

Annual Review of Neuroscience

The Role of the Medial Prefrontal Cortex in Moderating Neural Representations of Self and Other in Primates

Masaki Isoda^{1,2}

¹Division of Behavioral Development, Department of System Neuroscience, National Institute for Physiological Sciences, National Institutes of Natural Sciences, Okazaki, Aichi 444-8585, Japan; email: isodam@nips.ac.jp

²Department of Physiological Sciences, School of Life Science, The Graduate University for Advanced Studies (SOKENDAI), Hayama, Kanagawa 240-0193, Japan

Annu. Rev. Neurosci. 2021. 44:295–313

First published as a Review in Advance on March 22, 2021

The *Annual Review of Neuroscience* is online at neuro.annualreviews.org

<https://doi.org/10.1146/annurev-neuro-101420-011820>

Copyright © 2021 by Annual Reviews.
All rights reserved

Keywords

social cognition, medial prefrontal cortex, other-mentalizing, mental model, autism spectrum disorder

Abstract

As a frontal node in the primate social brain, the medial prefrontal cortex (MPFC) plays a critical role in coordinating one's own behavior with respect to that of others. Current literature demonstrates that single neurons in the MPFC encode behavior-related variables such as intentions, actions, and rewards, specifically for self and other, and that the MPFC comes into play when reflecting upon oneself and others. The social moderator account of MPFC function can explain maladaptive social cognition in people with autism spectrum disorder, which tips the balance in favor of self-centered perspectives rather than taking into consideration the perspective of others. Several strands of evidence suggest a hypothesis that the MPFC represents different other mental models, depending on the context at hand, to better predict others' emotions and behaviors. This hypothesis also accounts for aberrant MPFC activity in autistic individuals while they are mentalizing others.

ANNUAL
REVIEWS **CONNECT**

www.annualreviews.org

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

Contents

| | |
|--|-----|
| INTRODUCTION | 296 |
| THE MPFC AS A CRITICAL NODE IN THE PRIMATE SOCIAL BRAIN | 296 |
| SELF-OTHER DISTINCTION FOR MODERATING THE SELF AND OTHERS | 298 |
| AGENT-SELECTIVE BEHAVIORAL SIGNALS IN THE MPFC | 299 |
| SOCIAL PREFERENCE OF MPFC NEURONS | 302 |
| MENTAL MODELS IN THE MPFC: EVIDENCE AND HYPOTHESIS | 303 |
| CLINICAL CONSIDERATIONS ON THE SOCIAL MODERATOR ROLE: EVIDENCE AND HYPOTHESIS | 305 |
| CONCLUDING REMARKS | 308 |

INTRODUCTION

Primates, and especially humans, are social by nature. Their presence and fitness as social entities depend on how well they can deal with others. In doing so, seeing the world solely from self-centered perspectives is not adaptive; one also needs to consider what others think and feel about the world. Thus, the key to survival from a long-term perspective is to coordinate one's own behavior with proper regard for others. This other-regarding social cognition is of vital importance for maintaining group cohesion, which eventually leads to success as a species in the animal kingdom (Byrne & Whiten 1988, Dunbar 2003, Herrmann et al. 2007).

Which brain regions are involved in coordinating one's own behavior with respect to that of others so as to ensure productive social exchanges? Here, I review literature converging on the notion that the medial prefrontal cortex (MPFC) plays such a moderator role. The MPFC refers to midline cortical structures, including the dorsomedial prefrontal cortex [mainly Brodmann area (BA)8 and BA9], the presupplementary motor area (pre-SMA; BA6), and the sulcus and gyrus regions of the anterior cingulate cortex (ACC; BA24 and BA32) (Ninomiya et al. 2018). In the following sections, I first describe how the MPFC originally attracted attention for its role in behavioral organization in nonsocial contexts but is now considered a critical node in what is termed the social brain. The MPFC is consistently activated when people infer mental states of others and predict their behavior, a cognitive process known as other-mentalizing. After reviewing how self-other distinction is the basis of other-mentalizing, I highlight electrophysiological findings in nonhuman primates showing that individual MPFC neurons encode a range of behavioral variables important for social decision-making—such as intentions, actions, and rewards—selectively for the self and for others. The activity of MPFC neurons encoding such behavioral variables from others is diminished when interacting with nonsocial objects compared to social agents. I then discuss a hypothesis that the MPFC represents mental models that predict how others would behave in a given context on the basis of other-specific behavioral variables. Finally, I provide a hypothetical account of what aberrant MPFC activity in the brains of autism spectrum disorder (ASD) patients might imply in the context of its role in moderating neural representations of self and others.

THE MPFC AS A CRITICAL NODE IN THE PRIMATE SOCIAL BRAIN

Experimental work during the past few decades has documented that the MPFC plays diverse roles in coordinating one's own behavior. Specifically, the MPFC has evaluative roles for behavioral control, which are collectively referred to as performance monitoring (for reviews, see Ridderinkhof

et al. 2004, Schall et al. 2002, Ullsperger et al. 2014). These roles include the detection of error commission (Falkenstein et al. 1990), external negative feedback indicating performance error (Holroyd & Coles 2002), and conflict in information processing (Botvinick et al. 1999). In addition to these evaluative roles, the MPFC has executive roles for behavioral adjustment. First, the MPFC plays an important role in the arrangement of multiple actions in an appropriate temporal order (Tanji 2001). Second, the MPFC is activated when switching from automatic actions to nonautomatic actions (Isoda & Hikosaka 2007); this switching ability is impaired in a patient with a focal lesion in the MPFC (Roberts & Husain 2015). Third, activation in the MPFC when facing a choice is increased with increasing uncertainty in event prediction (Volz et al. 2003). Finally, the MPFC encodes tactics for action selection, for example, reaching either toward or away from an illuminated target, and tactic-selective activity is further transformed into choice-selective activity (Matsuzaka et al. 2016). Models of underlying neural computations in the MPFC generally converge on principles of reinforcement learning. MPFC activity is correlated with errors in the prediction of reward, expectancy violation such as surprise, and the weighting of individual events for guiding future behavior (Alexander & Brown 2011, Behrens et al. 2007, Fischer & Ullsperger 2013, Holroyd & Coles 2002, Jocham et al. 2009, Matsumoto et al. 2007, McGuire et al. 2014). With these proposed roles, the MPFC enables organisms to update the value of response options and guide optimal behaviors in a constantly changing environment.

These studies have been performed using nonsocial task paradigms, in which a subject plans, generates, and monitors their own actions and the resulting consequences in the absence of others. Under these isolated conditions, the subject's action has no impact on others. In real life, however, others are ubiquitous. Under such social conditions, what others think, feel, and do are critical variables for decision-making. Moreover, what the subject does can affect others' states, be it positively or negatively, directly or indirectly. If the role of the MPFC in behavioral coordination is versatile, then one would expect that this cortical region also comes into play when coordinating one's own behavior while taking into consideration that of others. How then might the MPFC monitor, evaluate, and predict the behavior of others?

Accumulating evidence indicates that the MPFC of primates, in both humans and macaques, is implicated in a wide range of social cognitive functions (Ninomiya et al. 2018, Wittmann et al. 2018). First, the MPFC is involved in social valuation and learning (Apps & Ramnani 2014, Apps et al. 2015, Behrens et al. 2008, Hampton et al. 2008, Rudebeck et al. 2006, Seid-Fatemi & Tobler 2015, Suzuki et al. 2012). Second, the MPFC is consistently activated during other-mentalizing in humans (Amodio & Frith 2006, Gallagher & Frith 2003). This high-level cognition includes the capacity to attend to, evaluate, and predict others' behaviors, thereby adaptively coordinating one's own behavior. The foci of activation during such other-mentalizing range from BA6, including the pre-SMA (Baron-Cohen et al. 1999), to BA8 and BA9 (Fletcher et al. 1995, Gallagher et al. 2000, Goel et al. 1995, Happe et al. 1996) and to areas that are even more anterior such as BA10 (Gilbert et al. 2006). Activation of more anterior MPFC regions may occur when other-mentalizing requires higher-order recursions (Coricelli & Nagel 2009, Hampton et al. 2008). Finally, the MPFC is recruited when a social comparison is made (Bault et al. 2011, Kedia et al. 2014, Luo et al. 2018, Swencionis & Fiske 2014, Takahashi et al. 2009). A social comparison helps people make sense of themselves by looking to others for comparison standards (Festinger 1954). Although people high in social comparison hierarchies tend to reflect on their own feelings and thoughts, they also have a strong interest in what others feel, a strong empathy for others, and a general sensitivity to the needs of others (Buunk & Gibbons 2006).

In addition to noninvasive neuroimaging methods, researchers have developed novel experimental paradigms using two interactive macaques and investigated neuronal representations of other-derived social signals in the MPFC (Isoda et al. 2018). These studies revealed that, in at

least some species of monkeys, the MPFC is where self-related and other-related signals coexist (see below). Importantly, macaques display self-recognition behaviors toward a mirror, although this is controversial (Anderson & Gallup 2015), after training of visuo-somatosensory associations (Chang et al. 2015, 2017). Moreover, macaques can display prosocial behaviors by considering the welfare of their peers (Ballesta & Duhamel 2015, Basile et al. 2020, Chang et al. 2011). MPFC lesions significantly reduce prosocial preferences as indexed by a greater tendency to give rewards to others relative to no one (Basile et al. 2020). Furthermore, macaques can show false belief attribution-like behavior (Hayashi et al. 2020). In this study, animals were presented with a video in which a human actor observed where an object was hidden in one of two locations. The object was then moved to the other location or removed in the absence of the actor. When the actor returned to the scene to find the object, the animals looked at the false-belief location first, that is, the location where the actor last saw the object, before the actor directed his gaze there. Critically, deactivation of the MPFC eliminated the animals' anticipatory looking at the false-belief location (Hayashi et al. 2020). Together, these findings suggest that macaques have sophisticated social cognitive capabilities and that the MPFC mediates these functions. From an anatomical viewpoint, connectivity patterns of the frontal cortex, including the MPFC, with the rest of the brain are surprisingly similar in humans and macaques, even for regions thought to perform uniquely human functions (Sallet et al. 2013). Given these parallels, research in two closely related primate species, humans and macaques, provides important insights into the functional organization of the MPFC for moderating neural representations of self and others.

SELF-OTHER DISTINCTION FOR MODERATING THE SELF AND OTHERS

The capacity to mentalize is thought to require the ability to form mental representations of others' mental states that are distinct from one's own (Frith & Frith 1999). This claim makes great sense if one considers ordinary experiences whereby different people can have different or even opposing views of the same situation (Isoda 2016). Crucially, mindsets that afford a focus on self-other differences facilitate taking the perspective of others, suggesting that self-other distinction enables observers to look beyond the limits of their own perspective and provides a means to effectively reflect on others (Todd et al. 2011). The self-other distinction—the capacity to correctly distinguish between the representations of one's own actions, perceptions, and emotions and those of others—is considered the basis of many aspects of social cognition (Lamm et al. 2015). In the laboratory setting, the other-mentalizing capacity is assessed using false-belief tasks that require a distinction between one's own true beliefs and others' false beliefs (Baron-Cohen et al. 1985).

These cognitive requirements of the self-other distinction, as well as its consistent activation during other-mentalizing, raise the possibility that the MPFC is equipped with neural machinery capable of separating observed aspects of others from those of oneself. Such separation at the neural level helps to avoid self-other confusion, to see the world from others' perspectives (often by inhibiting prepotent self-perspectives), and to better predict others' future behavior. Lower levels of this ability to distinguish between the self and others, in both cognitive and neural terms, may be associated with the difficulties that children with ASD have performing false-belief tasks compared to those without ASD (Lee & Hobson 2006, Lombardo et al. 2010).

What then might be the neuronal substrate of the self-other distinction? It is unknown whether the self-other distinction is underpinned by a unitary or distinct mechanism across cognitive, affective, and motor domains (Lamm et al. 2015). As outlined in the following sections, however, the form of neuronal representations of agent-specific information seems similar at least in the coding of intentions, actions, and rewards. In the coding of actions, the so-called mirror neuron

theory of social cognition has been based on the unique property of mirror neurons to encode both the actions of oneself and those of others (Rizzolatti & Craighero 2004, Rizzolatti & Sinigaglia 2016). Specifically, this class of neurons, originally identified in the macaque ventral premotor cortex (PMv), fires when an animal both performs a motor action and observes the same action being performed by another agent (di Pellegrino et al. 1992). Because of this firing property, it seems that the brain's mirror neuron system alone cannot discriminate reliably between oneself and others. To prevent possible self–other confusion, a neural mechanism must exist that can keep others' behavioral signals separate from one's own. Despite ample evidence for the existence of shared coding, the neuronal substrate for distinctly encoding aspects of self and other is much less clear.

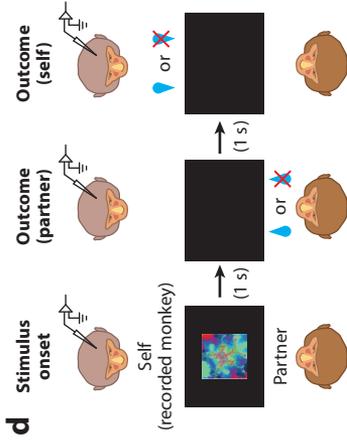
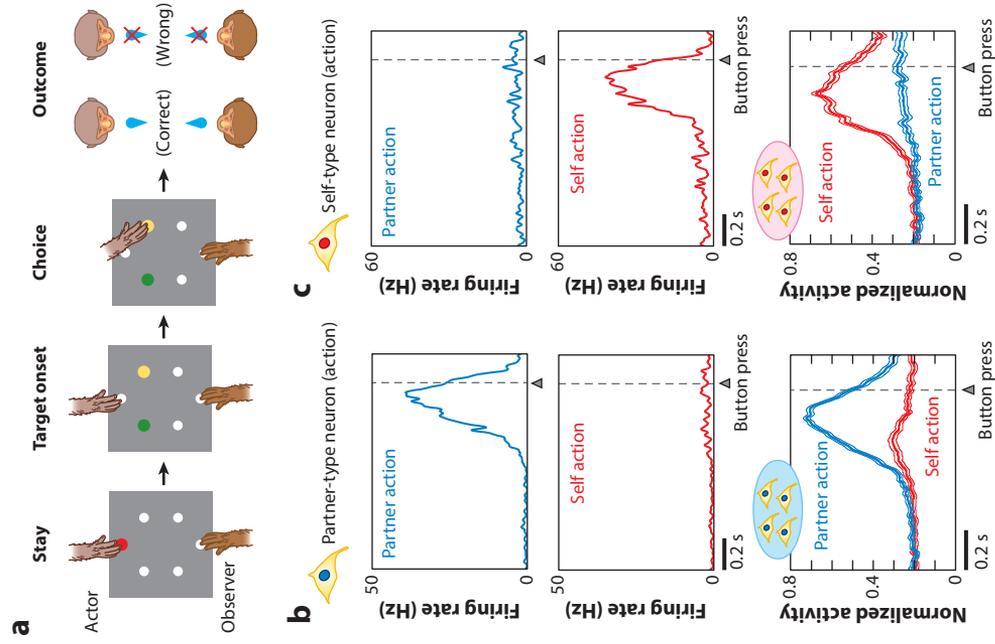
AGENT-SELECTIVE BEHAVIORAL SIGNALS IN THE MPFC

To explore a single-neuron substrate of agent-selective behavioral signals in the MPFC, Yoshida et al. (2011) devised a role-reversal choice task in which two macaques sitting face-to-face monitored each other's actions to guide their own optimal actions (**Figure 1a**). In each trial, one animal was assigned the role of actor and the other the role of observer; these two roles alternated every two trials. The actor was required to choose between a yellow or green button. If the actor made a correct choice, both animals received a reward; if the actor made a wrong choice, neither animal was rewarded. Thus, reward expectation was constant across the two roles in each trial, and the examiners were able to identify agent-selective and action-related neuronal signals.

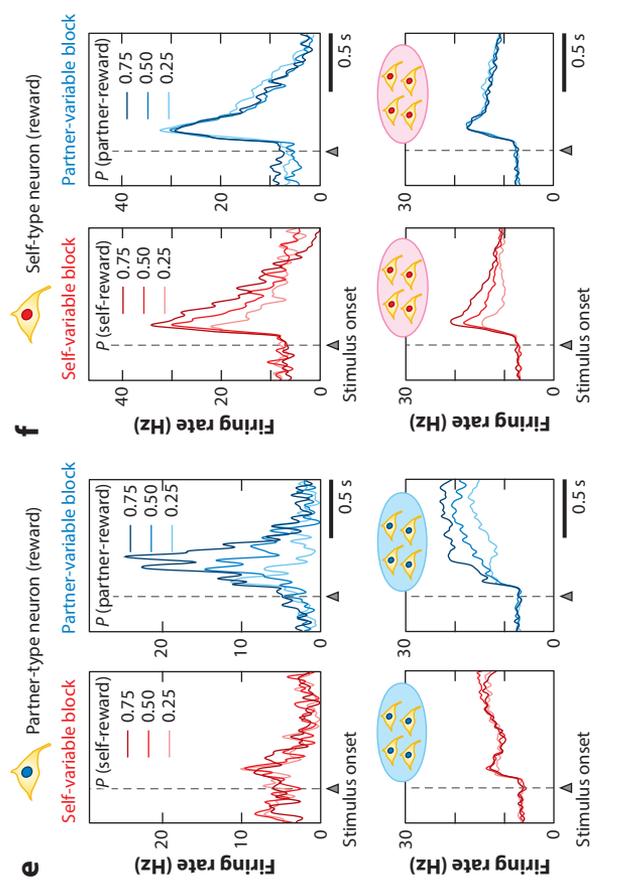
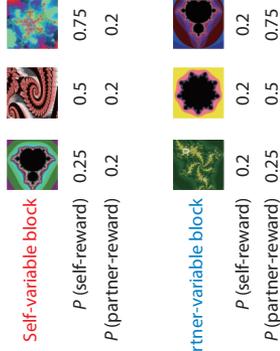
Yoshida et al. (2011) found that a group of MPFC neurons responded selectively or preferentially to the other's action (partner-type neurons; **Figure 1b**), while another group of neurons responded selectively or preferentially to the self-action (self-type neurons; **Figure 1c**). In addition to these agent-selective populations, neurons exhibiting agent-nonselective coding were also identified (mirror-type neurons). These findings demonstrate that the MPFC can distinguish between self-actions and other-actions at the single-neuron level.

The recording sites in these experiments encompassed the pre-SMA, area 9m (Miyachi et al. 2005) in the dorsomedial convexity region, and area 24c (Dum & Strick 1991, Matelli et al. 1991) in the cingulate sulcus region. Within these MPFC regions, the partner-type neurons were significantly more prevalent in the dorsomedial convexity region than in the cingulate sulcus region, whereas the self-type neurons were significantly more frequent in the cingulate sulcus region than in the dorsomedial convexity region (Yoshida et al. 2011). Furthermore, the partner-type neurons not only exhibited agent-specific response properties but also encoded the direction (Yoshida et al. 2011) and correctness (Yoshida et al. 2012) of the other-actions.

The role of MPFC neurons in the self–other distinction extends beyond observable actions and can also include covert intentions and beliefs. Haroush & Williams (2015) devised a strategic game between two macaques in which essential elements of the prisoner's dilemma (Rilling et al. 2002) were incorporated. Here, each animal chose between cooperating and defecting by pressing one of two buttons. When both animals cooperated, the highest mutual reward was delivered, but when one of them defected, the highest individual reward was delivered to the defector. When neither animal cooperated, a smaller reward was given to each animal compared to when they both cooperated. A critical point was that the animals were unable to see each other's choices until they had made their final choices. The investigators found that the animals were most likely to cooperate if both of them chose cooperating in the preceding trial, whereas they were less likely to cooperate if the other animal previously defected. Thus, the macaques took into account the other's past action when selecting their own. The investigators further found that neurons in the dorsal ACC co-fluctuated with the other's yet unknown choice. Specifically, about one-third



Stimulus-reward association



(Caption appears on following page)

Figure 1 (Figure appears on preceding page)

Agent-selective neuronal coding. (a) A role-reversal choice task is used to study agent-selective action coding. The two roles (actor and observer) alternate every two trials. (b) Charts show the activity of a single partner-type neuron (*top* and *middle*) and the population of partner-type neurons (*bottom*). For constructing population activity, firing rates of individual neurons are normalized using their maximum and minimum values and then averaged. (c) Charts show the activity of a single self-type neuron (*top* and *middle*) and the population of self-type neurons (*bottom*). (d) A social Pavlovian conditioning procedure is used to study agent-selective reward coding. The presentation of a conditioned stimulus (1 s) is followed by reward outcomes (water or nothing), first to the partner and then to the self (*left*). A low-pitched and high-pitched tone occurred together with the reward delivery to the partner and to self, respectively. The stimulus-reward association for the self and partner is also shown (*right*). $P(\text{self-reward})$ and $P(\text{partner-reward})$ denote the probability of self-rewards and partner-rewards, respectively. (e) Charts show the activity of a single partner-type neuron (*top*) and the population of partner-type neurons (*bottom*). Neuronal activity was differential only for partner-reward probabilities. (f) Charts show the activity of a single self-type neuron (*top*) and the population of self-type neurons (*bottom*). Neuronal activity was differential only for self-reward probabilities. Panels *a–c* adapted with permission from Yoshida et al. (2011) and panels *d–f* adapted with permission from Noritake et al. (2018).

of the recorded neurons demonstrated significant differences in activity when the other animal selected cooperation versus defection. This finding implies that dorsal ACC neurons encode predictions of others' actions or others' covert intentions. Neurons with similar functional properties were later identified in the amygdala (Grabenhorst et al. 2019), a subcortical region with which the ACC and other MPFC regions interact both anatomically (Ishida et al. 2018) and functionally (Dal Monte et al. 2020, Klavir et al. 2013, Rudebeck et al. 2017). Agent-specific anticipatory activity before the initiation of actions is also reported in the dorsomedial convexity region, including BA9 and pre-SMA (Falcone et al. 2017). The existence of neuronal responses to other-related intentions is reinforced by the finding that the inactivation of MPFC activity can eliminate a spontaneous gaze bias reflecting others' false beliefs (Hayashi et al. 2020). Most critically, Jamali et al. (2021) have very recently identified single neurons in the human dorsal MPFC that encode information about others' beliefs during a verbal false-belief task and that distinguish between representations of beliefs for self and others.

Livi et al. (2019) reported a novel aspect of agent coding in the MPFC. They showed that the mirror-type neurons in the pre-SMA [their self- and other-type (SOT) action-related neurons] conveyed more detailed information about self-actions than about other-actions, suggesting that MPFC neurons with mirror-like properties may also contribute to the self–other distinction. Moreover, the pre-SMA has an agent-based representation of graspable objects in the observer's peripersonal, but not extrapersonal, space. Most interestingly, despite the apparent similarity between neuronal coding of objects grasped by the self and by others, agent information is still embedded in, and thus decodable from, the activity of SOT object-related neurons.

In addition to motor actions and intentions, the activity of MPFC neurons is agent selective with respect to reward coding. Chang et al. (2011, 2013) devised a social reward–allocation task for two macaques. In this task, one of the animals chose between different options, each of which resulted in one of the following outcomes: self-reward only, other-reward only, or neither-reward. They found that neurons in the AC sulcus preferentially encoded reward allocations to nonself (i.e., another animal or neither), while neurons in the AC gyrus preferentially encoded reward allocations to the self, another animal, or both. AC gyrus neurons may encode reward signals in frames of reference for both the self and others (Chang 2013, Chang et al. 2013). By affecting activity in the AC gyrus, oxytocin and vasopressin can affect other-regarding behavior (Jiang & Platt 2018).

To study agent-selective reward coding, Noritake et al. (2018) devised a Pavlovian conditioning procedure within a social framework. Two macaques sat face-to-face and were conditioned with different visual stimuli (see the left side of **Figure 1d**), each of which predicted their reward outcomes with different probabilities. Specifically, two blocks of trials were run alternately. In one

block, the probability of self-rewards changed depending on the stimulus presented, but the probability of partner-rewards remained unchanged (self-variable block; see the top right of **Figure 1d**). In the other block, the probability of partner-rewards changed depending on the stimulus, but the probability of self-rewards remained unchanged (partner-variable block; see the bottom right of **Figure 1d**). In the partner-variable block, the subjective value of upcoming self-rewards, as indexed by both anticipatory licking and choice behavior during a stimulus presentation period, decreased as the probability of partner-rewards increased. During this conditioning, individual neurons in the dorsal MPFC (area 9m and pre-SMA) specifically tracked the probabilities of partner-rewards (partner-type neurons; **Figure 1e**) or self-rewards (self-type neurons; **Figure 1f**). Thus, in addition to information about who gets a reward (reward allocation), MPFC neurons also encoded the likelihood of rewards in an agent-selective manner.

In social life, people care deeply about rewards to both themselves and others. Such information about self-rewards and other-rewards and their comparisons is a key determinant in inducing positive other-regarding responses, such as empathy (Buunk & Gibbons 2006) and vicarious reward (Mobbs et al. 2009), or negative other-regarding responses, such as envy and schadenfreude (Ben-Ze'ev 2000, Dvash et al. 2010, Takahashi et al. 2009). The social reward signals in the MPFC are conveyed to subcortical reward motivation centers such as the dopaminergic midbrain nuclei (Noritake et al. 2018), amygdala (Dal Monte et al. 2020), and lateral hypothalamus (Noritake et al. 2020). The MPFC might play a role in switching other-regarding responses, depending on social contexts, via specific coordination with subcortical regions.

SOCIAL PREFERENCE OF MPFC NEURONS

To what extent is the activity of MPFC neurons in response to the behavior of others selective for, or influenced by, their social nature? This question is fundamental to understanding the organizing principles of the primate social brain. In the mirror neuron system of both humans and macaques, some studies showed that mirror neurons exhibited activity preferences for the action of biological agents over similar movements of nonbiological objects (di Pellegrino et al. 1992, Gallese et al. 1996, Rizzolatti et al. 1996, Tai et al. 2004), while others reported no such social preference (Caggiano et al. 2016, Oberman et al. 2007). In the MPFC, neuroimaging studies in humans identified strong activation in response to moving geometric shapes that evoke attribution of intentions (Castelli et al. 2000, Wheatley et al. 2007). Neuroimaging of the macaque MPFC, however, demonstrated higher levels of activation in response to the actions made by biological agents compared to nonbiological objects (Sliwa & Freiwald 2017). This issue has never been addressed at the single-neuron level in the MPFC.

To study the extent to which action-related activity in the MPFC is socially oriented, Ninomiya et al. (2020) devised a behavioral procedure in which essential components of the Yoshida et al. (2011) task were incorporated, but in which the partner's social nature was manipulated instead. Specifically, the subject macaques were faced with either a real macaque, a filmed macaque, or a filmed object (stick) and performed a turn-taking choice task. At the level of individual neurons, many of the self-, partner-, and mirror-type neurons in the MPFC exhibited different action-related activity depending on the partner condition. At a population level, however, only the partner-type neurons exhibited a systematic bias in favor of the partner with a higher social nature (i.e., real macaque > filmed macaque > filmed object). Such a population-level bias in favor of the social partner was much less evident in the PMv.

The authors also analyzed local field potentials to clarify the nature of mu suppression, a decrease in neural signals in the alpha band over the human sensorimotor cortex (Muthukumaraswamy et al. 2004, Oberman et al. 2005) and in the high beta band in the macaque

PMv (Bimbi et al. 2018, Caggiano et al. 2015, Kilner et al. 2014, Waldert et al. 2015), during both action execution and action observation. Mu suppression has attracted attention as a neural marker of mirror neuron activity (Muthukumaraswamy et al. 2004) and as a potential diagnostic tool for ASD (Oberman et al. 2005). In the new study (Ninomiya et al. 2020), two notable findings were obtained. First, mu suppression also occurred in the MPFC. Second, the magnitude of mu suppression in the MPFC, but not in the PMv, increased as the target's social nature increased. These electrophysiological findings obtained at both single-neuron and field potential levels support the social nature of MPFC activity, which is consistent with a previous study using neuroimaging methods (Sliwa & Freiwald 2017).

The social nature of macaque MPFC activity was also demonstrated in behavioral aspects other than action. Specifically, regarding motor-intention coding, significantly fewer other-choice predictive neurons in the dorsal ACC fired when a partner animal was in a separate room compared to when the partner was directly in front of the subject (Haroush & Williams 2015). Regarding reward coding, the activity of MPFC neurons selective for others' reward probability was significantly lower when a partner animal was replaced with a nonsocial object such as a water-collecting bottle (Noritake et al. 2018).

MENTAL MODELS IN THE MPFC: EVIDENCE AND HYPOTHESIS

Together, the above findings obtained in macaques indicate that individual MPFC neurons can selectively encode behavioral variables for the self and others. As discussed, such self–other distinction helps individuals to reflect efficiently on others and to predict their emotions and actions. To accomplish this mentalizing role and thereby moderate brain representations of the self and others, I hypothesize that a set of mental models, similar to simulation functions, exist in the MPFC (**Figure 2**). I conjecture that other-mentalizing is implemented by mental models that predict others' behavior using other-specific behavioral variables. The mental model processes in the MPFC may require functional coordination with other brain regions such as the temporoparietal junction (Denny et al. 2012, Hill et al. 2017), another key region in the human mentalizing networks (Van Overwalle & Baetens 2009).

Van Overwalle & Baetens (2009) performed a meta-analysis of neuroimaging studies in humans, demonstrating that other-mentalizing preferentially recruits more dorsal divisions of the MPFC, whereas self-mentalizing preferentially recruits more ventral MPFC divisions. In support of this finding, another meta-analysis conducted by a different research group demonstrated that whether subjects performed self-mentalizing or other-mentalizing was predictable on the basis of the *z*-coordinate of activation foci in the ventral or dorsal MPFC, respectively, in humans (Denny et al. 2012). The spatial segregation or, more precisely, the spatial gradient along the dorsoventral axis in the human MPFC is consistent with the macaque study, reporting a biased distribution of the partner-type neurons and the self-type neurons in the dorsal MPFC and the ventral MPFC, respectively (Yoshida et al. 2011). This spatial gradient was also observed when human subjects monitored the welfare (Sul et al. 2015) and performance ability (Wittmann et al. 2016) of the self and others, and it provides a critical viewpoint when interpreting aberrant MPFC activity seen in people with ASD (see below).

It should be noted that the activation foci in the above-mentioned human studies are not determined solely by a categorical dichotomy of the self versus others. They are also affected by a perceived overlap between self and other. For example, mentalizing a similar other, in terms of sociopolitical views, activated the ventral MPFC, while mentalizing a dissimilar other activated the dorsal MPFC (Mitchell et al. 2005, 2006). Thus, which MPFC division is activated during other-mentalizing can be influenced by the psychological distance between self and other.

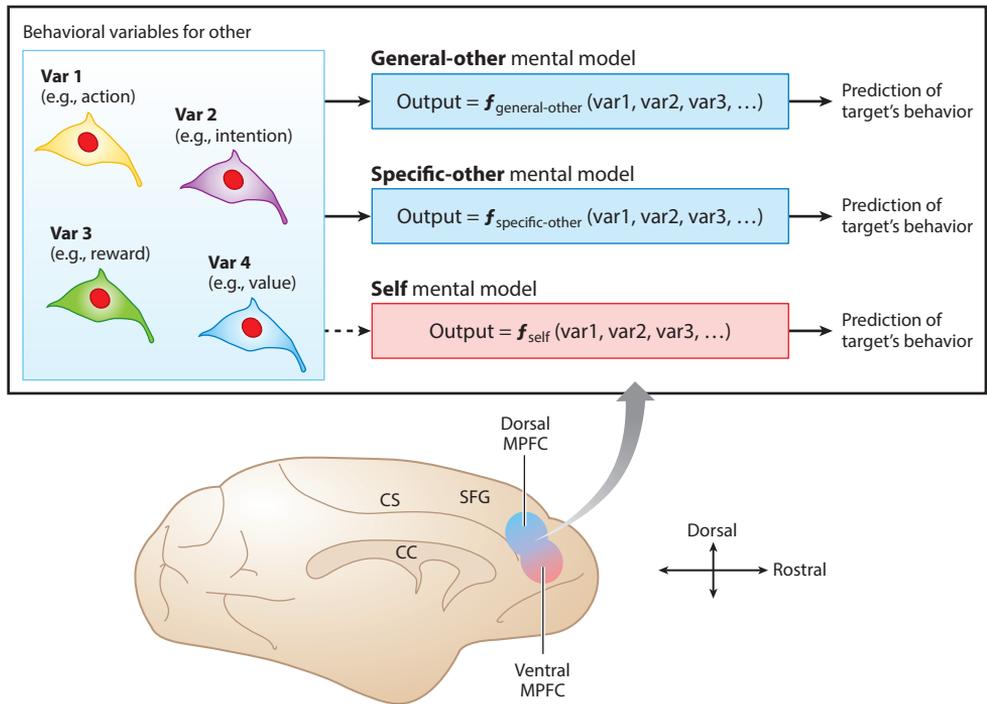


Figure 2

Schematic illustration of mental models in the MPFC. The dorsal MPFC is shown in blue and the ventral MPFC in pink on the medial surface of the macaque brain. Other mental models are formed in the dorsal MPFC and self mental models in the ventral MPFC. Individual MPFC neurons selectively encode behavioral variables for others (*left*), which are used for different mental models (*middle*). These models predict how others would feel and behave in a given context (*right*). Note that others' behavioral variables can be used for the self mental model (*dashed arrow*) when a subject has prior knowledge that a target individual shares the subject's perspective or when the subject cannot form other mental models. Abbreviations: CC, corpus callosum; CS, cingulate sulcus; MPFC, medial prefrontal cortex; SFG, superior frontal gyrus; var, variable.

Inspired by the above findings and discussion, I propose a speculative working hypothesis that the MPFC encodes a mental model that is most adaptive for reflecting on a target individual under a given context. If the subject is asked to reflect on an unfamiliar or anonymous target, the MPFC operates a general-other mental model (see the top of **Figure 2**). By referencing other-specific behavioral variables represented in MPFC neurons, this model outputs how others in general might feel and behave in a given context. The general-other mental model is acquired and updated based on others' default behavioral characteristics via social experiences. Alternately, if the subject is asked to reflect on a quite familiar target such as close relatives, friends, and colleagues, the MPFC encodes specific-other mental models, each of which is uniquely recruited for the corresponding individual (see the middle of **Figure 2**). Developed via a long-term, intimate relationship, the specific-other mental model outputs how the target in question might feel and behave given the behavioral variables encoded in MPFC neurons. Use of the specific-other mental model, instead of the general-other mental model, enables more accurate prediction of the target's future behavior. Finally, if the subject is asked to reflect on an unfamiliar or anonymous target, but is informed that psychobehavioral characteristics of the target are similar to those of the subject, the MPFC encodes a self mental model (see the bottom of **Figure 2**). The self

mental model outputs how I (as the subject) might feel and behave if I were the target with the given behavioral variables. In preschool children whose other-mentalizing ability is not fully developed (Callaghan et al. 2005, Wellman et al. 2001), use of the self mental model might be a default mechanism, owing to privileged access to self-related processing (Lockwood et al. 2018, Sui & Humphreys 2015). Considering the spatial gradient highlighted above, the general-other and specific-other mental models are likely to be formed mainly in the dorsal MPFC, while the self mental model is mainly in the ventral MPFC.

This working hypothesis invites two critical questions. The first question concerns how different specific-other mental models are represented in the dorsal MPFC. Suppose a person has three close friends, each of whom has unique traits, and the person is asked to mentalize each of them. One possibility is that each mental model is represented at a unique subregion, and therefore different subregions are activated. Another possibility is that a shared region is activated, but spatiotemporal activation patterns are different within the region, which may be distinguishable using multivoxel pattern analysis (Norman et al. 2006). These two possibilities are testable using functional neuroimaging. A related question in macaques is whether the activity of partner/other-type neurons is selective for the behavior of a particular individual, and if so, how such neurons encoding different individuals are distributed in the MPFC. The second question concerns the neural circuitry and computation underlying the formation and updating of other mental models. Here, cortical microcircuits for predictive coding (Bastos et al. 2012) may serve as a reference. A proposed mechanism for predictive coding is that top-down predictions are compared with bottom-up sensory information, which generates a prediction error if there is a mismatch (Bastos et al. 2012, Rao & Ballard 1999). The prediction error is then used to update higher-level representations on which predictions are based. Whether a similar mechanism can apply to the prediction of others' behavior is an interesting question. One possibility is that other mental models are updated on the basis of error in the prediction of others' behavior. The neural signature of other-action prediction errors is identified in the dorsal MPFC in both humans (Suzuki et al. 2012) and macaques (Ninomiya et al. 2020, Yoshida et al. 2012). MPFC regions involved in other-mentalizing, that is, mainly BA8 and BA9, are granular cortex, just like visual cortical areas for which predictive coding was originally developed. A neurobiologically plausible account of how other mental models are formed and updated could be provided by combining layer-specific local field potential analysis (Sajad et al. 2019), pathway-selective intervention (Inoue et al. 2015, Kinoshita et al. 2012, Ninomiya et al. 2020), and computational modeling in the macaque. A rare opportunity to record neuronal activity from neurosurgical patients performing a theory-of-mind task (Jamali et al. 2021) will also help us understand the representation of other mental models in the human brain.

CLINICAL CONSIDERATIONS ON THE SOCIAL MODERATOR ROLE: EVIDENCE AND HYPOTHESIS

ASD is a neurodevelopmental disorder caused by complex associations between genetic and environmental factors (Kim & Leventhal 2015). Individuals with ASD show persistent deficits in social communication and interactions (APA 2013) and exhibit deficient performance in other-mentalizing as assessed using false-belief tasks (Baron-Cohen et al. 1985). These signs suggest that ASD can be considered a representative clinical condition exhibiting difficulty in moderating brain representations of self and others.

Happé et al. (1996) were the first to report aberrant activation patterns in the autistic brain during other-mentalizing. They conducted positron emission tomography imaging and observed that people with Asperger's syndrome, a mild subtype of ASD, exhibited a lack of activation in

the dorsal MPFC, the border between BA8 and BA9, whereas normal volunteers showed clear activation. Instead, there was an increase in activity in more ventral MPFC regions (**Figure 3a**). Later work reported consistent findings; the ventral MPFC and middle ACC (regions with greater responses during self-mentalizing than during other-mentalizing in normal control subjects) responded preferentially during other-mentalizing in people with ASD (Lombardo et al. 2010). The authors hypothesized that the neural circuitry that processes self-information is atypically organized in people with ASD.

While several studies suggest deficient self-referential processing in ASD, known as the absent-self hypothesis (Gillespie-Smith et al. 2018, Nijhof & Bird 2019), a possible alternative view posits that aberrant activation in the ventral MPFC during other-mentalizing is a consequence of malfunctioning within the dorsal MPFC (**Figure 3b**). I hypothesize that the dorsal MPFC is faulty in two respects. First, individual neurons in the dorsal MPFC are not capable of selectively encoding others' behavioral variables. In support of this view, partner-type neurons were virtually absent in a macaque with a spontaneous autistic phenotype (Yoshida et al. 2016). In this autistic-like macaque, loss-of-function mutations were identified in two genes linked to human psychiatric disorders (Yoshida et al. 2016), and cortical-layer construction was disorganized (Iritani et al. 2018). Second, the dorsal MPFC is not capable of developing other mental models. Deficits in performing theory-of-mind tasks were consistently reported to be the cardinal clinical sign in ASD (Baron-Cohen et al. 1985, Frith 2001, Velikonja et al. 2019). Under these conditions, the dorsal MPFC cannot come into play when other-mentalizing is required. Instead, only the self mental models in the ventral MPFC are available (**Figure 3b**). In this scheme, aberrant activation in the ventral MPFC during other-mentalizing reflects a compensatory mechanism. As a consequence, individuals with ASD may have self-centered perspectives and may experience difficulty separating aspects of others from themselves (Lee & Hobson 2006, Lombardo et al. 2010, Mitchell & O'Keefe 2008). I tentatively refer to this idea of deficient other-related processing in the dorsal MPFC as the broken other hypothesis.

One might ask whether this broken other hypothesis can coexist with the so-called broken mirror hypothesis. The broken mirror hypothesis posits that people with ASD have a dysfunctional mirror neuron system (Ramachandran & Oberman 2006). Two lines of studies in macaques suggest that the two hypotheses are not mutually exclusive. First, in the MPFC of the autistic macaque described above, mirror-type neurons were also rarely observed (Yoshida et al. 2016). Second, selective blockade of the pathway from the PMv to the dorsal MPFC (including BA9 and pre-SMA), using a double viral vector infection technique (Kinoshita et al. 2012), caused severe performance errors only after the partner's turn in the above-mentioned turn-taking choice task, suggesting the emergence of deficits in monitoring and in utilizing behavioral signals derived from observed, but not executed, actions (Ninomiya et al. 2020). The degree of this deficit was most pronounced when the task was performed with a real macaque compared to when the task was performed with a filmed macaque or a filmed object. The animals in both studies developed a self-centered behavioral strategy—either spontaneously (Yoshida et al. 2016) or after intervention (Ninomiya et al. 2020)—relying only on the information about one's own actions and the resulting consequences. These findings raise an interesting possibility: What might be broken in the autistic brain is the functional coordination between the dorsal MPFC and PMv. These two cortical regions have rich anatomical connections in macaques (Albertini et al. 2020, Bruni et al. 2018). In neurotypical subjects, coherent activity between the dorsal MPFC and PMv is markedly enhanced in response to observed actions, and causal information flow increases from the PMv to the dorsal MPFC, especially when the biological nature of observed actions is high (Ninomiya et al. 2020). These findings lead to the fascinating hypothesis that the social moderator role of the MPFC is underpinned by and strengthened via functional coordination with the PMv.

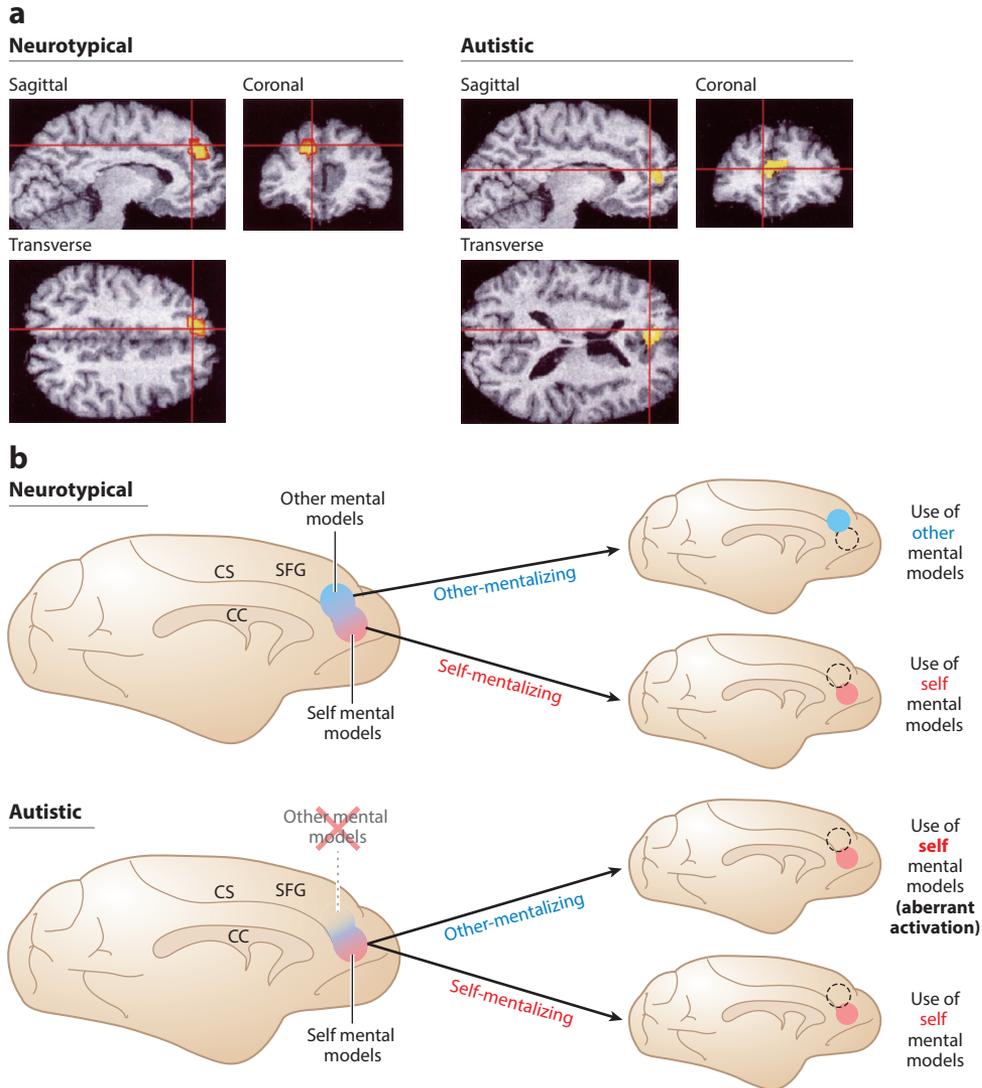


Figure 3

Aberrant activation of the ventral MPFC in the autistic brain during other-mentalizing. (a) In neurotypical subjects, the dorsal MPFC is activated during other-mentalizing (*left three panels*), while in autistic subjects, the ventral MPFC is activated during other-mentalizing (*right three panels*). Red crosshairs denote the location of the peak activation in neurotypical (coordinates $-12, 36, 36$) and autistic (coordinates $-10, 44, 16$) subjects, respectively. Panel *a* adapted from Happe et al. (1996) with permission from Wolters Kluwer Health, Inc. (b) Schematic illustration of MPFC activation during other-mentalizing and self-mentalizing. In neurotypical subjects (*top*), other mental models in the dorsal MPFC and self mental models in the ventral MPFC are both functional. In autistic subjects (*bottom*), the dorsal MPFC is dysfunctional, which induces compensatory recruitment of the ventral MPFC during other-mentalizing (aberrant activation). Abbreviations: CC, corpus callosum; CS, cingulate sulcus; MPFC, medial prefrontal cortex; SFG, superior frontal gyrus.

CONCLUDING REMARKS

Although research into the role of the MPFC in behavioral coordination originally started in nonsocial contexts, it is now widely accepted that the MPFC constitutes a critical node in the primate social brain. As the cortical crossroad where self-specific and other-specific signals meet, the MPFC organizes the behavior of oneself with proper consideration of the behavior of others. The social nature of MPFC neurons is demonstrated by their selective or preferential responses toward the behavior of real biological agents compared to nonbiological objects. These observations validate the view that the MPFC plays a pivotal role in moderating brain representations of self and other, thereby facilitating productive social exchanges. This social moderator account of MPFC function is backed by cognitive-behavioral manifestations in people with ASD, a clinical condition characterized by inferior other-mentalizing capacities and heightened self-centered perspectives. This shift from other-regarding to self-regarding cognition occurs in parallel with aberrant activation in the ventral MPFC during other-mentalizing. In my view, such aberrant activation emerges as a consequence of, and compensation for, malfunctioning within the dorsal MPFC, in which other mental models are normally formed. Despite being a challenging issue to study, future studies should explore the way in which mental models in the MPFC are implemented by ensembles of neural activities and how other brain regions might contribute to such mental model operations.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

The author is grateful to Drs. Taihei Ninomiya, Atsushi Noritake, and Saeka Tomatsu for helpful discussions and to Ms. Ayumi Shibata for artwork. This research was supported by the Japan Agency for Medical Research and Development under grant numbers JP20dm0107145 (to M.I.) and JP20dm0307005 (to M.I.).

LITERATURE CITED

- Albertini D, Gerbella M, Lanzilotto M, Livi A, Maranesi M, et al. 2020. Connectional gradients underlie functional transitions in monkey pre-supplementary motor area. *Prog. Neurobiol.* 184:101699
- Alexander WH, Brown JW. 2011. Medial prefrontal cortex as an action-outcome predictor. *Nat. Neurosci.* 14:1338–44
- Amodio DM, Frith CD. 2006. Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* 7:268–77
- Anderson JR, Gallup GG Jr. 2015. Mirror self-recognition: a review and critique of attempts to promote and engineer self-recognition in primates. *Primates* 56:317–26
- APA (Am. Psychiatr. Assoc.). 2013. *Diagnostic and Statistical Manual of Mental Disorders*. Arlington, VA: Am. Psychiatr. Publ. 5th ed.
- Apps MA, Lesage E, Ramnani N. 2015. Vicarious reinforcement learning signals when instructing others. *J. Neurosci.* 35:2904–13
- Apps MA, Ramnani N. 2014. The anterior cingulate gyrus signals the net value of others' rewards. *J. Neurosci.* 34:6190–200
- Ballesta S, Duhamel JR. 2015. Rudimentary empathy in macaques' social decision-making. *PNAS* 112:15516–21

- Baron-Cohen S, Leslie AM, Frith U. 1985. Does the autistic child have a “theory of mind”? *Cognition* 21:37–46
- Baron-Cohen S, Ring HA, Wheelwright S, Bullmore ET, Brammer MJ, et al. 1999. Social intelligence in the normal and autistic brain: an fMRI study. *Eur. J. Neurosci.* 11:1891–98
- Basile BM, Schafroth JL, Karaskiewicz CL, Chang SWC, Murray EA. 2020. The anterior cingulate cortex is necessary for forming prosocial preferences from vicarious reinforcement in monkeys. *PLoS Biol.* 18:e3000677
- Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ. 2012. Canonical microcircuits for predictive coding. *Neuron* 76:695–711
- Bault N, Joffily M, Rustichini A, Coricelli G. 2011. Medial prefrontal cortex and striatum mediate the influence of social comparison on the decision process. *PNAS* 108:16044–49
- Behrens TE, Hunt LT, Woolrich MW, Rushworth MF. 2008. Associative learning of social value. *Nature* 456:245–49
- Behrens TE, Woolrich MW, Walton ME, Rushworth MF. 2007. Learning the value of information in an uncertain world. *Nat. Neurosci.* 10:1214–21
- Ben-Ze’ev A. 2000. *The Subtlety of Emotions*. Cambridge, MA: MIT Press
- Bimbi M, Festante F, Coude G, Vanderwert RE, Fox NA, Ferrari PF. 2018. Simultaneous scalp recorded EEG and local field potentials from monkey ventral premotor cortex during action observation and execution reveals the contribution of mirror and motor neurons to the mu-rhythm. *Neuroimage* 175:22–31
- Botvinick M, Nystrom LE, Fissell K, Carter CS, Cohen JD. 1999. Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature* 402:179–81
- Bruni S, Gerbella M, Bonini L, Borra E, Coude G, et al. 2018. Cortical and subcortical connections of parietal and premotor nodes of the monkey hand mirror neuron network. *Brain Struct. Funct.* 223:1713–29
- Buunk AP, Gibbons FX. 2006. Social comparison orientation: a new perspective on those who do and those who don’t compare with others. In *Social Comparison and Social Psychology: Understanding Cognition, Intergroup Relations and Culture*, ed. S Guimond, pp. 15–21. Cambridge, UK: Cambridge Univ. Press
- Byrne RW, Whiten A. 1988. *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes, and Humans*. Oxford, UK: Clarendon Press
- Caggiano V, Fleischer F, Pomper JK, Giese MA, Thier P. 2016. Mirror neurons in monkey premotor area F5 show tuning for critical features of visual causality perception. *Curr. Biol.* 26:3077–82
- Caggiano V, Giese M, Thier P, Casile A. 2015. Encoding of point of view during action observation in the local field potentials of macaque area F5. *Eur. J. Neurosci.* 41:466–76
- Callaghan T, Rochat P, Lillard A, Claux ML, Odden H, et al. 2005. Synchrony in the onset of mental-state reasoning: evidence from five cultures. *Psychol. Sci.* 16:378–84
- Castelli F, Happe F, Frith U, Frith C. 2000. Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *Neuroimage* 12:314–25
- Chang L, Fang Q, Zhang S, Poo MM, Gong N. 2015. Mirror-induced self-directed behaviors in rhesus monkeys after visual-somatosensory training. *Curr. Biol.* 25:212–17
- Chang L, Zhang S, Poo MM, Gong N. 2017. Spontaneous expression of mirror self-recognition in monkeys after learning precise visual-proprioceptive association for mirror images. *PNAS* 114:3258–63
- Chang SW. 2013. Coordinate transformation approach to social interactions. *Front. Neurosci.* 7:147
- Chang SW, Garipey JF, Platt ML. 2013. Neuronal reference frames for social decisions in primate frontal cortex. *Nat. Neurosci.* 16:243–50
- Chang SW, Winecoff AA, Platt ML. 2011. Vicarious reinforcement in rhesus macaques (*Macaca mulatta*). *Front. Neurosci.* 5:27
- Coricelli G, Nagel R. 2009. Neural correlates of depth of strategic reasoning in medial prefrontal cortex. *PNAS* 106:9163–68
- Dal Monte O, Chu CCJ, Fagan NA, Chang SWC. 2020. Specialized medial prefrontal-amygdala coordination in other-regarding decision preference. *Nat. Neurosci.* 23:565–74
- Denny BT, Kober H, Wager TD, Ochsner KN. 2012. A meta-analysis of functional neuroimaging studies of self- and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. *J. Cogn. Neurosci.* 24:1742–52
- di Pellegrino G, Fadiga L, Fogassi L, Gallese V, Rizzolatti G. 1992. Understanding motor events: a neurophysiological study. *Exp. Brain Res.* 91:176–80

- Dum RP, Strick PL. 1991. The origin of corticospinal projections from the premotor areas in the frontal lobe. *J. Neurosci.* 11:667–89
- Dunbar RIM. 2003. The social brain: mind, language, and society in evolutionary perspective. *Annu. Rev. Anthropol.* 32:161–81
- Dvash J, Gilam G, Ben-Ze'ev A, Hendler T, Shamay-Tsoory SG. 2010. The envious brain: the neural basis of social comparison. *Hum. Brain Mapp.* 31:1741–50
- Falcone R, Cirillo R, Ferraina S, Genovesio A. 2017. Neural activity in macaque medial frontal cortex represents others' choices. *Sci. Rep.* 7:12663
- Falkenstein M, Hohnsbein J, Hoormann J, Blanke L. 1990. Effects of errors in choice reaction tasks on the ERP under focused and divided attention. In *Psychophysiological Brain Research*, ed. CHM Brunia, AWK Gaillard, A Kok, pp. 192–95. Tilburg, Neth.: Tilburg Univ. Press
- Festinger L. 1954. A theory of social comparison processes. *Hum. Relat.* 7:117–40
- Fischer AG, Ullsperger M. 2013. Real and fictive outcomes are processed differently but converge on a common adaptive mechanism. *Neuron* 79:1243–55
- Fletcher PC, Happe F, Frith U, Baker SC, Dolan RJ, et al. 1995. Other minds in the brain: a functional imaging study of “theory of mind” in story comprehension. *Cognition* 57:109–28
- Frith CD, Frith U. 1999. Interacting minds—a biological basis. *Science* 286:1692–95
- Gallagher HL, Frith CD. 2003. Functional imaging of ‘theory of mind’. *Trends Cogn. Sci.* 7:77–83
- Frith U. 2001. Mind blindness and the brain in autism. *Neuron* 32:969–79
- Gallagher HL, Happe F, Brunswick N, Fletcher PC, Frith U, Frith CD. 2000. Reading the mind in cartoons and stories: an fMRI study of ‘theory of mind’ in verbal and nonverbal tasks. *Neuropsychologia* 38:11–21
- Gallese V, Fadiga L, Fogassi L, Rizzolatti G. 1996. Action recognition in the premotor cortex. *Brain* 119(2):593–609
- Gilbert SJ, Spengler S, Simons JS, Steele JD, Lawrie SM, et al. 2006. Functional specialization within rostral prefrontal cortex (area 10): a meta-analysis. *J. Cogn. Neurosci.* 18:932–48
- Gillespie-Smith K, Ballantyne C, Branigan HP, Turk DJ, Cunningham SJ. 2018. The I in autism: Severity and social functioning in autism are related to self-processing. *Br. J. Dev. Psychol.* 36:127–41
- Goel V, Grafman J, Sadato N, Hallett M. 1995. Modeling other minds. *Neuroreport* 6:1741–46
- Grabenhorst F, Baez-Mendoza R, Genest W, Deco G, Schultz W. 2019. Primate amygdala neurons simulate decision processes of social partners. *Cell* 177:986–98.e15
- Hampton AN, Bossaerts P, O’Doherty JP. 2008. Neural correlates of mentalizing-related computations during strategic interactions in humans. *PNAS* 105:6741–46
- Happe F, Ehlers S, Fletcher P, Frith U, Johansson M, et al. 1996. ‘Theory of mind’ in the brain. Evidence from a PET scan study of Asperger syndrome. *Neuroreport* 8:197–201
- Haroush K, Williams ZM. 2015. Neuronal prediction of opponent’s behavior during cooperative social interchange in primates. *Cell* 160:1233–45
- Hayashi T, Akikawa R, Kawasaki K, Egawa J, Minamimoto T, et al. 2020. Macaques exhibit implicit gaze bias anticipating others’ false-belief-driven actions via medial prefrontal cortex. *Cell Rep.* 30:4433–44.e5
- Herrmann E, Call J, Hernandez-Lloreda MV, Hare B, Tomasello M. 2007. Humans have evolved specialized skills of social cognition: the cultural intelligence hypothesis. *Science* 317:1360–66
- Hill CA, Suzuki S, Polania R, Moisa M, O’Doherty JP, Ruff CC. 2017. A causal account of the brain network computations underlying strategic social behavior. *Nat. Neurosci.* 20:1142–49
- Holroyd CB, Coles MGH. 2002. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.* 109:679–709
- Inoue KI, Takada M, Matsumoto M. 2015. Neuronal and behavioural modulations by pathway-selective optogenetic stimulation of the primate oculomotor system. *Nat. Commun.* 6:8378
- Iritani S, Torii Y, Habuchi C, Sekiguchi H, Fujishiro H, et al. 2018. The neuropathological investigation of the brain in a monkey model of autism spectrum disorder with ABCA13 deletion. *Int. J. Dev. Neurosci.* 71:130–39
- Ishida H, Inoue KI, Takada M. 2018. Multisynaptic projections from the amygdala to the ventral premotor cortex in macaque monkeys: anatomical substrate for feeding behavior. *Front. Neuroanat.* 12:3
- Isoda M. 2016. Understanding intentional actions from observers’ viewpoints: a social neuroscience perspective. *Neurosci. Res.* 112:1–9

- Isoda M, Hikosaka O. 2007. Switching from automatic to controlled action by monkey medial frontal cortex. *Nat. Neurosci.* 10:240–48
- Isoda M, Noritake A, Ninomiya T. 2018. Development of social systems neuroscience using macaques. *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.* 94:305–23
- Jamali M, Grannan BL, Fedorenko E, Saxe R, Báez-Mendoza R, Williams ZM. 2021. Single-neuronal predictions of others' beliefs in humans. *Nature* 591:610–14
- Jiang Y, Platt ML. 2018. Oxytocin and vasopressin flatten dominance hierarchy and enhance behavioral synchrony in part via anterior cingulate cortex. *Sci. Rep.* 8:8201
- Jocham G, Neumann J, Klein TA, Danielmeier C, Ullsperger M. 2009. Adaptive coding of action values in the human rostral cingulate zone. *J. Neurosci.* 29:7489–96
- Kedia G, Mussweiler T, Linden DE. 2014. Brain mechanisms of social comparison and their influence on the reward system. *Neuroreport* 25:1255–65
- Kilner JM, Kraskov A, Lemon RN. 2014. Do monkey F5 mirror neurons show changes in firing rate during repeated observation of natural actions? *J. Neurophysiol.* 111:1214–26
- Kim YS, Leventhal BL. 2015. Genetic epidemiology and insights into interactive genetic and environmental effects in autism spectrum disorders. *Biol. Psychiatry* 77:66–74
- Kinoshita M, Matsui R, Kato S, Hasegawa T, Kasahara H, et al. 2012. Genetic dissection of the circuit for hand dexterity in primates. *Nature* 487:235–38
- Klavir O, Genud-Gabai R, Paz R. 2013. Functional connectivity between amygdala and cingulate cortex for adaptive aversive learning. *Neuron* 80:1290–300
- Lamm C, Bukowski H, Silani G. 2015. From shared to distinct self–other representations in empathy: evidence from neurotypical function and socio-cognitive disorders. *Phil. Trans. R. Soc. B* 371:20150083
- Lee A, Hobson RP. 2006. Drawing self and others: How do children with autism differ from those with learning difficulties? *Br. J. Dev. Psychol.* 24:547–65
- Livi A, Lanzilotto M, Maranesi M, Fogassi L, Rizzolatti G, Bonini L. 2019. Agent-based representations of objects and actions in the monkey pre-supplementary motor area. *PNAS* 116:2691–700
- Lockwood PL, Wittmann MK, Apps MAJ, Klein-Flügge MC, Crockett MJ, et al. 2018. Neural mechanisms for learning self and other ownership. *Nat. Commun.* 9:4747
- Lombardo MV, Chakrabarti B, Bullmore ET, Sadek SA, Pasco G, et al. 2010. Atypical neural self-representation in autism. *Brain* 133:611–24
- Luo Y, Eickhoff SB, Hetu S, Feng C. 2018. Social comparison in the brain: a coordinate-based meta-analysis of functional brain imaging studies on the downward and upward comparisons. *Hum. Brain Mapp.* 39:440–58
- Matelli M, Luppino G, Rizzolatti G. 1991. Architecture of superior and mesial area 6 and the adjacent cingulate cortex in the macaque monkey. *J. Comp. Neurol.* 311:445–62
- Matsumoto M, Matsumoto K, Abe H, Tanaka K. 2007. Medial prefrontal cell activity signaling prediction errors of action values. *Nat. Neurosci.* 10:647–56
- Matsuzaka Y, Tanji J, Mushiake H. 2016. Representation of behavioral tactics and tactics-action transformation in the primate medial prefrontal cortex. *J. Neurosci.* 36:5974–87
- McGuire JT, Nassar MR, Gold JI, Kable JW. 2014. Functionally dissociable influences on learning rate in a dynamic environment. *Neuron* 84:870–81
- Mitchell JP, Banaji MR, Macrae CN. 2005. The link between social cognition and self-referential thought in the medial prefrontal cortex. *J. Cogn. Neurosci.* 17:1306–15
- Mitchell JP, Macrae CN, Banaji MR. 2006. Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron* 50:655–63
- Mitchell P, O'Keefe K. 2008. Brief report: Do individuals with autism spectrum disorder think they know their own minds? *J. Autism Dev. Disord.* 38:1591–97
- Miyachi S, Lu X, Inoue S, Iwasaki T, Koike S, et al. 2005. Organization of multisynaptic inputs from prefrontal cortex to primary motor cortex as revealed by retrograde transneuronal transport of rabies virus. *J. Neurosci.* 25:2547–56
- Mobbs D, Yu R, Meyer M, Passamonti L, Seymour B, et al. 2009. A key role for similarity in vicarious reward. *Science* 324:900

- Muthukumaraswamy SD, Johnson BW, McNair NA. 2004. Mu rhythm modulation during observation of an object-directed grasp. *Brain Res. Cogn. Brain Res.* 19:195–201
- Nijhof AD, Bird G. 2019. Self-processing in individuals with autism spectrum disorder. *Autism Res.* 12:1580–84
- Ninomiya T, Noritake A, Kobayashi K, Isoda M. 2020. A causal role for frontal cortico-cortical coordination in social action monitoring. *Nat. Commun.* 11:5233
- Ninomiya T, Noritake A, Ullsperger M, Isoda M. 2018. Performance monitoring in the medial frontal cortex and related neural networks: from monitoring self actions to understanding others' actions. *Neurosci. Res.* 137:1–10
- Noritake A, Ninomiya T, Isoda M. 2018. Social reward monitoring and valuation in the macaque brain. *Nat. Neurosci.* 21:1452–62
- Noritake A, Ninomiya T, Isoda M. 2020. Representation of distinct reward variables for self and other in primate lateral hypothalamus. *PNAS* 117:5516–24
- Norman KA, Polyn SM, Detre GJ, Haxby JV. 2006. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn. Sci.* 10:424–30
- Oberman LM, Hubbard EM, McCleery JP, Altschuler EL, Ramachandran VS, Pineda JA. 2005. EEG evidence for mirror neuron dysfunction in autism spectrum disorders. *Brain Res. Cogn. Brain Res.* 24:190–98
- Oberman LM, McCleery JP, Ramachandran VS, Pineda JA. 2007. EEG evidence for mirror neuron activity during the observation of human and robot actions: toward an analysis of the human qualities of interactive robots. *Neurocomputing* 70:2194–203
- Ramachandran VS, Oberman LM. 2006. Broken mirrors: a theory of autism. *Sci. Am.* 295:62–69
- Rao RP, Ballard DH. 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2:79–87
- Ridderinkhof KR, Ullsperger M, Crone EA, Nieuwenhuis S. 2004. The role of the medial frontal cortex in cognitive control. *Science* 306:443–47
- Rilling J, Gutman D, Zeh T, Pagnoni G, Berns G, Kilts C. 2002. A neural basis for social cooperation. *Neuron* 35:395–405
- Rizzolatti G, Craighero L. 2004. The mirror-neuron system. *Annu. Rev. Neurosci.* 27:169–92
- Rizzolatti G, Fadiga L, Gallese V, Fogassi L. 1996. Premotor cortex and the recognition of motor actions. *Brain Res. Cogn. Brain Res.* 3:131–41
- Rizzolatti G, Sinigaglia C. 2016. The mirror mechanism: a basic principle of brain function. *Nat. Rev. Neurosci.* 17:757–65
- Roberts RE, Husain M. 2015. A dissociation between stopping and switching actions following a lesion of the pre-supplementary motor area. *Cortex* 63:184–95
- Rudebeck PH, Buckley MJ, Walton ME, Rushworth MF. 2006. A role for the macaque anterior cingulate gyrus in social valuation. *Science* 313:1310–12
- Rudebeck PH, Ripple JA, Mitz AR, Averbeck BB, Murray EA. 2017. Amygdala contributions to stimulus-reward encoding in the macaque medial and orbital frontal cortex during learning. *J. Neurosci.* 37:2186–202
- Sajad A, Godlove DC, Schall JD. 2019. Cortical microcircuitry of performance monitoring. *Nat. Neurosci.* 22:265–74
- Sallet J, Mars RB, Noonan MP, Neubert FX, Jbabdi S, et al. 2013. The organization of dorsal frontal cortex in humans and macaques. *J. Neurosci.* 33:12255–74
- Schall JD, Stuphorn V, Brown JW. 2002. Monitoring and control of action by the frontal lobes. *Neuron* 36:309–22
- Seid-Fatemi A, Tobler PN. 2015. Efficient learning mechanisms hold in the social domain and are implemented in the medial prefrontal cortex. *Soc. Cogn. Affect. Neurosci.* 10:735–43
- Sliwa J, Freiwald WA. 2017. A dedicated network for social interaction processing in the primate brain. *Science* 356:745–49
- Sui J, Humphreys GW. 2015. The integrative self: how self-reference integrates perception and memory. *Trends Cogn. Sci.* 19:719–28
- Sul S, Tobler PN, Hein G, Leiberg S, Jung D, et al. 2015. Spatial gradient in value representation along the medial prefrontal cortex reflects individual differences in prosociality. *PNAS* 112:7851–56

- Suzuki S, Harasawa N, Ueno K, Gardner JL, Ichinohe N, et al. 2012. Learning to simulate others' decisions. *Neuron* 74:1125–37
- Swencionis JK, Fiske ST. 2014. How social neuroscience can inform theories of social comparison. *Neuropsychologia* 56:140–46
- Tai YF, Scherfler C, Brooks DJ, Sawamoto N, Castiello U. 2004. The human premotor cortex is 'mirror' only for biological actions. *Curr. Biol.* 14:117–20
- Takahashi H, Kato M, Matsuura M, Mobbs D, Suhara T, Okubo Y. 2009. When your gain is my pain and your pain is my gain: neural correlates of envy and schadenfreude. *Science* 323:937–39
- Tanji J. 2001. Sequential organization of multiple movements: involvement of cortical motor areas. *Annu. Rev. Neurosci.* 24:631–51
- Todd AR, Hanks K, Galinsky AD, Mussweiler T. 2011. When focusing on differences leads to similar perspectives. *Psychol. Sci.* 22:134–41
- Ullsperger M, Danielmeier C, Jochem G. 2014. Neurophysiology of performance monitoring and adaptive behavior. *Physiol. Rev.* 94:35–79
- Van Overwalle F, Baetens K. 2009. Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. *Neuroimage* 48:564–84
- Velikonja T, Fett AK, Velthorst E. 2019. Patterns of nonsocial and social cognitive functioning in adults with autism spectrum disorder: a systematic review and meta-analysis. *JAMA Psychiatry* 76:135–51
- Volz KG, Schubotz RI, von Cramon DY. 2003. Predicting events of varying probability: uncertainty investigated by fMRI. *Neuroimage* 19:271–80
- Waldert S, Vigneswaran G, Philipp R, Lemon RN, Kraskov A. 2015. Modulation of the intracortical LFP during action execution and observation. *J. Neurosci.* 35:8451–61
- Wellman HM, Cross D, Watson J. 2001. Meta-analysis of theory-of-mind development: the truth about false belief. *Child Dev.* 72:655–84
- Wheatley T, Milleville SC, Martin A. 2007. Understanding animate agents: distinct roles for the social network and mirror system. *Psychol. Sci.* 18:469–74
- Wittmann MK, Kolling N, Faber NS, Scholl J, Nelissen N, Rushworth MF. 2016. Self-other mergence in the frontal cortex during cooperation and competition. *Neuron* 91:482–93
- Wittmann MK, Lockwood PL, Rushworth MFS. 2018. Neural mechanisms of social cognition in primates. *Annu. Rev. Neurosci.* 41:99–118
- Yoshida K, Go Y, Kushima I, Toyoda A, Fujiyama A, et al. 2016. Single-neuron and genetic correlates of autistic behavior in macaque. *Sci. Adv.* 2:e1600558
- Yoshida K, Saito N, Iriki A, Isoda M. 2011. Representation of others' action by neurons in monkey medial frontal cortex. *Curr. Biol.* 21:249–53
- Yoshida K, Saito N, Iriki A, Isoda M. 2012. Social error monitoring in macaque frontal cortex. *Nat. Neurosci.* 15:1307–12