen in the arMFC (-8,42,20). Therefore, consistent with the subdivisions outlined above, thinking about private intentions elicits activity in the prMFC, whereas thinking about communicative intentions elicits activity in the arMFC.

Each of the tasks reviewed here that elicited activity in the arMFC involved thinking about the psychological attributes of people regardless of whether the person

Blood oxygen level dependent signal
(BOLD signal). fMRI measures local changes in the proportion of oxygenated blood in the brain; the blood oxygen level dependent, or BOLD, signal. This proportion changes in response to neural activity. Therefore, the BOLD signal, or haemodynamic response, indicates the location and magnitude of neural activity.

Working memory

Working memory refers to a set of processes involved in rehearsing and manipulating information that has either just been experienced or just been retrieved from long-term memory, often in the service of goal-directed behaviour. Working memory functions are typically associated with activity in the dorsolateral prefrontal cortex.

was the self or another person, or whether judgments pertained to dispositions or mental states. However, there are hints of further divisions within this area. Thinking about unfamiliar others and thinking about the simple actions of others activates the lower border of the prMFC and the most superior area of the arMFC. By contrast, thinking about familiar others activates the most inferior area of the arMFC and the upper border of the oMFC. Does this separation relate to the distinction between outcomes and actions? People may have ideas about how unfamiliar others might act. We can often predict actions on the basis of the situation a person is in without needing to know what sort of a person they are. By contrast, predicting how people will feel might depend more on having some familiarity with them. So, we can speculate that the more superior part of the region is more involved with actions, whereas the inferior part is more involved with feelings and outcomes. This division is, of course, commensurate with the likely function of the adjacent regions: the prMFC pertaining to actions, the oMFC pertaining to outcomes.

Value, pain and self-reflection

We have proposed a functional characterization of more posterior MFC regions, but the functional significance of the arMFC remains less clear. In the remainder of this review, we address this problem by trying to define a trajectory of cognitive processing in the MFC. Our proposal is that representations become more complex and abstract as we move forward through the MFC. A similar proposal has previously been made by Ochsner and Gross in relation to the representation of reinforcement contingencies⁸¹. Studies of pain provide clues to the form taken by this increasing degree of abstraction. For example, on the basis of pain research, Craig⁸² has proposed that high-resolution, modality-specific sensory representation of the physiological condition of the body in the posterior insula is re-represented in the anterior insula. This second-order re-representation in the right anterior insula is believed to subservise subjective feelings and the awareness of a physical self. We propose that a similar progression, which is involved in the broader process of social cognition, occurs in the MFC.

Pain controls our behaviour through a particularly primitive form of value, but even the negative value of pain is subject to top-down control. Rainville *et al.* studied changes in neural activity associated with the analgesic effects of hypnosis in different segments of the ACC^{83,84}. In the most caudal region (1,5,56/-1,3,39), activity elicited by a painful stimulus was unaffected by hypnosis. However, the reduced subjective experience of pain resulting from hypnosis was associated with decreased activity in the prACC (3,20,30/0,29,35). Similar results were reported by Wager *et al.*⁸⁵, such that treatment with placebo analgesia did not affect the caudal ACC (0,-4,50) and inferior ACC (0,54,-18) activity in response to pain, whereas placebo treatment was associated with reduced prACC activity (4,23,27/3,18,34). Furthermore, the change in prACC

activity was correlated with the change in subjective ratings of pain (see also REF. 86). These studies suggest that the caudal region of the ACC represents more objective aspects of pain (for example, the temperature of the stimulus), whereas the prACC represents subjective properties of pain.

This distinction between objective and subjective aspects of pain has also been suggested by research on empathy by Singer *et al.*⁸⁷. Activity in the caudal ACC (6,6,42) was elicited only by pain felt by the self. Activity in the prACC (-3,24,33) was elicited by pain to the self and also by the knowledge that a significant other was in pain (see also REFS 88,89). Here again, the response to pain in this more anterior region was independent of sensory input, suggesting a more abstract form of representation. The results of EEG studies show that this region is also involved when we observe the actions of others. A negative ERP component arising from the MFC is seen not only when we make an error, but also when we receive delayed error feedback⁹⁰⁻⁹² or observe someone else making an error^{93,94}. How can we link these results with our characterization of the prACC as being about the value of possible actions? An important facet of pain is our strong desire for action to escape or reduce it. Similarly, we also have strong drives to take action when we know that a significant other is in pain. Top-down influences that devalue the pain, such as hypnosis and placebo analgesia, should also devalue our drive to take action.

The stimuli used in studies of negative emotion induction can also be viewed as the application of painful stimuli. Typically, subjects are shown photographs of unknown people in painful or dangerous situations. Whether or not these stimuli elicit activity in the arMFC depends on the task participants are given. A stimulus with strong emotional valence^{95,96} activates the caudal ACC (14,6,30/-5,3,48) even when participants are simply reporting whether the scene is indoors or outdoors, but the arMFC is only activated by such stimuli when participants are asked to rate their emotional arousal in response to the stimuli; the activated regions are located at Talairach coordinates 0,50,16 (REF. 97) and -3,41,8 (REF. 96). As with pain, the caudal region of the ACC is activated by unpleasant stimuli across tasks, whereas a more anterior region is activated only when participants report how unpleasant the picture makes them feel. Again, we find that more anterior regions of the MFC seem to be concerned with subjective, more abstract representations of experience.

But why does the subjective experience of pain activate the prMFC, whereas reporting the subjective experience of emotionally arousing pictures activates the arMFC? The primary difference is that pain is intrinsically unpleasant, whereas a picture is only unpleasant when we think about the experiences of the people depicted in it. In the study of Singer *et al.*, the arMFC was not activated by application of pain to the participant⁸⁷. However, activity in this region (-6,45,21) was elicited by the knowledge that a loved one was in pain, with greater activity associated with higher self-ratings of empathy. Therefore, it is not the subjective unpleasantness

of the pain itself that activates this region, but rather thinking about the subjective unpleasantness of the pain. Thinking about pain is the same meta-cognitive process that is involved in thinking about the unpleasantness of emotionally-arousing pictures.

Recent evidence reported in the ERP literature also suggests that the more anterior regions of the MFC are involved in more complex social cognitive processing. ERP research on error processing and cognitive control has revealed two important neural components. The ERN, which occurs within milliseconds of a response, reflects an early, pre-conscious stage of conflict monitoring that is associated with dorsal regions of the ACC, whereas the error-positivity (P_e), which occurs ~200 ms post response, is associated with the awareness of error commission and has been linked to activity in the rostral anterior cingulate and paracingulate cortices^{98,99}. This distinction is consistent with functional organizations suggested by Eisenberger and Lieberman¹⁰⁰, and by Ochsner and his colleagues^{58,67}. In a study by Amodio *et al.*¹⁰¹ that capitalized on this distinction, participants performed a cognitive task purporting to measure their level of racial bias either confidentially (in private) or while being observed by an ostensibly non-prejudiced experimenter (in public). When in private, participants presumably regulated their response according to their internal motives for accuracy on the task, whereas in public they additionally regulated behaviour according to the perceived social demand to appear non-prejudiced. Amodio *et al.* found that better response control in private was predicted solely by larger ERN amplitudes, replicating past work^{32,38}. By contrast, when responding in public, better response control was predicted by larger P_e amplitudes only among participants who had previously reported being highly sensitive to social pressures to appear non-prejudiced. Importantly, the condition in which the P_e strongly predicted behaviour involved monitoring the value that others put on the actions of the self. This is certainly a more complex and abstract representation of the value of actions: it is a meta-representation that enables us to reflect on the value of an action and contrast this with the value that others would place on the same action.

Morality, reputation and the self

If we characterize the role of the arMFC as allowing the meta-cognitive process of reflecting on feelings and intentions, then we can provide a unitary account of the wide range of different tasks that activate this region. The ability to reflect on our subjective experience is of great importance for many aspects of social cognition. For example, when confronted with a moral dilemma, we base our decision on what feels like the right thing to do rather than on a logical analysis. Such decisions are associated with activity in the arMFC (0,50,17)³². People often reflect on their feelings for information when making decisions about what they like and dislike¹⁰², and such reflection has been found to activate this same region (-6,55,13)⁵⁶.

An important, but often underappreciated, aspect of moral decisions is that they are not based solely on

reflections about the self, but also relate to the image of the self we want to project into the minds of others: our reputation. Much social psychological research shows that there is a distinction between our actual behaviour and the image we wish to have of ourselves and to present to others^{103,104}. This concept of reputation is essentially a representation of how others represent us (although note that this re-representation of subordinate knowledge structures might not be conscious or deliberative, and is therefore distinguished from meta-cognition). Such a representation goes beyond the first-level representations of our own attributes or the attributes of others. The representation of our own reputation requires that we close the social loop and form a second-level representation of the attributes that others apply to us. Ochsner *et al.* refer to this as reflected self-knowledge⁶⁷. We have to think about how others think about us. This is perhaps why the same region is activated whether we are thinking about our own psychological attributes or those of others.

Self- and other-referencing, and counter-referencing are even more obviously involved when playing economic games that involve trust and reciprocity¹⁰⁵. Before we invest we must decide not just whether we trust the other player, but also whether the other player trusts us. Consistent with our formulation, the arMFC is activated when playing these games, especially when participants are cooperating, as long as they believe they are playing against a person rather than a computer. In these studies, arMFC activation was observed at Talairach coordinates 5,52,10 (REF. 106) and 3,44,20 (REF. 107).

Conclusions

The meeting of neuroscientists and social psychologists in research on the MFC has led to a remarkably rich and varied set of experimental data from which to speculate about the function of this region. Our assumption has been that the different functions instantiated in this region are not placed randomly, but form a systematic map. As part of the frontal cortex, this region is concerned with determining future behaviour. More specifically, it is concerned with determining behaviour on the basis of anticipated value. In the more caudal region of the MFC value is associated with actions, whereas in the more orbital region value is associated with outcomes. These representations become more abstract as we move forward, such that the most anterior region of the MFC is associated with meta-cognitive representations that enable us to reflect on the values linked to outcomes and actions (that is, thinking about thinking). These high level representations have a major role in many aspects of social cognition. Not only do they allow us to reflect on the values that other people attach to actions and outcomes, they also allow us to reflect on what other people think about us. These speculations remain to be confirmed, but we hope we have provided a framework for future research on the role of the MFC in social cognition that will permit fruitful interactions between neuroscientists and social psychologists.

1. Cohen, S. Social relationships and health. *Am. Psychol.* **59**, 676–684 (2004).
2. Silk, J. B., Alberts, S. C. & Altman, J. Social bonds of female baboons enhance infant survival. *Science* **302**, 1231–1234 (2003).
3. Fiske, S. T. & Taylor, S. E. *Social Cognition* (McGraw-Hill, New York, 1991).
4. Bargh, J. A. in *Handbook of Social Cognition: Basic Processes* (Eds Wyer, R. S. Jr & Srull, T. K.) 1–40 (Hillsdale, New Jersey/Lawrence Erlbaum Associates, Inc. England, 1994).
5. Devine, P. G. Stereotypes and prejudice: their automatic and controlled components. *J. Pers. Soc. Psychol.* **56**, 5–18 (1989).
6. Kihlstrom, J. F. The cognitive unconscious. *Science* **237**, 1445–1452 (1987).
7. Eslinger, P. J. & Damasio, A. R. Severe disturbance of higher cognition after bilateral frontal lobe ablation: patient EVR. *Neurology* **35**, 1731–1741 (1985).
8. Farah, M. J. Agnosia. *Curr. Opin. Neurobiol.* **2**, 162–164 (1992).
9. Ochsner, K. N. & Lieberman, M. D. The emergence of social cognitive neuroscience. *Am. Psychol.* **56**, 717–734 (2001).
- The authors propose the theoretical and methodological integration of neuroscience, cognitive science and social psychology, and introduce the new field of social cognitive neuroscience.**
10. Bush, G., Luu, P. & Posner, M. I. Cognitive and emotional influences in anterior cingulate cortex. *Trends Cogn. Sci.* **4**, 215–222 (2000).
11. Ullsperger, M. & von Cramon, D. Y. Neuroimaging of performance monitoring: error detection and beyond. *Cortex* **40**, 593–604 (2004).
- Important evidence on the connectivity of medial prefrontal cortex obtained from studies of the monkey that provides a model of anatomical connectivity of the MFC in humans.**
12. Barbas, H., Ghashghaei, H., Dombrowski, S. M. & Rempel-Clower, N. L. Medial prefrontal cortices are unified by common connections with superior temporal cortices and distinguished by input from memory-related areas in the rhesus monkey. *J. Comp. Neurol.* **410**, 343–367 (1999).
13. Carmichael, S. T. & Price, J. L. Connectional networks within the orbital and medial prefrontal cortex of macaque monkeys. *J. Comp. Neurol.* **371**, 179–207 (1996).
14. Bates, J. F. & Goldman-Rakic, P. S. Prefrontal connections of medial motor areas in the rhesus monkey. *J. Comp. Neurol.* **336**, 211–228 (1993).
15. Amaral, D. G. & Price, J. L. Amygdalo-cortical projections in the monkey (*Macaca fascicularis*). *J. Comp. Neurol.* **230**, 465–496 (1984).
16. Vogt, B. A. & Pandya, D. N. Cingulate cortex of the rhesus monkey: II. Cortical afferents. *J. Comp. Neurol.* **262**, 271–289 (1987).
17. Ray, J. P. & Price, J. L. The organization of projections from the mediodorsal nucleus of the thalamus to orbital and medial prefrontal cortex in macaque monkeys. *J. Comp. Neurol.* **337**, 1–31 (1993).
18. Petrides, M. & Pandya, D. N. Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. *Eur. J. Neurosci.* **11**, 1011–1036 (1999).
19. Ongur, D., Ferry, A. T. & Price, J. L. Architectonic subdivision of the human orbital and medial prefrontal cortex. *J. Comp. Neurol.* **460**, 425–449 (2003).
20. Sherwood, C. C., Holloway, R. L., Semendeferi, K. & Hof, P. R. Is prefrontal white matter enlargement a human evolutionary specialization? *Nature Neurosci.* **8**, 537–538; author reply 538 (2005).
21. Koski, L. & Paus, T. Functional connectivity of the anterior cingulate cortex within the human frontal lobe: a brain-mapping meta-analysis. *Exp. Brain Res.* **133**, 55–65 (2000).
- Provides a useful account of the connectivity of the ACC based on human functional data.**
22. Croxson, P. L. *et al.* Quantitative investigation of connections of the prefrontal cortex in the human and macaque using probabilistic diffusion tractography. *J. Neurosci.* **25**, 8854–8866 (2005).
23. Paus, T., Petrides, M., Evans, A. C. & Meyer, E. Role of the human anterior cingulate cortex in the control of oculomotor, manual, and speech responses: a positron emission tomography study. *J. Neurophysiol.* **70**, 453–469 (1993).
24. Picard, N. & Strick, P. L. Motor areas of the medial wall: a review of their location and functional activation. *Cereb. Cortex* **6**, 342–353 (1996).
25. Paus, T., Koski, L., Caramanos, Z. & Westbury, C. Regional differences in the effects of task difficulty and motor output on blood flow response in the human anterior cingulate cortex: a review of 107 PET activation studies. *Neuroreport* **9**, R37–R47 (1998).
26. Steele, J. D. & Lawrie, S. M. Segregation of cognitive and emotional function in the prefrontal cortex: a stereotactic meta-analysis. *Neuroimage* **21**, 868–875 (2004).
- A careful meta-analysis of studies activating the MFC that reveals clear functional segregation of cognitive and emotional processes.**
27. Botvinick, M. M., Cohen, J. D. & Carter, C. S. Conflict monitoring and anterior cingulate cortex: an update. *Trends Cogn. Sci.* **8**, 539–546 (2004).
28. Botvinick, M., Nystrom, L. E., Fissell, K., Carter, C. S. & Cohen, J. D. Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature* **402**, 179–181 (1999).
29. Carter, C. S. *et al.* Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science* **280**, 747–749 (1998).
30. Liddle, P. F., Kiehl, K. A. & Smith, A. M. Event-related fMRI study of response inhibition. *Hum. Brain Mapp.* **12**, 100–109 (2001).
31. Barch, D. M. *et al.* Anterior cingulate cortex and response conflict: effects of response modality and processing domain. *Cereb. Cortex* **11**, 837–848 (2001).
32. Gehring, W., Goss, B., Coles, M., Meyer, D. & Donchin, E. A neural system for error detection and compensation. *Psychol. Sci.* **4**, 385–390 (1993).
33. Brázdil, M. *et al.* Combined event-related fMRI and intracerebral ERP study of an auditory oddball task. *Neuroimage* **26**, 285–293 (2005).
34. Dehaene, S., Posner, M. & Tucker, D. Localization of a neural system for error detection and compensation. *Psychol. Sci.* **5**, 303–305 (1994).
35. Yeung, N., Cohen, J. D. & Botvinick, M. M. The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychol. Rev.* **111**, 931–959 (2004).
36. Bartholow, B. D. *et al.* Strategic control and medial frontal negativity: beyond errors and response conflict. *Psychophysiology* **42**, 33–42 (2005).
37. Hajcak, G., Moser, J. S., Yeung, N. & Simons, R. F. On the ERN and the significance of errors. *Psychophysiology* **42**, 151–160 (2005).
38. Amodio, D. M. *et al.* Neural signals for the detection of unintentional race bias. *Psychol. Sci.* **15**, 88–93 (2004).
39. Walton, M. E., Devlin, J. T. & Rushworth, M. F. Interactions between decision making and performance monitoring within prefrontal cortex. *Nature Neurosci.* **7**, 1259–1265 (2004).
- Reveals an important distinction between monitoring actions and monitoring outcomes in the prefrontal cortex.**
40. Knutson, B., Taylor, J., Kaufman, M., Peterson, R. & Glover, G. Distributed neural representation of expected value. *J. Neurosci.* **25**, 4806–4812 (2005).
41. Coricelli, G. *et al.* Regret and its avoidance: a neuroimaging study of choice behavior. *Nature Neurosci.* **8**, 1255–1262 (2005).
- An important demonstration of the role of emotion (anticipated regret) in decision making**
42. Brown, J. W. & Braver, T. S. Learned predictions of error likelihood in the anterior cingulate cortex. *Science* **307**, 1118–1121 (2005).
43. Holroyd, C. B. & Coles, M. G. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.* **109**, 679–709 (2002).
44. Rolls, E. T. The orbitofrontal cortex. *Phil. Trans. R. Soc. Lond. B* **351**, 1433–1443; discussion 1443–1444 (1996).
45. Elliott, R., Dolan, R. J. & Frith, C. D. Dissociable functions in the medial and lateral orbitofrontal cortex: evidence from human neuroimaging studies. *Cereb. Cortex* **10**, 308–317 (2000).
46. Camille, N. *et al.* The involvement of the orbitofrontal cortex in the experience of regret. *Science* **304**, 1167–1170 (2004).
47. Fellows, L. K. & Farah, M. J. Different underlying impairments in decision-making following ventromedial and dorsolateral frontal lobe damage in humans. *Cereb. Cortex* **15**, 58–63 (2005).
48. Ridderinkhof, K. R., Ullsperger, M., Crone, E. A. & Nieuwenhuis, S. The role of the medial frontal cortex in cognitive control. *Science* **306**, 443–447 (2004).
49. Alexander, G. E., DeLong, M. R. & Strick, P. L. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.* **9**, 357–381 (1986).
50. Lehericy, S. *et al.* Diffusion tensor fiber tracking shows distinct corticostriatal circuits in humans. *Ann. Neurol.* **55**, 522–529 (2004).
51. Baumeister, R. F. in *The Handbook of Social Psychology* Vol. 1, 4th edn (eds Gilbert, D. T., Fiske, S. T. & Lindzey, G.) 680–740 (McGraw Hill, New York, USA, 1998).
52. Johnson, S. C. *et al.* Neural correlates of self-reflection. *Brain* **125**, 1808–1814 (2002).
53. Kelley, W. M. *et al.* Finding the self? An event-related fMRI study. *J. Cogn. Neurosci.* **14**, 785–794 (2002).
54. Lieberman, M. D., Jarcho, J. M. & Satpute, A. B. Evidence-based and intuition-based self-knowledge: an fMRI study. *J. Pers. Soc. Psychol.* **87**, 421–435 (2004).
55. Schmitz, T. W., Kawahara-Baccus, T. N. & Johnson, S. C. Metacognitive evaluation, self-relevance, and the right prefrontal cortex. *Neuroimage* **22**, 941–947 (2004).
56. Zysset, S., Huber, O., Ferstl, E. & von Cramon, D. Y. The anterior frontomedian cortex and evaluative judgment: an fMRI study. *Neuroimage* **15**, 983–991 (2002).
57. Macrae, C. N., Moran, J. M., Heatherton, T. F., Banfield, J. F. & Kelley, W. M. Medial prefrontal activity predicts memory for self. *Cereb. Cortex* **14**, 647–654 (2004).
58. Ochsner, K. N. *et al.* Reflecting upon feelings: an fMRI study of neural systems supporting the attribution of emotion to self and other. *J. Cogn. Neurosci.* **16**, 1746–1772 (2004).
- An important demonstration of the role of the MFC in emotion attributions, and a detailed review of MFC activation associated with inferring emotional responses of the self and others.**
59. Phan, K. L., Wager, T., Taylor, S. F. & Liberzon, I. Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage* **16**, 331–348 (2002).
60. Lieberman, M. D. & Pfeifer, J. H. in *Cognitive Neuroscience of Emotional and Social Behavior* (eds Easton, A. & Emery, N.) 195–235 (Psychology Press, Philadelphia, in the press).
61. Mitchell, J. P., Heatherton, T. F. & Macrae, C. N. Distinct neural systems subserving person and object knowledge. *Proc. Natl Acad. Sci. USA* **99**, 15238–15243 (2002).
62. Mason, M. F., Banfield, J. F. & Macrae, C. N. Thinking about actions: the neural substrates of person knowledge. *Cereb. Cortex* **14**, 209–214 (2004).
63. Mitchell, J., Macrae, C. & Banaji, M. Forming impressions of people versus inanimate objects: social-cognitive processing in the medial prefrontal cortex. *Neuroimage* **26**, 251–257 (2005).
64. Iacoboni, M. *et al.* Watching social interactions produces dorsomedial prefrontal and medial parietal BOLD fMRI signal increases compared to a resting baseline. *Neuroimage* **21**, 1167–1173 (2004).
65. Gobbini, M. I., Leibenluft, E., Santiago, N. & Haxby, J. V. Social and emotional attachment in the neural representation of faces. *Neuroimage* **22**, 1628–1635 (2004).
66. Mitchell, J. P., Banaji, M. R. & Macrae, C. N. General and specific contributions of the medial prefrontal cortex to knowledge about mental states. *Neuroimage* **28**, 757–762 (2005).
- An important demonstration that activity in the MFC is associated with the inference of others' mental states, rather than specific person-related attributes.**
67. Ochsner, K. N. *et al.* The neural correlates of direct and reflected self-knowledge. *Neuroimage* **28**, 797–814 (2005).
68. Mitchell, J. P., Banaji, M. R. & Macrae, C. N. The link between social cognition and self-referential thought in the medial prefrontal cortex. *J. Cogn. Neurosci.* **17**, 1306–1315 (2005).
69. Craik, F. I. M. *et al.* In search of the self: a positron emission tomography study. *Psychol. Sci.* **10**, 26–34 (1999).
70. Frith, U. & Frith, C. D. Development and neurophysiology of mentalizing. *Phil. Trans. R. Soc. Lond. B* **358**, 459–473 (2003).
71. Premack, D. & Woodruff, G. Chimpanzee problem-solving: a test for comprehension. *Science* **202**, 532–535 (1978).

72. Fletcher, P. C. *et al.* Other minds in the brain: a functional imaging study of 'theory of mind' in story comprehension. *Cognition* **57**, 109–128 (1995).
73. Goel, V., Grafman, J., Sadato, N. & Hallett, M. Modeling other minds. *Neuroreport* **6**, 1741–1746 (1995).
74. Saxe, R., Carey, S. & Kanwisher, N. Understanding other minds: linking developmental psychology and functional neuroimaging. *Annu. Rev. Psychol.* **55**, 87–124 (2004).
75. Gilbert, D. T. in *The Handbook of Social Psychology* Vol. 2, 4th edn (eds Gilbert, D. T., Fiske, S. T. & Lindzey, G.) 89–150 (McGraw Hill, New York, USA, 1998).
76. Jones, E. E. & Davis, K. E. in *Advances in Experimental Social Psychology* (ed. Berkowitz, L.) 220–266 (Academic, New York, 1965).
77. Harris, L. T., Todorov, A. & Fiske, S. T. Attributions on the brain: neuro-imaging dispositional inferences, beyond theory of mind. *Neuroimage* **28**, 763–769 (2005).
78. Walter, H. *et al.* Understanding intentions in social interaction: the role of the anterior paracingulate cortex. *J. Cogn. Neurosci.* **16**, 1854–1863 (2004).
79. Grezes, J., Frith, C. & Passingham, R. E. Brain mechanisms for inferring deceit in the actions of others. *J. Neurosci.* **24**, 5500–5505 (2004).
Demonstrates an interesting distinction between the role of the MFC and the pSTS in social cognition that involves mentalizing.
80. Grezes, J., Frith, C. D. & Passingham, R. E. Inferring false beliefs from the actions of oneself and others: an fMRI study. *Neuroimage* **21**, 744–750 (2004).
81. Ochsner, K. N. & Gross, J. J. The cognitive control of emotion. *Trends Cogn. Sci.* **9**, 242–249 (2005).
82. Craig, A. D. How do you feel? Interoception: the sense of the physiological condition of the body. *Nature Rev. Neurosci.* **3**, 655–666 (2002).
Proposes that interoceptive information is represented in an explicit form in the anterior insula.
83. Rainville, P., Duncan, G. H., Price, D. D., Carrier, B. & Bushnell, M. C. Pain affect encoded in human anterior cingulate but not somatosensory cortex. *Science* **277**, 968–971 (1997).
84. Rainville, P. *et al.* Cerebral mechanisms of hypnotic induction and suggestion. *J. Cogn. Neurosci.* **11**, 110–125 (1999).
85. Wager, T. D. *et al.* Placebo-induced changes in fMRI in the anticipation and experience of pain. *Science* **303**, 1162–1167 (2004).
86. Salomons, T. V., Johnstone, T., Backonja, M. M. & Davidson, R. J. Perceived controllability modulates the neural response to pain. *J. Neurosci.* **24**, 7199–7203 (2004).
87. Singer, T. *et al.* Empathy for pain involves the affective but not sensory components of pain. *Science* **303**, 1157–1162 (2004).
Reveals functional segregation of the anterior cingulate in relation to sensory and affective properties of pain.
88. Jackson, P. L., Meltzoff, A. N. & Decety, J. How do we perceive the pain of others? A window into the neural processes involved in empathy. *Neuroimage* **24**, 771–779 (2005).
89. Morrison, I., Lloyd, D., di Pellegrino, G. & Roberts, N. Vicarious responses to pain in anterior cingulate cortex: is empathy a multisensory issue? *Cogn. Affect. Behav. Neurosci.* **4**, 270–278 (2004).
90. Donkers, F. C., Nieuwenhuis, S. & van Boxtel, G. J. Mediofrontal negativities in the absence of responding. *Brain Res. Cogn. Brain Res.* **25**, 777–787 (2005).
91. Luu, P., Tucker, D. M., Derryberry, D., Reed, M. & Poulsen, C. Electrophysiological responses to errors and feedback in the process of action regulation. *Psychol. Sci.* **14**, 47–53 (2003).
92. Gehring, W. J. & Willoughby, A. R. The medial frontal cortex and the rapid processing of monetary gains and losses. *Science* **295**, 2279–2282 (2002).
93. van Schie, H. T., Mars, R. B., Coles, M. G. & Bekkering, H. Modulation of activity in medial frontal and motor cortices during error observation. *Nature Neurosci.* **7**, 549–554 (2004).
94. Bates, A. T., Patel, T. P. & Liddle, P. F. External behavior monitoring mirrors internal behavior monitoring: error-related negativity for observed errors. *J. Psychophysiol.* **19**, 281–288 (2005).
95. Lane, R. D. *et al.* Neural correlates of levels of emotional awareness. Evidence of an interaction between emotion and attention in the anterior cingulate cortex. *J. Cogn. Neurosci.* **10**, 525–535 (1998).
96. Gusnard, D. A., Akbudak, E., Shulman, G. L. & Raichle, M. E. Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proc. Natl Acad. Sci. USA* **98**, 4259–4264 (2001).
A classic paper examining the functional significance of 'resting activity' in fMRI studies. The authors suggest that self-referential processes correspond to the default mode of function in the MFC.
97. Lane, R. D., Fink, G. R., Chau, P. M. & Dolan, R. J. Neural activation during selective attention to subjective emotional responses. *Neuroreport* **8**, 3969–3972 (1997).
98. Nieuwenhuis, S., Ridderinkhof, K. R., Blom, J., Band, G. P. & Kok, A. Error-related brain potentials are differentially related to awareness of response errors: evidence from an antisaccade task. *Psychophysiology* **38**, 752–760 (2001).
99. Van Veen, V. & Carter, C. S. The timing of action-monitoring processes in the anterior cingulate cortex. *J. Cogn. Neurosci.* **14**, 593–602 (2002).
100. Eisenberger, N. I. & Lieberman, M. D. Why rejection hurts: a common neural alarm system for physical and social pain. *Trends Cogn. Sci.* **8**, 294–300 (2004).
101. Amodio, D. M., Kubota, J. T., Harmon-Jones, E. & Devine, P. G. Alternative mechanisms for regulating racial responses according to internal vs. external cues. *Soc. Cogn. Affect. Neurosci.* (in press).
102. Schwarz, N. & Clore, G. L. Mood, misattribution, and judgments of well-being: informative and directive functions of affective states. *J. Pers. Soc. Psychol.* **45**, 513–523 (1983).
103. Devine, P. G., Monteith, M. M., Zuwerink, J. R. & Elliot, A. J. Prejudice with and without compunction. *J. Pers. Soc. Psychol.* **60**, 817–830 (1991).
104. Higgins, E. T. Self-discrepancy: a theory relating self and affect. *Psychol. Rev.* **94**, 319–340 (1987).
105. Camerer, C. F. Psychology and economics. Strategizing in the brain. *Science* **300**, 1673–1675 (2003).
106. McCabe, K., Houser, D., Ryan, L., Smith, V. & Trouard, T. A functional imaging study of cooperation in two-person reciprocal exchange. *Proc. Natl Acad. Sci. USA* **98**, 11832–11835 (2001).
The first demonstration of a role for the MFC in trust and reciprocity games.
107. Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E. & Cohen, J. D. The neural correlates of theory of mind within interpersonal interactions. *Neuroimage* **22**, 1694–1703 (2004).
108. Wimmer, H. & Perner, J. Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* **13**, 103–128 (1983).
109. Brunet, E., Sarfati, Y., Hardy-Bayle, M. C. & Decety, J. A PET investigation of the attribution of intentions with a nonverbal task. *Neuroimage* **11**, 157–166 (2000).
110. Heider, F. & Simmel, M. An experimental study of apparent behavior. *Am. J. Psychol.* **57**, 243–249 (1944).
111. Castelli, F., Happe, F., Frith, U. & Frith, C. Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *Neuroimage* **12**, 314–325 (2000).
112. Gallagher, H. L., Jack, A. I., Roepstorff, A. & Frith, C. D. Imaging the intentional stance in a competitive game. *Neuroimage* **16**, 814–821 (2002).
113. Corcoran, R., Cahill, C. & Frith, C. D. The appreciation of visual jokes in people with schizophrenia: a study of 'mentalizing' ability. *Schizophr. Res.* **24**, 319–327 (1997).
114. Gallagher, H. L. *et al.* Reading the mind in cartoons and stories: an fMRI study of 'theory of mind' in verbal and nonverbal tasks. *Neuropsychologia* **38**, 11–21 (2000).
115. Adolphs, R. Cognitive neuroscience of human social behaviour. *Nature Rev. Neurosci.* **4**, 165–178 (2003).
116. Gusnard, D. A. & Raichle, M. E. Searching for a baseline: functional imaging and the resting human brain. *Nature Rev. Neurosci.* **2**, 685–694 (2001).
117. Fransson, P. Spontaneous low-frequency BOLD signal fluctuations: an fMRI investigation of the resting-state default mode of brain function hypothesis. *Hum. Brain Mapp.* **26**, 15–29 (2005).
118. Lieberman, M. D., Gaunt, R., Gilbert, D. T. & Trope, Y. in *Advances in Experimental Social Psychology* (ed. Zanna, M.) 199–249 (Academic, New York, 2002).
119. D'Argebeau, A. *et al.* Self-referential reflective activity and its relationship with rest: a PET study. *Neuroimage* **25**, 616–624 (2005).
120. Maguire, E. A. & Mummery, C. J. Differential modulation of a common memory retrieval network revealed by positron emission tomography. *Hippocampus* **9**, 54–61 (1999).
121. Maguire, E. A., Mummery, C. J. & Buchel, C. Patterns of hippocampal–cortical interaction dissociate temporal lobe memory subsystems. *Hippocampus* **10**, 475–482 (2000).
122. Bonda, E., Petrides, M., Ostry, D. & Evans, A. Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *J. Neurosci.* **16**, 3737–3744 (1996).
123. Turk, D. J. *et al.* From facial cue to dinner for two: the neural substrates of personal choice. *Neuroimage* **22**, 1281–1290 (2004).
124. Bottini, G. *et al.* The role of the right hemisphere in the interpretation of figurative aspects of language. A positron emission tomography activation study. *Brain* **117**, 1241–1253 (1994).
125. Ferstl, E. C. & von Cramon, D. Y. What does the frontomedian cortex contribute to language processing: coherence or theory of mind? *Neuroimage* **17**, 1599–1612 (2002).
126. German, T. P., Niehaus, J. L., Roarty, M. P., Giesbrecht, B. & Miller, M. B. Neural correlates of detecting pretense: automatic engagement of the intentional stance under covert conditions. *J. Cogn. Neurosci.* **16**, 1805–1817 (2004).
127. Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M. & Cohen, J. D. An fMRI investigation of emotional engagement in moral judgment. *Science* **293**, 2105–2108 (2001).
128. Heekeren, H. R. *et al.* Influence of bodily harm on neural correlates of semantic and moral decision-making. *Neuroimage* **24**, 887–897 (2005).
129. Kampe, K. K., Frith, C. D. & Frith, U. "Hey John": signals conveying communicative intention toward the self activate brain regions associated with "mentalizing," regardless of modality. *J. Neurosci.* **23**, 5258–5263 (2003).
130. Schilbach, L. *et al.* Being with virtual others: neural correlates of social interaction. *Neuropsychologia* (in press).
131. Voegeley, K. *et al.* Mind reading: neural mechanisms of theory of mind and self-perspective. *Neuroimage* **14**, 170–181 (2001).
132. Braver, T. S., Barch, D. M., Gray, J. R., Molfese, D. L. & Snyder, A. Anterior cingulate cortex and response conflict: effects of frequency, inhibition and errors. *Cereb. Cortex* **11**, 825–836 (2001).
133. Bunge, S. A., Hazeltine, E., Scanlon, M. D., Rosen, A. C. & Gabrieli, J. D. Dissociable contributions of prefrontal and parietal cortices to response selection. *Neuroimage* **17**, 1562–1571 (2002).
134. Casey, B. J. *et al.* Dissociation of response conflict, attentional selection, and expectancy with functional magnetic resonance imaging. *Proc. Natl Acad. Sci. USA* **97**, 8728–8733 (2000).
135. Eisenberger, N. I., Lieberman, M. D. & Williams, K. D. Does rejection hurt? An fMRI study of social exclusion. *Science* **302**, 290–292 (2003).
136. Kerns, J. G. *et al.* Anterior cingulate conflict monitoring and adjustments in control. *Science* **303**, 1023–1026 (2004).

Acknowledgements

We are grateful to D. Passingham, J. Mitchell, K. Ochsner and an anonymous reviewer for their comments on early versions of this review.

Competing interests statement

The authors declare no competing financial interests.

FURTHER INFORMATION

Amodio's laboratory: <http://www.psych.nyu.edu/amodiolab>
Frith's laboratory: http://www.fil.ion.ucl.ac.uk/Frith_Lab/
Access to this links box is available online.