CHAPTER EIGHT

Social processing by the primate medial frontal cortex

Philip T. Putnam[a,*] **and Steve W.C. Chang**[a,b,c]
[a]Department of Psychology, Yale University, New Haven, CT, United States
[b]Department of Neuroscience, Yale University School of Medicine, New Haven, CT, United States
[c]Kavli Institute for Neuroscience, Yale University School of Medicine, New Haven, CT, United States
*Corresponding author: e-mail address: philip.putnam@yale.edu

Contents

Abstract

The primate medial frontal cortex is comprised of several brain regions that are consistently implicated in regulating complex social behaviors. The medial frontal cortex is also critically involved in many non-social behaviors, such as those involved in reward, affective, and decision-making processes, broadly implicating the fundamental role of the medial frontal cortex in internally guided cognition. An essential question therefore is what unique contributions, if any, does the medial frontal cortex make to social behaviors? In this chapter, we outline several neural algorithms necessary for mediating adaptive social interactions and discuss selected evidence from behavioral neurophysiology experiments supporting the role of the medial frontal cortex

International Review of Neurobiology, Volume 158
ISSN 0074-7742
https://doi.org/10.1016/bs.irn.2020.12.003

in implementing these algorithms. By doing so, we primarily focus on research in non-human primates and examine several key attributes of the medial frontal cortex. Specifically, we review neuronal substrates in the medial frontal cortex uniquely suitable for enabling social monitoring, observational and vicarious learning, as well as predicting the behaviors of social partners. Moreover, by utilizing the three levels of organization in information processing systems proposed by Marr (1982) and recently adapted by Lockwood, Apps, and Chang (2020) for social information processing, we survey selected social functions of the medial frontal cortex through the lens of socially relevant algorithms and implementations. Overall, this chapter provides a broad overview of the behavioral neurophysiology literature endorsing the importance of socially relevant neural algorithms implemented by the primate medial frontal cortex for regulating social interactions.

1. Introduction

Social interactions cannot be understood simply through the framework of classical learning processes (Bandura & Walters, 1977; Frith & Frith, 2010; Redcay & Schilbach, 2019). Rather, highly adaptative and contingent social behaviors require specialized neural processes to observe and predict the actions of others, to learn from these actions, and make inferences about a social partner's motivations and intentions. Here we will present emerging evidence from behavioral neurophysiology research that suggests that the medial frontal cortex (MFC) performs a unique role in shaping social behaviors by performing many of these processes. Additionally, by applying the three levels of organization proposed by David Marr in his seminal 1982 book *Vision* (Marr, 1982) to social information processing (Lockwood, Apps, & Chang, 2020), we will examine in detail why certain regions within the MFC are particularly suited to support socially relevant computations with specific algorithms and implementations.

First, it is important to recognize that the MFC is just one collection of brain regions that is critically implicated in social functions out of many other collections of cortical and subcortical regions implicated in social behaviors (Adolphs, 2009; Behrens, Hunt, & Rushworth, 2009). Accumulating evidence strongly suggests that social behaviors are accomplished through the coordination of many different brain regions working in tandem (Chen & Hong, 2018; Dal Monte, Chu, Fagan, & Chang, 2020; Dölen, Darvishzadeh, Huang, & Malenka, 2013; Sliwa & Freiwald, 2017; Zhou et al., 2017). Based on existing evidence it is plausible to argue that the regions in the MFC serve as central nodes of social processing networks in the brain across multiple animal species. These MFC brain regions include

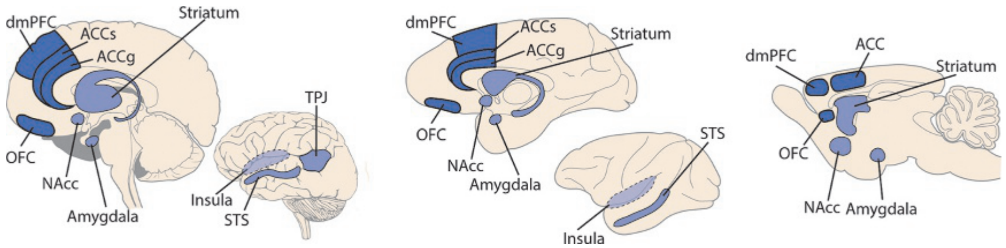

Fig. 1 Sagittal brain illustrations with selected brain structures implicated in social functions in humans (A), rhesus macaques (B), and mice (C). Highlighted are the key components of the MFC (darker shading) and other regions (lighter shading) that are implicated in various social behaviors in each species. ACCg, anterior cingulate gyrus; ACCs, anterior cingulate sulcus; dmPFC, dorsomedial prefrontal cortex; OFC, orbitofrontal cortex; NAcc, nucleus accumbens; STS, superior temporal sulcus; TPJ, temporal parietal junction.

the anterior cingulate cortex (ACC), dorsal, medial, and ventral sub-divisions of the medial prefrontal cortex (mPFC), supplemental motor area (SMA), pre-SMA, as well as the medial aspects of the orbitofrontal cortex (OFC), among others (Wise, 2008). The MFC is centrally positioned as a key node in a network of brain regions involved in social behaviors (Fig. 1), which some scholars have referred to as the "social brain" (Adolphs, 2009; Lockwood et al., 2020), and is a major cortical node of the broader limbic network (Morecraft & Van Hoesen, 1998; Porrino, Crane, & Goldman-Rakic, 1981), allowing integrations of reward and affective information critical for social functions. As part of this network (Fig. 1), the MFC is highly interconnected with other frontal regions (Carmichael & Price, 1996), other cortical areas such as the superior temporal sulcus (STS) and temporoparietal junction (TPJ) (Bachevalier, Meunier, Lu, & Ungerleider, 1997; Carmichael & Price, 1995b), and subcortical structures such as the amygdala (Amaral & Insausti, 1992; Carmichael & Price, 1995a; Morecraft et al., 2012) and nucleus accumbens (NAcc).

Experimental evidence supports a causal role for the MFC in social behaviors (Amodio & Frith, 2006). In both humans (Bechara, Damasio, & Damasio, 2000; Saver & Damasio, 1991) and rhesus monkeys (Basile, Schafroth, Karaskiewicz, Chang, & Murray, 2020; Rudebeck, Buckley, Walton, & Rushworth, 2006), lesions within the MFC regions alter social behaviors, such as social valuation and forming prosocial preferences from vicarious social reward. However, lesion-induced social behavioral deficits often do not truly distinguish the MFC from other interconnected brain areas, such as the amygdala, that are also implicated in social behaviors (Adolphs, 2010). On the other hand, evidence from behavioral neurophysiology has revealed that the MFC may be uniquely positioned to perform several key algorithms necessary for social interactions. In this chapter, we first briefly outline selected key algorithms necessary for social interactions in humans and non-human primates, and then provide experimental evidence suggesting that the neurons in the primate MFC are performing these algorithms. Finally, we discuss why the MFC is structurally and evolutionarily suited for social functions. In doing so, we will frequently reference back to the three levels of information processing—computational, algorithmic, and implementational—in discussing how the MFC regions are relevant for social functions.

2. Key algorithms in social interactions

As a framework to analyze information processing systems, David Marr proposed three levels of organization that should be considered (Fig. 2) (Marr, 1982). The most macroscopic level is *computational*, or understanding the ultimate goal of the process, and the logic of the strategy by which it will be accomplished. Supporting this goal is the *algorithm*, the logical operation or set of rules that is applied to accomplish the goals of the system. Finally, there is the *implementation*, or how these algorithms are realized physically. As an illustration, Marr applied this framework to a cash register, where the *computational* goal is to accomplish arithmetic and thus it must "master the theory of addition." The specific *algorithms* used to accomplish this can take various forms (e.g. Arabic numerals require "adding the least significant digits first and 'carrying' if the sum exceeds 9" while other numeric representations would require different parallel algorithms) but must ultimately serve the exactly same computational goal. These algorithms must be instantiated in a physical form at the *implementational* level, which for a mechanical cash register may be tracking numerical digits by a position on a notched metal wheel or the logic gates on a computer chip inside a digital cash register. This three-level framework is particularly advantageous for studying neuroscience when considering many varied *computational goals* the brain must accomplish. For example, an individual may be engaging in cooperative social behaviors in one context, and then must quickly transition to competitive behaviors in the next. Similarly, an individual may learn a choice-outcome association in an isolated context, and then later needs to apply that knowledge in a social context. To support these varied social computations, many different algorithmic processes have been proposed, such as monitoring the actions of others in an agent-specific reference

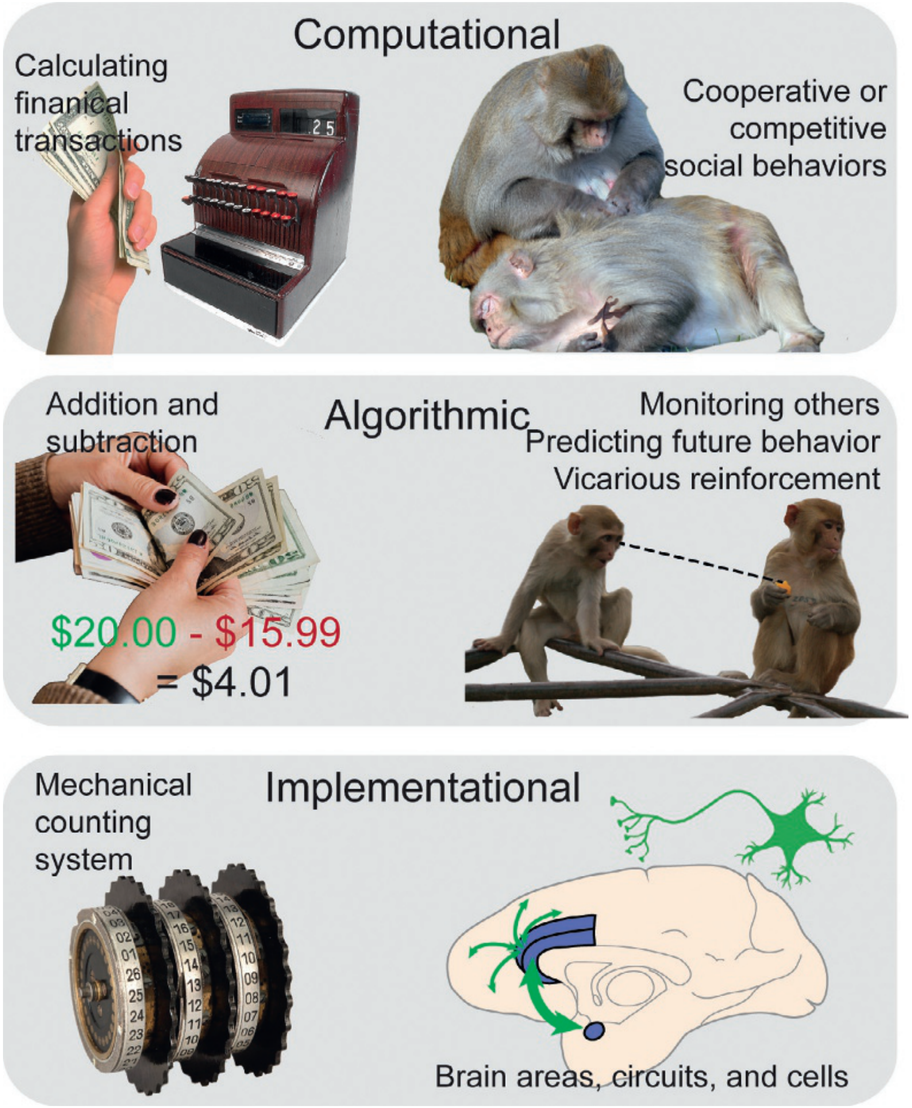

Fig. 2 Levels of organization proposed by David Marr to understand information processing systems. The computational level (top) encompasses the ultimate goal of the process, for example calculating financial transactions is the goal of a cash register (left) or engaging in cooperative social behaviors is the goal of certain brains areas. The algorithmic level (middle) is the logical processes by which this goal is accomplished. For a cash register to tally purchases it must perform arithmetic (left), while social brain regions must perform certain processes such as monitoring others or predicting their future behaviors (right). Finally, these processes must be realized at the implementational level (bottom). In a cash register the actual arithmetic is performed by a mechanical counting system (left), while in the brain different areas, circuits, and cells work together to process information (right).

frame or predicting the actions of others. These algorithms are implemented in circuits of neurons, linked by synaptic connections and oftentimes clustered in specific brain areas. Applying this framework to social functions, it has been recently suggested that some neural processes are socially specialized at the level of algorithms and implementations, and that altering the computational goals can change social specialization at these other levels (Lockwood et al., 2020). Given the enormous diversity of behaviors which the MFC is implicated in, it is likely that many of these same algorithms and implementations are shared between social and non-social behaviors and can be flexibly recruited to serve different computational goals. Below we outline a subset of the computational goals required for adaptative social behaviors, before providing experimental evidence supporting the role of the MFC at the algorithmic and implementational levels.

2.1 Monitoring the actions of others with agent-specific computation

Fundamental to any social interaction is an individual's ability to observe the actions of another, and to understand these actions from the perspective of the social partner. These actions may result in rewarding outcomes for the social partner, which the individual must process in reference to the partner in order to accurately guide their own social behaviors. This agent-specific computation, i.e. processing a reward in a "self-oriented" or "other-oriented" reference frame is theorized to be supported by a coordinate transformation framework (Chang, 2013), analogous to how distinctly referenced information is transformed from one frame of reference to another to perform sensorimotor transformation (Snyder, 2000). Even the most basic social interactions require agent-specific referencing of social information, as otherwise representing others separately from self in the brain would be impossible. These other-oriented signals may represent various domains of information; such as the overt actions of a social partner, errors in those actions, or information about rewarding or aversive outcomes received by others. At an implementational level, it would be predicted that populations of individual neurons would separately represent other-referenced information, self-referenced information, or commonly (both-referenced) information in order to perform both self- and other-regarding behaviors. Importantly, monitoring the actions of others in an agent-specific manner is also crucial to other algorithms used to shape social behaviors. For example, monitoring reward received by another is crucial for observational learning as well as deriving vicarious reinforcement.

2.2 Predicting the actions of others

Building upon the ability to monitor the actions of others, being able to predict the actions of others is a crucial computation that grants an individual the ability to engage in strategic, cooperative, or competitive social behaviors, or to learn through observing another's actions and outcomes. This prediction extends beyond merely learning the sensorimotor relationships between another's actions and associated outcomes but requires the observer to model the intentions that motivate these actions to discern what outcomes the partner values and expects (Apps, Balsters, & Ramnani, 2012; Apps, Rushworth, & Chang, 2016; Haroush & Williams, 2015; Joiner, Piva, Turrin, & Chang, 2017). An individual may choose to use these predictions in turn to guide their own behaviors, such as choosing to cooperate with others, or to compete for those same valued outcomes (Haroush & Williams, 2015). Predictions like these allow an individual to prospectively predict how their own actions will affect the behaviors of others, where errors in these predictions permit generalized algorithms, such as reinforcement learning, to guide learning about the internal states and motivations of others to support more complex forms of social learning and mentalizing (Joiner et al., 2017).

2.3 Observational learning

Observing the actions of others is a key component of how many species learn about the world. Crucially, observational learning enables the observer to learn about potential dangers or rewards from conspecifics while circumventing risks associated with direct experience. Additionally, observers can also learn about more abstract information such as the motivations or intentions of others and even specific personal characteristics of other individuals. The process of observational learning builds upon both monitoring and predicting the behaviors of another individual, necessitating both tracking the partner's actions and their associated outcomes in a self-referenced or other-referenced processing framework, and using this perspective to discern what outcomes the partner values and thus to predict their future actions. Observational learning begins early in development, with young children imitating the actions of older individuals (Bandura, Ross, & Ross, 1961), and continues into adulthood in multiple animal species, including rats (Heyes, Jaldow, Nokes, & Dawson, 1994), birds (Akins, Klein, & Zentall, 2002), monkeys (Grabenhorst, Hernádi, & Schultz, 2012;

Tomasello, Davis-Dasilva, Camak, & Bard, 1987) and humans (Bandura, 2008; Debiec & Olsson, 2017; Hill, Boorman, & Fried, 2016; Olsson, Knapska, & Lindström, 2020). Moreover, a higher degree of both attentional (Pagnotta, Laland, & Coco, 2020), autonomic (Pärnamets, Espinosa, & Olsson, 2020), or brain-to-brain (Bevilacqua et al., 2019; Dikker et al., 2017; Kostorz, Flanagin, & Glasauer, 2020) synchrony has been shown to enhance learning rates during observational learning. Such synchrony between a demonstrator and an observer may, in turn, facilitate socially specific algorithms such as theory of mind.

2.4 Theory of mind

Theory of mind is the process by which one attributes mental states, beliefs, and intentions to others. While the presence of a theory of mind in non-human primates is an ongoing and controversial question (Horschler, MacLean, & Santos, 2020), some components of theory of mind may be shared across primates (Drayton & Santos, 2016; Horschler et al., 2020). The most common test for theory of mind is the *false belief task*, where participants are tested to see if they can understand whether another social agent has a false belief about an object's location. In the false belief task, the subject observes an agent placing an object in set location, but another agent subsequently moves that object to a different location (unknown to the first agent). When the first agent returns to look for the object, the subject may anticipate that the first agent will search for the original but wrong location or the new and correct location depending on whether the subject possesses the capacity to represent a false belief. In brief, it is theorized that having a false belief representation would lead for the subject to anticipate that the first agent would search for the original yet wrong location. This specialized task thus attempts to examine if an individual can model the beliefs of another social agent distinctly from one's own, which can be argued to be the most complex form of agent-specific referencing. Human infants, even from a very young age, are generally able to use the mental states of others to explain their behaviors (Onishi & Baillargeon, 2005). However, whether non-human primates can "pass" the false belief task remains controversial (Hayashi et al., 2020; Horschler et al., 2020; Krupenye, Kano, Hirata, Call, & Tomasello, 2016). Overall, a theory of mind is a highly adaptive social algorithm, allowing an individual to model another's psychological states and can be used to predict, understand, and influence the behavior of others. Given the highly specific nature of theory of mind, it is exceedingly likely to be a socially specific algorithm that critically builds upon monitoring others in an agent-specific fashion, predicting other's behaviors, and learning about these mental states through observation.

3. Evidence from behavioral neurophysiology during social interactions

Examining the neuronal activity of MFC neurons while individuals participate in social behaviors grants a direct window into how various socially relevant algorithms are implemented in the MFC subregions. Uniquely suited to addressing these question is the emerging field of primate social neurophysiology; the investigation of the neurophysiological mechanisms underlying social interactions using a nonhuman primate model system (Chang, 2017; Gothard et al., 2017). Here we present studies from this subfield of neuroscience that provide insight into how the MFC regulate social behaviors by implementing socially relevant algorithms.

3.1 Neurons in the MFC monitor others in an agent-specific manner

Representations of value are of primary importance for the survival of a species, and thus are computed across many different brain regions including the medial and orbital frontal areas (Azab & Hayden, 2017; Hayden & Platt, 2010; Kennerley, Behrens, & Wallis, 2011; Padoa-Schioppa & Assad, 2006; Paton, Belova, Morrison, & Salzman, 2006; Rogers et al., 2004). However, the computations of value in a social context, especially those involving the process of deriving reward from others such as vicarious reward appears to be more selectively mediated by specific brain regions. Monitoring rewarding outcomes of others and estimating internal states of others is arguably the most fundamental computation that is necessary for normal social interactions. It goes without saying that such other-referenced processes are central to observational learning, empathy, and theory of mind.

To examine single neuron correlates of other-referenced reward processing in the MFC, researchers have recorded single unit activity from the anterior cingulate cortex (ACC) while monkeys were engaged in various social reward paradigms. In one of such paradigms known as the social reward allocation task (Fig. 3A), an actor monkey expressed either prosocial or antisocial preference by deciding to share with or withhold juice rewards from a conspecific recipient monkey in the same setup (Chang, Winecoff, & Platt, 2011). In one context, the actor monkey chose between donating

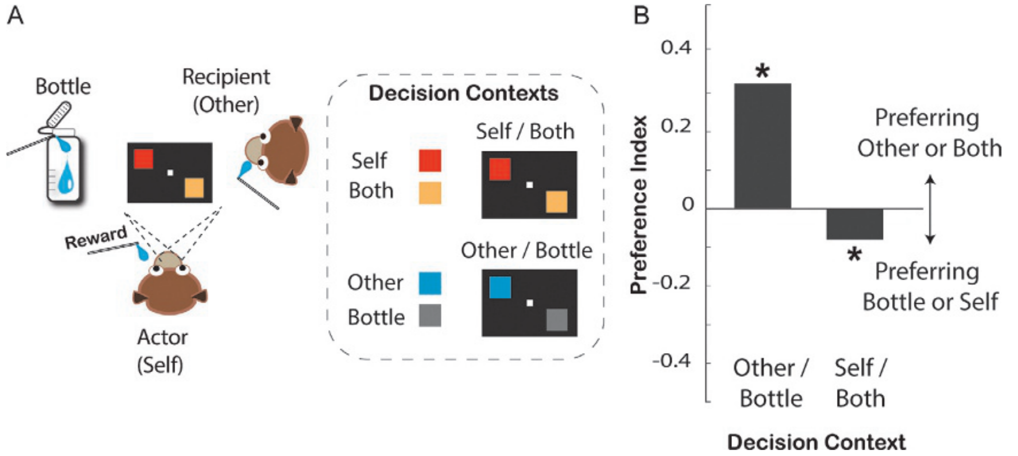

Fig. 3 Social reward allocation paradigm. (A) Task design. An actor monkey makes decisions to share with, or withhold juice rewards from, a recipient monkey seated adjacent to the actor. In the Self/Both context, the actor chooses between delivering a juice reward to himself and to both himself and the recipient. In the Other/Bottle context, the actor chooses between delivering a juice rewards to the recipient and wasting it in an empty bottle. (B) Social decision preference. Actor monkeys prefer to deliver juice rewards to the recipient over wasting them in the Other/Bottle context (exhibiting prosocial preference), whereas they prefer to deliver juice rewards just to themselves over both monkeys (exhibiting antisocial preference). *Figures adapted from Dal Monte, O., Chu, C. C. J., Fagan, N. A., & Chang, S. W. C. (2020). Specialized medial prefrontal–amygdala coordination in other-regarding decision preference.* Nature Neuroscience, 23*(4), 565–574. doi:10.1038/s41593-020-0593-y.*

juice rewards to the recipient and no one (Other/Bottle or Other/Neither context), whereas in the other context the actor chose between drinking juice rewards alone and together with the recipient (Self/Both context). Importantly, the choices in each context were paired such that choosing one or the other option did not have any direct reward impact to the actor—that is, the actor always received juice rewards in the Self/Both context, but never received juice rewards in the Other/Bottle context. Behaviorally, actor monkeys have been consistently found to prefer donating juice to the recipient over wasting, exhibiting a prosocial preference (Basile et al., 2020; Chang, Barter, Ebitz, Watson, & Platt, 2012; Chang et al., 2011, 2015; Chang, Gariépy, & Platt, 2013; Dal Monte et al., 2020) (Fig. 3B). By contrast, when the options for delivering juice rewards to the actors and both the actor and recipient monkeys were available, the actors consistently preferred drinking alone, exhibiting an antisocial preference (Chang et al., 2011, 2015; Dal Monte et al., 2020) (Fig. 3B).

Using the social reward allocation task, Chang and colleagues found that distinct neurons in the anterior cingulate gyrus (ACCg) exclusively encoded

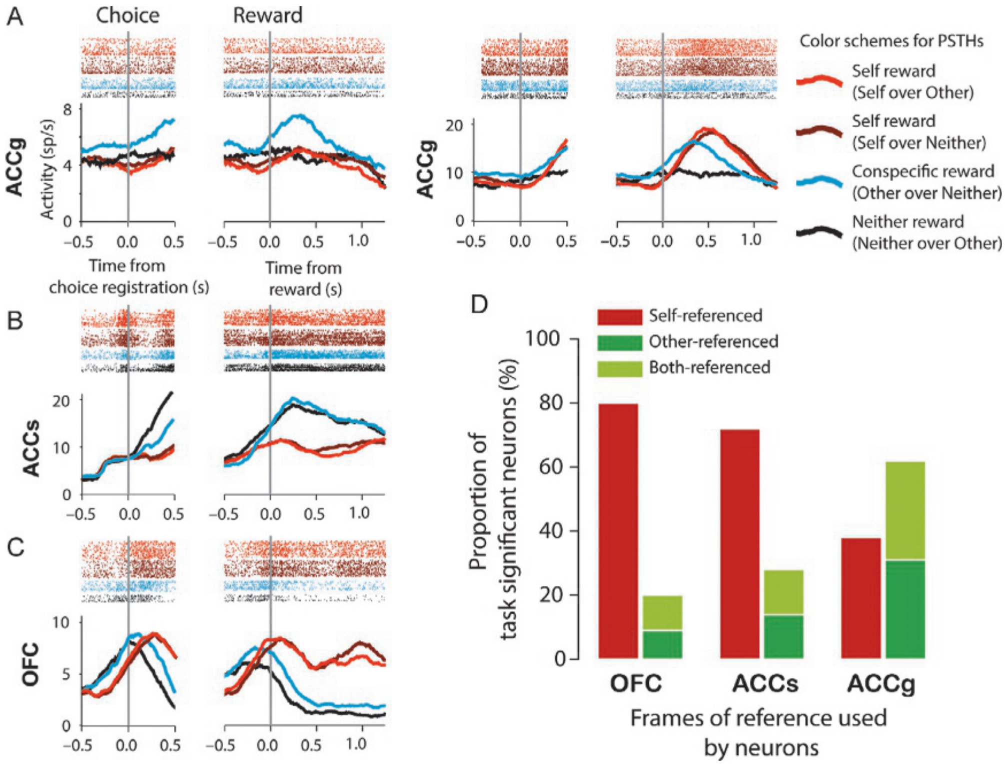

Fig. 4 Single unit activity of ACCg, ACCs, and OFC neurons collected during the social reward allocation task. (A) Single unit activity of two example ACCg cells depicted as a raster (above) and peri-stimulus time histogram (below) aligned to the time when the actor monkey made the choice (left) and when the reward was delivered (right). In the "other-referenced" ACCg cell shown on the left, note the specific increase of firing rate for the decisions to deliver juice rewards to the recipient and the reward outcome of the recipient. In the "both-referenced" ACCg cell shown on the right, note the increase in firing rate for the decisions to deliver juice rewards to the actor, the recipient, as well as both monkeys, but not for delivering juice to no one, and the associated reward outcomes. (B) Single unit activity of an example ACCs cell. In this "self-referenced" cell, note the increase in firing rate for decisions and reward outcomes that resulted in the foregoing of actor's rewards. (C) Single unit activity of an example OFC cell. In this "self-referenced" cell, note the increase in firing rate for all decisions but only for reward outcomes that resulted in the receipt of actor's rewards. (D) Plots summarizing the proportion of neurons encoding reward outcomes in self-referenced, other-referenced, or both-referenced frames. *Figures adapted from Chang, S. W. C., Gariépy, J.-F., & Platt, M. L. (2013). Neuronal reference frames for social decisions in primate frontal cortex.* Nature Neuroscience, 16*(2), 243–250. doi:10.1038/nn.3287.*

reward allocations to the other monkey (other-referenced), reward allocations to oneself (self-referenced), or both monkeys (both-referenced) (Chang et al., 2013) (Figs. 4A and 5A). Notably, the vast majority of neurons in the neighboring anterior cingulate sulcus (ACCs) signaled actor's foregone rewards by increasing activity similarly for the other monkey's rewards

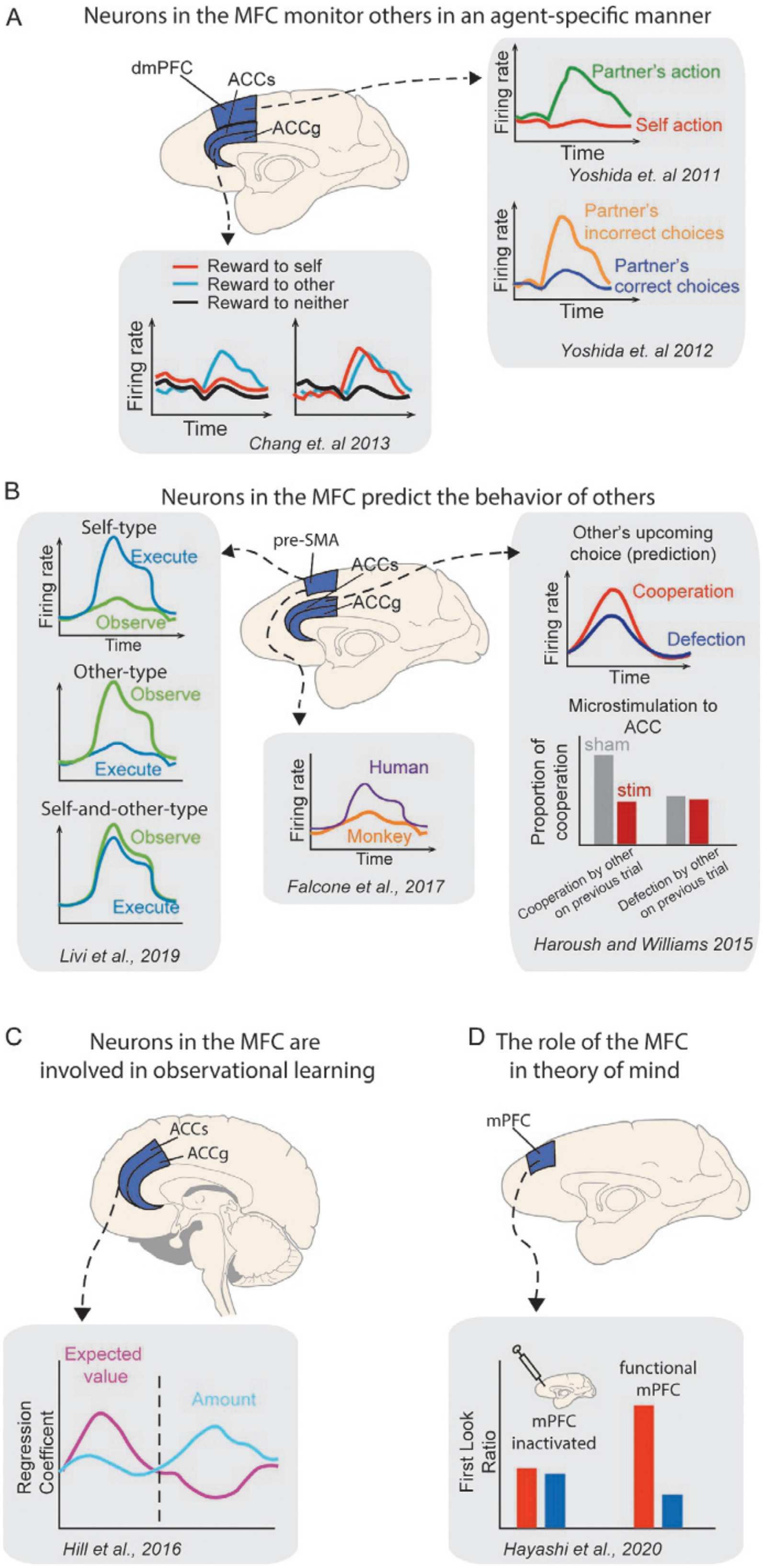

Fig. 5 Illustrations of selected findings from MFC during social interactions. *These illustrations are adapted from Gangopadhyay, P., Chawla, M., Dal Monte, O., & Chang, S. W. C. (2020). Neural mechanisms guiding social decision-making in the prefrontal-amygdala circuits.* Nature Neuroscience, *(in press).*

and wasted rewards (self-referenced) (Fig. 4B). Neurons in the orbitofrontal cortex (OFC), on the other hand, principally encoded rewards delivered to oneself (self-referenced) (Fig. 4C). Across the three neural populations, ACCg contained the highest number of neurons that were classified as either other-referenced or both-referenced, whereas the majority of ACCs and OFC neurons were classified as self-referenced (Fig. 4D). This evidence shows the ACCg is a specialized MFC region tracking rewarding outcomes of others, a key computation for monitoring others in an agent-specific manner (Apps et al., 2016; Chang et al., 2015). However, agent-specific signals are not confined to ACCg.

In the domains of performing and observing reward-guided actions, neurons in the dorsal bank of ACCs and pre-SMA have been found to exhibit robust correlates of social monitoring—that is, monitoring actions and errors of another individual. Yoshida, Saito, Iriki, and Isoda (2011) first devised an experimental paradigm in which an actor monkey chose between a pair of targets, and the partner monkey observed the actor to gain helpful information to guide their subsequent actions to obtain rewards. By recording single unit activity in the MFC, they identified a selective population of cells that encoded the actions of the other monkey in an agent-specific manner (Yoshida et al., 2011) (Fig. 5A). Notably, subsequent analyses of this same task found a special population of MFC neurons (dorsal bank of ACCs and pre-SMA) that substantially modulated their activity exclusively in response to another's errors (Yoshida, Saito, Iriki, & Isoda, 2012) (Fig. 5A). This agent-specific encoding of others' actions and errors is a core algorithmic component required for monitoring social partners. However, the social monitoring functions of MFC are not limited to the action and error monitoring domain. A recent study by Noritake, Ninomiya, and Isoda (2018) employed a social conditioning procedure where different amounts of juice rewards were delivered between a pair of monkeys. In this study, they found that pre-SMA and its rostrally adjacent area 9 of the MFC encoded reward values in either self-referenced or other-reference frames in an agent-specific manner, whereas midbrain dopaminergic neurons encoded an integrated subjective value (as modulated by the likelihood that another individual would receive a larger or smaller reward) (Noritake et al., 2018). Furthermore, single neurons in the MFC (pre-SMA) have been shown to encode representations of objects in an agent-specific manner. In a study by Livi et al. (2019), the activity of pre-SMA neurons was compared between trials where monkeys were instructed to grasp (or not grasp) one of three possible objects and trials where subject monkeys observed a human doing the same action to the objects. In this experimental setting, pre-SMA neurons either selectively encoded monkey's own actions, the actions of another agent, or both the monkey and other agent's actions (Livi et al., 2019) (Fig. 5B), again supporting the role of MFC in agent-specific representation of behaviors involving self-, other-, and both-reference frames. Therefore, social monitoring functions by the MFC seems to be more generalized to include social value monitoring, positioning MFC as a core brain region implicated in other-referenced computations that might be central to mentalizing and empathetic operations in the primate brain.

Taken together, these behavioral neurophysiology studies parallel functional neuroimaging and electroencephalographic data from human subjects that has identified neural signals of reward outcome or value across self and other in the MFC (Apps et al., 2012; Apps, Lesage, & Ramnani, 2015; Janowski, Camerer, & Rangel, 2013; Jenkins, Macrae, & Mitchell, 2008; Lockwood, Apps, Roiser, & Viding, 2015; Piva et al., 2019). However, there seems to be differently specialized functions across distinct subregions in the primate MFC with respect to their agent-specific processing. For example, in the case of ACCg, converging evidence from behavioral neurophysiology in monkeys and functional neuroimaging in humans suggests a specialized role of ACCg in computing other-referenced reward information used for tracking the motivations of others to possibly enable empathy and prosocial decision-making (Apps et al., 2016). By contrast, again based on both behavioral neurophysiology in monkeys and functional neuroimaging in humans, more dorsal MFC regions relative to ACCg, such as pre-SMA, seems to be more dedicated to understanding other's mental states through observing and predicting other's actions and intentions (Isoda & Noritake, 2013).

3.2 Neurons in the MFC predict the behavior of others

Although monitoring the actions of others is a supremely useful function, using past knowledge gleaned from doing so to predict their *future* behaviors has even greater behavioral utility. Neurons in the MFC have been shown to do exactly this; predict the future behaviors of social partners. A study by Haroush and Williams (2015) revealed that primate ACC neurons predict the behavior of an opponent during a cooperative social exchange. The authors trained pairs of rhesus monkeys to play an interactive prisoner's dilemma game, and recorded the activity of single neurons in the ACC. By analyzing the activity of these ACC neurons, they found two non-overlapping neuronal populations that encoded either the subject's (the monkey from whom the neurons were recorded) own choice, or the other monkey's upcoming choice, predicted by the previous pattern of cooperative behaviors (Fig. 5B). Importantly, a relatively high percentage of neurons (about 28%) in the ACC encoded the other's upcoming choice (Haroush & Williams, 2015). The activity of these ACC neurons found in this study can be interpreted as a prediction of the other's choice during social interactions, as it was yet unknown to the subject. Moreover, when decoding the other's choice using the simultaneously recorded population of neurons, it was possible to classify the opponent's choice with high accuracy (about 80%). Causally, microstimulations to the ACC reduced the frequency of cooperative decisions following past cooperative interactions by others, likely by disrupting the predictability necessary to guide cooperative decisions in the iterative prisoner's dilemma task. These results support the view that ACC neurons encode information about other social agents and make predictions about their behavior during social interaction.

Another direct neurophysiological evidence supporting the role of MFC in social predictions was reported during a monkey-human interaction paradigm used by Falcone, Cirillo, Ferraina, and Genovesio (2017). In this study, a human experimenter took turns alternating with a rhesus monkey in selecting between pairs of targets (from a potential pool of four), where choosing the target that did not match previously chosen target led to rewards, thus requiring the observer to monitor the other actor's choice when it was not their turn. Recordings of single unit activity in the MFC (pre-SMA, SMA, and a region rostral to them) of the participating monkey revealed that selective neurons encoded the future choice with in an agent-specific and often target-direction selective manner (Falcone et al., 2017) (Fig. 5B). This encoding of a prediction for a partner's future behaviors by MFC neurons represents a necessary algorithmic process to regulate social learning and adaptative social behaviors and supports the notion that a key implantation of these algorithm occurs in the MFC.

3.3 Neurons in the MFC are involved in observational and learning from vicarious reinforcement

The ability to predict other's future behaviors is, in turn, a necessary component of social learning that requires learning derived from other-referenced predictions and errors (Joiner et al., 2017). The role of MFC in observational learning has been documented in neurons in the brains of epileptic patients by Hill et al. (2016). In this study, human subjects performed a card game where they attempted to differentiate between two decks of cards as high- or low-probability of winning. Subjects learned through both self-experienced trials (i.e. they picked for themselves) and observational trials (i.e. they observed two other players pick between the decks of cards). Although neurons in the amygdala and prefrontal regions encoded the outcome (winning or losing) of trials, only neurons in rostral ACC neurons encoded observational reward prediction errors (Fig. 5C). These observational prediction errors, or the difference between the expected and actual reward, likely serve as an important algorithm supporting observational learning (Apps et al., 2012, 2015, 2016; Burke, Tobler, Baddeley, & Schultz, 2010; Joiner et al., 2017; Lockwood et al., 2015).

Learning from what happens to others is also important for developing and maintaining social preference. In the previously mentioned social reward allocation task, monkeys will form a prosocial preference (i.e. choosing to donate juice to the other monkey instead of no one) and antisocial preference (delivering juice to themselves over both monkeys) in a context-dependent manner (Chang et al., 2011). The process by which these preferences develop requires learning from vicarious reinforcement, where the actor monkey must monitor their social partner and understand the rewarding impact of their choices to experience vicarious reinforcement in order to develop and maintain a prosocial or an antisocial preference. Does the MFC contribute causally to learning to develop a social preference derived from vicarious reward? Recently, Basile and colleagues addressed this precise question using a selective excitotoxic lesion approach to the ACC. In this study, all six actors in a modified version of the social allocation task showed a prosocial preference (preferring other's reward over wasted reward) before excitotoxic lesion to ACC. However, postoperatively, the ACC-lesioned actors ($n=3$), but not control actors ($n=3$), could no longer learn to acquire a prosocial preference based on vicarious reinforcement derived from the reward delivered to the recipient monkeys (Basile et al., 2020) (Fig. 6). Importantly, the ACC-lesioned actors had no problem on simple learning to prefer one's own rewards over other's rewards or wasted rewards to an empty bottle (Basile et al., 2020). This deficit in forming a prosocial preference is possibly due to the lack of a social prediction error signal originating from the ACC. Already formed prosocial preferences (those with preoperatively learned cues) were reduced compared to the control animals but were still present, suggesting that the ACC is necessary for learning from vicarious reinforcement but perhaps only partially involved in the process of maintaining social preferences.

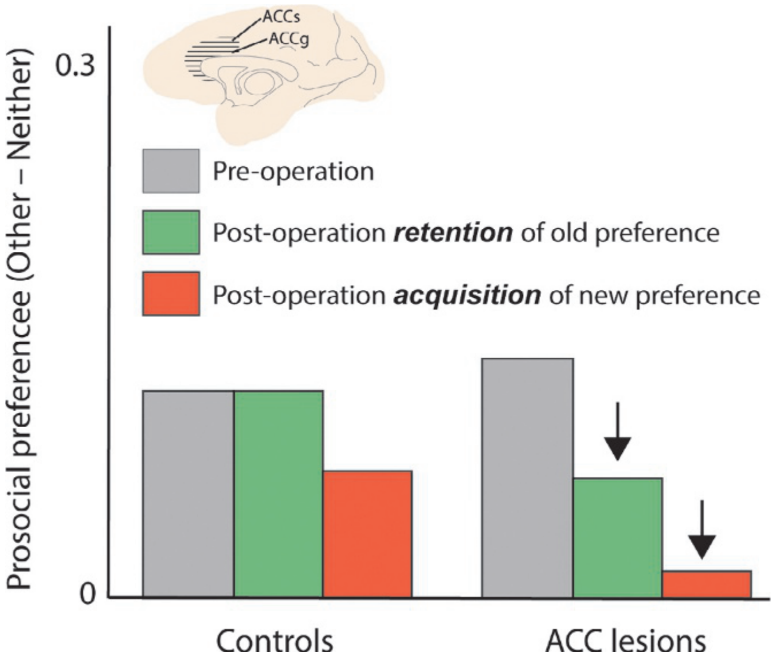

Fig. 6 ACC lesion (indicated in the inset brain diagram) reduced preoperatively learned prosocial preference and eliminated formation of postoperative prosocial preference in a vicarious social reward task (i.e., a preference for a juice reward delivered to a conspecific monkey ["Other"] over an empty bottle ["Neither"]). There was no effect from the ACC lesion on learning to prefer a juice reward to self over the conspecific monkey or over the empty bottle. Further, there was no overall changes to pupillary responses from the lesion. Data from three ACC lesion animals and three control animals. *Illustration based on Basile, B. M., Schafroth, J. L., Karaskiewicz, C. L., Chang, S. W. C., & Murray, E. A. (2020). The anterior cingulate cortex is necessary for forming prosocial preferences from vicarious reinforcement in monkeys.* PLoS Biology, 18*(6), e3000677. doi:10.1371/journal.pbio.3000677.*

3.4 The MFC and theory of mind

It remains unclear if any primate species aside from humans truly exhibit theory of mind (Drayton & Santos, 2016; Horschler et al., 2020; Krupenye et al., 2016), making laboratory experiments aimed at studying the neural mechanisms underlying theory of mind extremely challenging. However, a recent study Hayashi et al. (2020) implicates the MFC as being causally involved in theory of mind based on the false belief task (Wimmer & Perner, 1983) in Japanese macaques. In this experiment employing a dramatized and highly salient version of the false belief task similar to the one used recently by Krupenye and colleagues in apes (Krupenye et al., 2016), the eye movements of monkeys were recorded while they watched videos showing the aforementioned scenarios. During each video, the agent pursued a target that was placed in one of two different locations. By comparing anticipatory looking between the false-belief location (where the first agent last saw the object) and the new location truly containing the object, the authors showed that macaques looked first and more often at the false-belief location (Hayashi et al., 2020). The authors then tested the role of the MFC in this false belief mentalizing by disrupting MFC function using designer receptor exclusively activated by designer drugs (DREADDs). They found that MFC inactivation abolished the anticipatory looking toward the false-belief location (Fig. 5D), suggesting that the MFC plays a casual role in this mentalizing behavior. However, it is worthwhile to note that the behavioral demonstration of such a false belief representation in monkeys still remains controversial (see Horschler et al., 2020). Nevertheless, this study still lends a causal support for other-referenced information processing in the MFC.

4. Attributes of the MFC for social processing

The role of the MFC in performing specific algorithms that support necessary social computations is further supported by anatomical and comparative evidence. First, the MFC regions have bidirectional anatomical connections with the limbic system, allowing integrations across affective and reward-related information (Barbas, 2000; Carmichael & Price, 1995a). This is in contrast to the lateral aspects of the frontal cortex which interfaces more strongly with the sensorimotor processing regions in the brain, including the temporal and parietal cortices (Carmichael & Price, 1995b) for predominantly guiding externally guided cognition. Therefore, the MFC is anatomically equipped to synergistically work with the limbic systems to enable internally guided computations involving affective and reward-related variables suitable for social interactions. Second, possibly as a result of the anatomical connections, functional characteristics of the MFC in relation to executing socially relevant algorithms seem to differ from many other key brain regions implicated in social behaviors.

4.1 Bidirectional anatomical connections with the limbic networks

The interconnectedness of the MFC with limbic structures (Amaral & Insausti, 1992; Carmichael & Price, 1995a; Morecraft et al., 2012), particularly the amygdala, is a key feature that enables the MFC to perform these social computations. The amygdala contains many neural specializations for social perception, such as neurons that respond selectively to faces

(Gothard, Battaglia, Erickson, Spitler, & Amaral, 2007; Minxha et al., 2017; Rutishauser et al., 2011), emotions (Kuraoka & Nakamura, 2007; Wang et al., 2017), and eye contact (Mosher, Zimmerman, & Gothard, 2014; Putnam & Gothard, 2019). Amygdala damage in humans (Adolphs, Tranel, Damasio, & Damasio, 1994; Spezio, Huang, Castelli, & Adolphs, 2007) and monkeys (Emery et al., 2001; Taubert et al., 2018) impairs social perception and social interactions. Given the anatomical connections between MFC and the amygdala (Amaral & Insausti, 1992; Carmichael & Price, 1995a; Morecraft et al., 2012), it is not surprising that recent studies suggest functional relationships between MFC and the amygdala in the production of social behaviors. A recent study by Dal Monte et al. (2020) revealed that prosocial decisions were associated with enhanced synchronization of spikes and local field potential (LFP) activity between the MFC (ACCg) and the amygdala, whereas antisocial decisions measured from a different context were associated with suppressed synchronization of the same processes (Fig. 7) (Dal Monte et al., 2020). Moreover, this spike-LFP synchrony between the two areas was frequency-specific, where the frequency channel used was determined by the area contributing the spikes (i.e., spikes from the amygdala was synchronized to the beta frequency of ACCg LFP, whereas spikes from ACCg was synchronized to the gamma frequency of the amygdala) (Dal Monte et al., 2020). Interestingly, the researchers also found that prosocial decisions were associated with an increase in directional information transfer from the amygdala to ACCg in the same beta frequency band, whereas antisocial decisions were associated with an increased in information flow from ACCg to the amygdala more broadly across different frequency bands (Dal Monte et al., 2020). Other studies examining non-social behaviors collaborate the importance of the interactions between the ACC and the amygdala in guiding behaviors. For example, these interactions are important in shaping reward learning (Fiuzat, Rhodes, & Murray, 2017; Klavir, Genud-Gabai, & Paz, 2013; Livneh & Paz, 2012; Rudebeck, Mitz, Chacko, & Murray, 2013; Rudebeck, Ripple, Mitz, Averbeck, & Murray, 2017) and observational learning (Burgos-Robles, Gothard, Monfils, Morozov, & Vicentic, 2019). The amygdalo-cingulate connectivity is excellently reviewed in depth by Burgos-Robles and colleagues, where the authors discuss how these cytoarchitectonic features and functional connections support the role of the ACC in observational learning (Burgos-Robles et al., 2019).

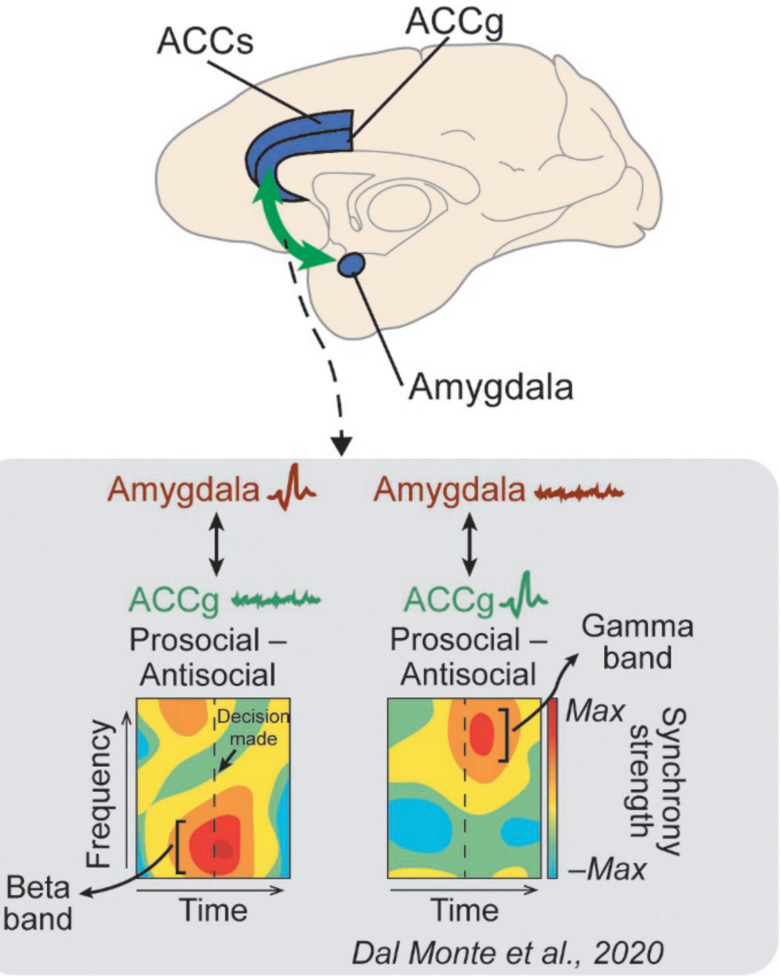

Fig. 7 Coherence between spiking activity and local field potential (LFP) signals between ACCg and the basolateral amygdala was found to be enhanced for making prosocial decisions but suppressed for making antisocial decisions in monkeys. This coherence occurred in dedicated frequency channels critically depending on the brain region contributing the spikes to the coherence. *Illustrations based on the data from Dal Monte, O., Chu, C. C. J., Fagan, N. A., & Chang, S. W. C. (2020). Specialized medial prefrontal–amygdala coordination in other-regarding decision preference.* Nature Neuroscience, 23*(4), 565–574. doi:10.1038/s41593-020-0593-y and the summary figure from Gangopadhyay, P., Chawla, M., Dal Monte, O., & Chang, S. W. C. (2020). Neural mechanisms guiding social decision-making in the prefrontal-amygdala circuits.* Nature Neuroscience, *(in press).*

4.2 Distinctive attributes of the MFC for socially relevant algorithms

Although the MFC represents a collection of brain regions out of many that are implicated in guiding social behaviors, several functional characteristics of the MFC distinguish its contributions to guiding social behaviors from other brain regions. Here we will compare the MFC to a small selection of other brain regions and example studies to explore the potential uniqueness of algorithms employed by the MFC. Our goal in this section is not to be exhaustive but rather to briefly highlight one or two studies from other brain regions in order to inform how neurons in the MFC might exhibit different socially relevant algorithms from other brain regions.

4.2.1 Compared to the lateral frontal cortex

The lateral frontal cortex is also comprised of a collection of brain regions, including the ventral premotor cortex (PMv) and dorsolateral prefrontal cortex (dlPFC). These areas are strongly implicated in cognitive operations such as working memory (Barbey, Koenigs, & Grafman, 2013; Curtis & D'Esposito, 2003; Funahashi, Bruce, & Goldman-Rakic, 1989; Fuster & Alexander, 1971; Miller, Erickson, & Desimone, 1996), outcome valuation (Cai & Padoa-Schioppa, 2014; Yim, Cai, & Wang, 2019), and strategy implementation (Baxter, Gaffan, Kyriazis, & Mitchell, 2009; Seo, Cai, Donahue, & Lee, 2014). Most pertinently to social processing, the lateral frontal cortex contains a network of neurons, known as "mirror neurons," that respond to both the observation of an action by another, and the performance of the same action by oneself (Fabbri-Destro & Rizzolatti, 2008; Rizzolatti & Craighero, 2004; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996; Rizzolatti, Fogassi, & Gallese, 2001). Mirror neurons were first identified in PMv but also since found in many different brain regions such as the inferior parietal lobule, the primary motor cortex, and the superior temporal sulcus (Dushanova & Donoghue, 2010; Ferrari, Bonini, & Fogassi, 2009; Hamilton, 2013; Rizzolatti et al., 2001; Suzuki et al., 2015), forming a mirror neuron network in the brain (Bonini, 2017). Mirror neurons have many parallels to neurons found in the MFC; mirror neurons seem to represent the actions of self and other and necessitate monitoring another individual's actions with a theorized role of understanding other's intentions. However, one key differentiation exists; mirror neurons can be algorithmically described as a stimulation of other's behavior based on externally guided information. A key feature of mirror neurons is that observing and executing the same actions are encoded in the same manner, hence the term mirroring. By contrast, a consistent feature of agent-specific processing in the MFC is the independent and coexistent representations of self-referenced, other-referenced, and a mirror-like combined self-and-other-referenced representations. Having separate representations of self-referenced and other-referenced signals within different population of neurons in the MFC is critical for many of the algorithms implemented for social behaviors such that the brain can differentiate self from other or represent information in a relative fashion between self and other. Mirroring by itself does not provide such separate representations and would not be possible to predict other's behaviors during social interactions (except when the interactions consists of simply mimicking one another). While the mirror neuron network is implicated as one mechanism for transforming of visual information into a motor format (Fabbri-Destro & Rizzolatti, 2008), it remains undecided if this network is required for visuomotor transformations (Hickok, 2009), and if this system causally impacts social behaviors (Hamilton, 2013).

Multiple studies also suggest that the dlPFC, another subdivision of the lateral frontal cortex, is centrally involved in strategic social interactions. In experiments where a monkey was strategically pitted against a computer opponent, neurons in the dlPFC tracked not only the subject's previous choices but also previous choices made by the opponent (Seo, Barraclough, & Lee, 2007; Seo et al., 2014). Moreover, many neurons in this region have been shown to encode winning and losing resulting from competition, but not from non-competition, when two monkeys played a missile commando game (Hosokawa & Watanabe, 2012). In fact, tracking the actions of opponents and the respective reward outcomes emerging from competition are both important algorithmic processes required for strategic behaviors. The involvement of the dlPFC in strategic decision-making is consistent with the broader role of the dlPFC in executive functions (Lee & Seo, 2007). One important contrast is that, compared to MFC neurons with stronger associations with the limbic networks and possibly under a greater influence of internal states, the information represented by dlPFC neurons might be more tightly linked with sensorimotor or externally guided processing of actions linked to rewarding outcomes (Donahue & Lee, 2015; Lee & Seo, 2007).

4.2.2 Compared to the amygdala

The amygdala is a subcortical region strongly implicated in social behaviors across many species (Adolphs, 2010). Like the MFC, the amygdala plays a role non-social learning such as reward valuation (Baxter & Murray, 2002; Bermudez, Göbel, & Schultz, 2012; Bermudez & Schultz, 2010; Burgos-Robles et al., 2017; Fiuzat et al., 2017; Paton et al., 2006; Putnam & Gothard, 2019; Rudebeck et al., 2013, 2017). A recent study by

Grabenhorst et al. (2012) provides strong evidence of functional overlap between the amygdala and MFC. In this study, two monkeys alternated between making reward-based choices, choosing between one of two fractals unique to each monkey that were associated with different reward probabilities. Halfway through each recording session, the pairs of fractals switched between the monkeys so that the observational learning could be tested. By recording single unit activity from the amygdala in one of the two monkeys, it was found that amygdala neurons encoded the reward values of the fractals, regardless of if this value was extract from personal experience or observation (Grabenhorst et al., 2012). The neuronal selectivity to the fractal stimuli was not merely for certain visual patterns, as this reward-value encoding tracked the unannounced probability reversal during each of the two halves. This study provides evidence that amygdala neurons, like neurons in the ACC (Haroush & Williams, 2015), are capable of predicting the actions of others by simulating their decisions. Indeed, it would be interesting to understand if these predictive signals of other's behaviors observed in the ACC (Haroush & Williams, 2015) originate from the amygdala, or vice versa, given the rich connections between the basolateral amygdala (the amygdala subdivision with the neural population showing the strongest shared value representation between self and another individual based on cross-decoding of reward value from self to other; Grabenhorst et al., 2012) and the ACC (Morecraft & Van Hoesen, 1998). While this seemingly shared algorithm in both the ACC and the amygdala provides evidence for functional overlaps between the two regions, there are still notable differences between them. It is possible that this algorithm is shared between the two different regions, however the overall computational goal might be different. One hypothesis is that the computational goal of the amygdala in observational learning is to support one own's future goal-directed behaviors, i.e. the subjects in the task were implicitly aware of an impending stimulus switch and thus tracking the stimulus-reward associations of the other's monkey was beneficial. However, this can be contrasted with neurons in the ACCg of the MFC that signal other's reward outcome even though representing this reward outcome had no behavioral contingencies for one's own reward (Chang et al., 2013), an independent representation of other's outcome that may be directly relevant for enabling vicarious reinforcement. This hypothesis, however, is mostly speculative and future studies will be needed to dissect the similarities and differences between the MFC and amygdala in representing social variables with respect to self and others.

4.2.3 Compared to the lateral intraparietal cortex

The lateral intraparietal cortex (LIP) is a region broadly associated with spatial attention (Bisley & Goldberg, 2003) and planning eye movements (Snyder, Batista, & Andersen, 2000). There exists evidence that, additionally, LIP neurons track the gaze of others. Shepherd, Klein, Deaner, and Platt (2009) recorded single unit activity from the LIP in a monkey while the subject participated in a simple visual orienting task. In this task, the monkey fixated on a center target, and once fixation was initiated, an image of a conspecific was presented at the center serving as a cue for the subject to shift their gaze toward a peripheral target. The direction of this saccade was either congruent or incongruent with the conspecific's gaze direction in the image. The results showed that LIP neurons are modulated by social gaze cues by exhibiting enhanced or suppressed firing rates when the social gaze was directed toward the neuron's receptive field (Shepherd et al., 2009). These results may capture a social gaze-related sensorimotor transformation by LIP neurons, which may underlie the transformation required for gaze-following behaviors. As of present, no data suggests that LIP neurons are involved in predicting other's behaviors in social interactions, nor do LIP neurons represent the gaze of others using a dissociated representation from one's own gaze direction. One plausible explanation is that LIP neurons may perform an attentional allocation or motor planning algorithm based on other's gaze direction to accomplish the sensorimotor transformation according to the information available from others. By contrast, social monitoring algorithms performed by the neurons in the MFC might be purposed for understanding other's preferences and other's internal (e.g., motivational) states (Haroush & Williams, 2015; Joiner et al., 2017; Lockwood et al., 2015) to support social learning and social decision-making (Apps et al., 2012, 2015, 2016; Joiner et al., 2017).

4.2.4 Compared to the striatum

The striatum is key node of decision-making and reward circuity in the brain (Burton, Nakamura, & Roesch, 2015). Unsurprisingly, as rewards powerfully shape social interactions, the striatum is critically implicated in social behaviors (Báez-Mendoza & Schultz, 2013). Báez-Mendoza and colleagues tested how neurons in the monkey striatum represented social action and reward by implementing a task where the subject monkey and another partner (either a second money or a computer actor) took turns making operant responses for rewards. They found that striatal neurons encoded signals for the actions of both self and other, but only encoded reward signals for self

(Báez-Mendoza, Harris, & Schultz, 2013). This coding of social action in the striatum parallels the findings from the MFC (Yoshida et al., 2011, 2012), and may suggest that striatal neurons work in close communication with the MFC to associate the information about other's actions and rewards given the well-known role of the striatum in linking action to reward (Balleine, Delgado, & Hikosaka, 2007; Monosov, Leopold, & Hikosaka, 2015). However, the absence of a localized other-referenced reward signal in that study also highlights the difference of the striatum from the MFC (Chang et al., 2013; Dal Monte et al., 2020). There is another excellent study from the ventral striatum for comparing representations of other's reward by the ventral striatum and the MFC. A study by Kashtelyan, Lichtenberg, Chen, Cheer, and Roesch (2014) measured dopamine release in rats when they observed rewards being delivered to conspecifics. When rats *first* observed conspecific receive rewards, there was an increase of dopamine release in the ventral striatum, along with appetitive ultrasonic vocalizations (Kashtelyan et al., 2014). Critically, later trials did not result in increased dopamine but rather a reduction in dopamine along with an increased rate of aversive ultrasonic vocalizations. These results suggest that representations of other's reward in the ventral striatum is highly dependent on affective states. In the MFC, representations of other's reward were found to be stable over the course of entire sessions (Chang et al., 2013), although existing studies have not yet tested how these socially relevant signals are gated by fluctuating internal states.

5. Social specificity of the MFC: Algorithms and implementations

An outstanding and critical question remains on understanding the role of the MFC in guiding social behaviors. Are these neuronal processes and algorithms in the MFC specific to social behaviors or are they shared across both social and non-social functions? Generally, the question of the existence of dedicated regions, circuits, or processes in the brain for social behaviors (compared to an alternative hypothesis that all of these are all shared between social and non-social faculties) is still under debate. Some researchers have put forward the concept of a "social brain," of which the MFC is a major node, that implies specialized systems for social cognition (Adolphs, 2009). Despite the presence of functionality that is inherently social, such as the perception of faces and emotions, there is increasing evidence of overlap between social and non-social functions at the neuronal level. For example, although the evaluation of social stimuli, particularly faces, is often cited as a key specialized module of the social brain, it was recently shown in rhesus monkeys that the same amygdala neurons encode faces (Putnam & Gothard, 2019) or social hierarchies (Munuera, Rigotti, & Salzman, 2018) also track non-social features such as reward and task dynamics. This outstanding question is where the framework of Marr is most useful as we can examine social specificity at different organizational levels, and this approach has been recently applied in order to redefine the concept of social specificity, or lack thereof, through the lens of computational, algorithmic, and implementational levels (Lockwood et al., 2020). In this last section, we examine the MFC in social functions specifically at the level of algorithms and implementations.

At the most macroscopic level, the computational goal of social interactions is compelled by the fundamental drives (e.g. needing to compete or cooperate with others to obtain food). While these goals can be innately social, the algorithms that support them are unlikely to be socially specific. Algorithmic process, such as reinforcement learning, have been used to model both social and non-social behaviors with high accuracy (Behrens, Hunt, Woolrich, & Rushworth, 2008; Charpentier, Iigaya, & O'Doherty, 2020; Ereira, Dolan, & Kurth-Nelson, 2018; Olsson et al., 2020). It has been argued that the ACCg, an area with a high degree of social specificity at the implementational level, may be utilizing reinforcement learning as an common underlying algorithm (Apps et al., 2015, 2016; Behrens et al., 2008; Hill et al., 2016; Joiner et al., 2017; Lockwood et al., 2015, 2018). However, there also exist algorithmic processes that appear to be exclusive to social functions. One example of this is the process of metalizing or theory of mind. In theory of mind tasks, such as the one used by Hayashi et al. (2020), general algorithms such as reinforcement learning cannot fully explain the experimental results (Devaine et al., 2017). In general, while it is clearer that the MFC implements socially non-specific algorithms in the production of social behaviors such as reinforcement learning algorithms, it remains to be further tested if there are also socially specific algorithms being implemented.

There is some evidence that there are socially specific implementations in the MFC. As referenced previously, a study by Chang et al. (2013) found that different proportions of neurons in the ACCg and ACCs encode other-referenced and self-referenced reward, respectively, with varying degrees of specialization. Neurons in the ACCg principally encoded rewards delivered to the *other* monkey, either through coding other- or both-referenced rewards. In contrast, neurons in the ACCs principally encoded rewards delivered to self, although other- or both-referenced reward coding neurons constituted a sizeable minority. In contrast still, nearly 80% of reward-encoding neurons recorded from the OFC were self-referenced. These findings support a division of socially specific implementations between the gyrus and the sulcus of the ACC at the neuronal level. A potentially specialized implementation in ACCg has been supported from human neuroimaging literature as well (Apps et al., 2015, 2016; Behrens et al., 2008; Hill et al., 2016; Joiner et al., 2017; Lockwood et al., 2015, 2018). Interestingly, in a recent study examining shared coding for reward value and hierarchical rank in monkeys, Munuera et al. (2018) found that the OFC and ACC encoded reward values but failed to robustly encode representations of hierarchical rank in a shared manner. This was in contrast to the neurons recorded from the amygdala that used the same neuronal ensembles to encode both hierarchical rank and reward values associated with non-social stimuli (Munuera et al., 2018). These findings suggest, accepting the premise that a single hierarchical algorithm can be used to sort information such as reward value or rank, the presence of potential socially specific implementations in the ACC and OFC but not in the amygdala. Given the range of functions associated with the MFC, it seems nearly certain that some MFC neurons implement generalized algorithms spanning both social and non-social behaviors during social interactions. Future studies may identify possibly differentiated neuronal ensembles in the MFC regions that implement socially specific and socially non-specific algorithms for guiding social behaviors.

6. Conclusions

In this chapter, we outline how the MFC contributes to shaping adaptive social behaviors through the use of several key algorithms. Experimental evidence, from behavioral neurophysiology, supports the involvement of MFC neurons in each of these algorithms. Single neuron recording studies from the MFC provide evidence for the function of the MFC in monitoring others in an agent-specific manner (Chang et al., 2013; Dal Monte et al., 2020; Noritake et al., 2018; Yoshida et al., 2011, 2012) and in predicting the behavior of social partners (Falcone, Brunamonti, Ferraina, & Genovesio, 2016; Falcone et al., 2017; Haroush & Williams, 2015; Livi et al., 2019). Moreover, single unit recording from human patients in the rostral ACC as well lesion studies in non-human primates implicate the

MFC in observational learning and learning from vicarious reinforcement (Basile et al., 2020; Hill et al., 2016). Finally, a recent study utilizing chemogenetic inactivation of the MFC suggests a role for the MFC in theory of mind like processes (Hayashi et al., 2020). Together, socially relevant algorithms found in MFC neurons support necessary social computations required for effective social interactions. Although the brain regions within the MFC are not the only brain areas implicated in shaping social functions, the MFC has various unique attributes that allow the MFC to play a central role in shaping interactive social behaviors. For example, strong connectivity with the limbic network supports the implantation of algorithmic processes such as observational learning (Burgos-Robles et al., 2019) and prosocial decision-making (Dal Monte et al., 2020). When compared to some of other brain areas, there seems to be some unique algorithms performed by MFC neurons. To continue to investigate and understand social functions in the MFC, applying the three levels of organization proposed by Marr (1982) to social behaviors (Lockwood et al., 2020) may help guide and refine research questions, designs, and interpretations. Fruitful research directions include examining socially relevant algorithms and implementations in the MFC by holding one level constant through innovative experimental designs (Lockwood et al., 2020). Continued technological innovations will in turn allow new discoveries of how socially specific and socially non-specific processes in different MFC ensembles interact to guide social interactions.

Acknowledgment

This work was supported by the National Institute of Mental Health (R01 MH120081, R01MH110750). We thank Colin Stanton for help with some of the figure illustrations.

References

Adolphs, R. (2009). The social brain: Neural basis of social knowledge. *Annual Review of Psychology*, *60*, 693–716. https://doi.org/10.1146/annurev.psych.60.110707.163514.

Adolphs, R. (2010). What does the amygdala contribute to social cognition? *Annals of the New York Academy of Sciences*, *1191*(1), 42–61. https://doi.org/10.1111/j.1749-6632.2010.05445.x.

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature*, *372*(6507), 669–672. https://doi.org/10.1038/372669a0.

Akins, C. K., Klein, E. D., & Zentall, T. R. (2002). Imitative learning in Japanese quail (Coturnix japonica) using the bidirectional control procedure. *Animal Learning & Behavior*, *30*(3), 275–281. https://doi.org/10.3758/BF03192836.

Amaral, D. G., & Insausti, R. (1992). Retrograde transport of D-[3H]-aspartate injected into the monkey amygdaloid complex. *Experimental Brain Research*, *88*(2), 375–388. https://doi.org/10.1007/BF02259113.

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews. Neuroscience*, 7(4), 268–277. https://doi.org/10.1038/nrn1884.

Apps, M. A. J., Balsters, J. H., & Ramnani, N. (2012). The anterior cingulate cortex: Monitoring the outcomes of others' decisions. *Social Neuroscience*, 7(4), 424–435. https://doi.org/10.1080/17470919.2011.638799.

Apps, M. A. J., Lesage, E., & Ramnani, N. (2015). Vicarious reinforcement learning signals when instructing others. *Journal of Neuroscience*, *35*(7), 2904–2913. https://doi.org/10.1523/JNEUROSCI.3669-14.2015.

Apps, M. A. J., Rushworth, M. F. S., & Chang, S. W. C. (2016). The anterior cingulate gyrus and social cognition: Tracking the motivation of others. *Neuron*, *90*(4), 692–707. https://doi.org/10.1016/j.neuron.2016.04.018.

Azab, H., & Hayden, B. Y. (2017). Correlates of decisional dynamics in the dorsal anterior cingulate cortex. *PLoS Biology*, *15*(11), e2003091. https://doi.org/10.1371/journal.pbio.2003091.

Bachevalier, J., Meunier, M., Lu, M. X., & Ungerleider, L. G. (1997). Thalamic and temporal cortex input to medial prefrontal cortex in rhesus monkeys. *Experimental Brain Research*, *115*(3), 430–444.

Báez-Mendoza, R., Harris, C. J., & Schultz, W. (2013). Activity of striatal neurons reflects social action and own reward. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(41), 16634–16639. https://doi.org/10.1073/pnas.1211342110.

Báez-Mendoza, R., & Schultz, W. (2013). The role of the striatum in social behavior. *Frontiers in Neuroscience*, 7, 233. https://doi.org/10.3389/fnins.2013.00233.

Balleine, B. W., Delgado, M. R., & Hikosaka, O. (2007). The role of the dorsal striatum in reward and decision-making. *Journal of Neuroscience*, *27*(31), 8161–8165. https://doi.org/10.1523/JNEUROSCI.1554-07.2007.

Bandura, A. (2008). Observational learning. In *The International encyclopedia of communication* American Cancer Society. https://doi.org/10.1002/9781405186407.wbieco004.

Bandura, A., Ross, D., & Ross, S. A. (1961). Transmission of aggression through imitation of aggressive models. *The Journal of Abnormal and Social Psychology*, *63*(3), 575–582. https://doi.org/10.1037/h0045925.

Bandura, A., & Walters, R. H. (1977). *Social learning theory*. *Vol. 1*. NJ: Prentice-Hall Englewood Cliffs.

Barbas, H. (2000). Connections underlying the synthesis of cognition, memory, and emotion in primate prefrontal cortices. *Brain Research Bulletin*, *52*(5), 319–330. https://doi.org/10.1016/s0361-9230(99)00245-2.

Barbey, A. K., Koenigs, M., & Grafman, J. (2013). Dorsolateral prefrontal contributions to human working memory. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, *49*(5), 1195–1205. https://doi.org/10.1016/j.cortex.2012.05.022.

Basile, B. M., Schafroth, J. L., Karaskiewicz, C. L., Chang, S. W. C., & Murray, E. A. (2020). The anterior cingulate cortex is necessary for forming prosocial preferences from vicarious reinforcement in monkeys. *PLoS Biology*, *18*(6). https://doi.org/10.1371/journal.pbio.3000677, e3000677.

Baxter, M. G., Gaffan, D., Kyriazis, D. A., & Mitchell, A. S. (2009). Ventrolateral prefrontal cortex is required for performance of a strategy implementation task but not reinforcer devaluation effects in rhesus monkeys. *The European Journal of Neuroscience*, *29*(10), 2049–2059. https://doi.org/10.1111/j.1460-9568.2009.06740.x.

Baxter, M. G., & Murray, E. A. (2002). The amygdala and reward. *Nature Reviews. Neuroscience*, *3*(7), 563–573. https://doi.org/10.1038/nrn875.

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, *10*(3), 295–307. https://doi.org/10.1093/cercor/10.3.295.

Behrens, T. E. J., Hunt, L. T., & Rushworth, M. F. S. (2009). The computation of social behavior. *Science*, *324*(5931), 1160–1164. https://doi.org/10.1126/science.1169694.

Behrens, T. E. J., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. S. (2008). Associative learning of social value. *Nature*, *456*(7219), 245–249. https://doi.org/10.1038/nature07538.

Bermudez, M. A., Göbel, C., & Schultz, W. (2012). Sensitivity to temporal reward structure in amygdala neurons. *Current Biology: CB*, *22*(19), 1839–1844. https://doi.org/10.1016/j.cub.2012.07.062.

Bermudez, M. A., & Schultz, W. (2010). Reward magnitude coding in primate amygdala neurons. *Journal of Neurophysiology*, *104*(6), 3424–3432. https://doi.org/10.1152/jn.00540.2010.

Bevilacqua, D., Davidesco, I., Wan, L., Chaloner, K., Rowland, J., Ding, M., et al. (2019). Brain-to-brain synchrony and learning outcomes vary by student-teacher dynamics: Evidence from a real-world classroom electroencephalography study. *Journal of Cognitive Neuroscience*, *31*(3), 401–411. https://doi.org/10.1162/jocn_a_01274.

Bisley, J. W., & Goldberg, M. E. (2003). Neuronal activity in the lateral intraparietal area and spatial attention. *Science (New York, N.Y.)*, *299*(5603), 81–86. https://doi.org/10.1126/science.1077395.

Bonini, L. (2017). The extended mirror neuron network: Anatomy, origin, and functions. *The Neuroscientist*, *23*(1), 56–67. https://doi.org/10.1177/1073858415626400.

Burgos-Robles, A., Gothard, K. M., Monfils, M. H., Morozov, A., & Vicentic, A. (2019). Conserved features of anterior cingulate networks support observational learning across species. *Neuroscience and Biobehavioral Reviews*, *107*, 215–228. https://doi.org/10.1016/j.neubiorev.2019.09.009.

Burgos-Robles, A., Kimchi, E. Y., Izadmehr, E. M., Porzenheim, M. J., Ramos-Guasp, W. A., Nieh, E. H., et al. (2017). Amygdala inputs to prefrontal cortex guide behavior amid conflicting cues of reward and punishment. *Nature Neuroscience*, *20*(6), 824–835. https://doi.org/10.1038/nn.4553.

Burke, C. J., Tobler, P. N., Baddeley, M., & Schultz, W. (2010). Neural mechanisms of observational learning. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(32), 14431–14436. https://doi.org/10.1073/pnas.1003111107.

Burton, A. C., Nakamura, K., & Roesch, M. R. (2015). From ventral-medial to dorsal-lateral striatum: Neural correlates of reward-guided decision-making. *Neurobiology of Learning and Memory*, *117*, 51–59. https://doi.org/10.1016/j.nlm.2014.05.003.

Cai, X., & Padoa-Schioppa, C. (2014). Contributions of orbitofrontal and lateral prefrontal cortices to economic choice and the good-to-action transformation. *Neuron*, *81*(5), 1140–1151. https://doi.org/10.1016/j.neuron.2014.01.008.

Carmichael, S. T., & Price, J. L. (1995a). Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *The Journal of Comparative Neurology*, *363*(4), 615–641. https://doi.org/10.1002/cne.903630408.

Carmichael, S. T., & Price, J. L. (1995b). Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. *The Journal of Comparative Neurology*, *363*(4), 642–664. https://doi.org/10.1002/cne.903630409.

Carmichael, S. T., & Price, J. L. (1996). Connectional networks within the orbital and medial prefrontal cortex of macaque monkeys. *The Journal of Comparative Neurology*, *371*(2), 179–207. https://doi.org/10.1002/(SICI)1096-9861(19960722)371:2<179::AID-CNE1>3.0.CO;2-#.

Chang, S. W. C. (2013). Coordinate transformation approach to social interactions. *Frontiers in Neuroscience*, 7, 147. https://doi.org/10.3389/fnins.2013.00147.

Chang, S. W. C. (2017). An emerging field of primate social neurophysiology: Current developments. *eNeuro*, *4*(5). https://doi.org/10.1523/ENEURO.0295-17.2017. ENEURO.0295-17.2017.

Chang, S. W. C., Barter, J. W., Ebitz, R. B., Watson, K. K., & Platt, M. L. (2012). Inhaled oxytocin amplifies both vicarious reinforcement and self reinforcement in rhesus macaques (Macaca mulatta). *Proceedings of the National Academy of Sciences of the United States of America*, *109*(3), 959–964. https://doi.org/10.1073/pnas.1114621109.

Chang, S. W. C., Fagan, N. A., Toda, K., Utevsky, A. V., Pearson, J. M., & Platt, M. L. (2015). Neural mechanisms of social decision-making in the primate amygdala. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(52), 16012–16017. https://doi.org/10.1073/pnas.1514761112.

Chang, S. W. C., Gariépy, J.-F., & Platt, M. L. (2013). Neuronal reference frames for social decisions in primate frontal cortex. *Nature Neuroscience*, *16*(2), 243–250. https://doi.org/10.1038/nn.3287.

Chang, S. W. C., Winecoff, A. A., & Platt, M. L. (2011). Vicarious reinforcement in Rhesus Macaques (Macaca Mulatta). *Frontiers in Neuroscience*, *5*, 27. https://doi.org/10.3389/fnins.2011.00027.

Charpentier, C. J., Iigaya, K., & O'Doherty, J. P. (2020). A neuro-computational account of arbitration between choice imitation and goal emulation during human observational learning. *Neuron*, *106*(4), 687–699.e7. https://doi.org/10.1016/j.neuron.2020.02.028.

Chen, P., & Hong, W. (2018). Neural circuit mechanisms of social behavior. *Neuron*, *98*(1), 16–30. https://doi.org/10.1016/j.neuron.2018.02.026.

Curtis, C. E., & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences*, 7(9), 415–423. https://doi.org/10.1016/S1364-6613(03)00197-9.

Dal Monte, O., Chu, C. C. J., Fagan, N. A., & Chang, S. W. C. (2020). Specialized medial prefrontal–amygdala coordination in other-regarding decision preference. *Nature Neuroscience*, *23*(4), 565–574. https://doi.org/10.1038/s41593-020-0593-y.

Debiec, J., & Olsson, A. (2017). Social fear learning: From animal models to human function. *Trends in Cognitive Sciences*, *21*(7), 546–555. https://doi.org/10.1016/j.tics.2017.04.010.

Devaine, M., San-Galli, A., Trapanese, C., Bardino, G., Hano, C., Jalme, M. S., et al. (2017). Reading wild minds: A computational assay of theory of mind sophistication across seven primate species. *PLoS Computational Biology*, *13*(11), e1005833. https://doi.org/10.1371/journal.pcbi.1005833.

Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., et al. (2017). Brain-to-brain synchrony tracks real-world dynamic group interactions in the classroom. *Current Biology*, *27*(9), 1375–1380. https://doi.org/10.1016/j.cub.2017.04.002.

Dölen, G., Darvishzadeh, A., Huang, K. W., & Malenka, R. C. (2013). Social reward requires coordinated activity of nucleus accumbens oxytocin and serotonin. *Nature*, *501*(7466), 179. https://doi.org/10.1038/nature12518.

Donahue, C. H., & Lee, D. (2015). Dynamic routing of task-relevant signals for decision making in dorsolateral prefrontal cortex. *Nature Neuroscience*, *18*(2), 295–301. https://doi.org/10.1038/nn.3918.

Drayton, L. A., & Santos, L. R. (2016). A decade of theory of mind research on Cayo Santiago: Insights into rhesus macaque social cognition. *American Journal of Primatology*, *78*(1), 106–116. https://doi.org/10.1002/ajp.22362.

Dushanova, J., & Donoghue, J. (2010). Neurons in primary motor cortex engaged during action observation. *The European Journal of Neuroscience*, *31*(2), 386–398. https://doi.org/10.1111/j.1460-9568.2009.07067.x.

Emery, N. J., Capitanio, J. P., Mason, W. A., Machado, C. J., Mendoza, S. P., & Amaral, D. G. (2001). The effects of bilateral lesions of the amygdala on dyadic social interactions in rhesus monkeys (Macaca mulatta). *Behavioral Neuroscience*, *115*(3), 515–544.

Ereira, S., Dolan, R. J., & Kurth-Nelson, Z. (2018). Agent-specific learning signals for self–other distinction during mentalising. *PLoS Biology*, *16*(4), e2004752. https://doi.org/10.1371/journal.pbio.2004752.

Fabbri-Destro, M., & Rizzolatti, G. (2008). Mirror neurons and mirror systems in monkeys and humans. *Physiology*, *23*(3), 171–179. https://doi.org/10.1152/physiol.00004.2008.

Falcone, R., Brunamonti, E., Ferraina, S., & Genovesio, A. (2016). Neural encoding of self and another agent's goal in the primate prefrontal cortex: Human-monkey interactions. *Cerebral Cortex (New York, N.Y.: 1991)*, *26*(12), 4613–4622. https://doi.org/10.1093/cercor/bhv224.

Falcone, R., Cirillo, R., Ferraina, S., & Genovesio, A. (2017). Neural activity in macaque medial frontal cortex represents others' choices. *Scientific Reports*, *7*, 12663. https://doi.org/10.1038/s41598-017-12822-5.

Ferrari, P. F., Bonini, L., & Fogassi, L. (2009). From monkey mirror neurons to primate behaviours: Possible 'direct' and 'indirect' pathways. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *364*(1528), 2311–2323. https://doi.org/10.1098/rstb.2009.0062.

Fiuzat, E. C., Rhodes, S. E. V., & Murray, E. A. (2017). The role of orbitofrontal-amygdala interactions in updating action-outcome valuations in macaques. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *37*(9), 2463–2470. https://doi.org/10.1523/JNEUROSCI.1839-16.2017.

Frith, C., & Frith, U. (2010). Learning from others: Introduction to the special review series on social neuroscience. *Neuron*, *65*(6), 739–743.

Funahashi, S., Bruce, C. J., & Goldman-Rakic, P. S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology*, *61*(2), 331–349. https://doi.org/10.1152/jn.1989.61.2.331.

Fuster, J. M., & Alexander, G. E. (1971). Neuron activity related to short-term memory. *Science*, *173*(3997), 652–654. https://doi.org/10.1126/science.173.3997.652.

Gothard, K. M., Battaglia, F. P., Erickson, C. A., Spitler, K. M., & Amaral, D. G. (2007). Neural responses to facial expression and face identity in the monkey amygdala. *Journal of Neurophysiology*, *97*(2), 1671–1683. https://doi.org/10.1152/jn.00714.2006.

Gothard, K. M., Mosher, C. P., Zimmerman, P. E., Putnam, P. T., Morrow, J. K., & Fuglevand, A. J. (2017). New perspectives on the neurophysiology of primate amygdala emerging from the study of naturalistic social behaviors. *Wiley Interdisciplinary Reviews: Cognitive Science*, *9*(1), e1449. https://doi.org/10.1002/wcs.1449.

Grabenhorst, F., Hernádi, I., & Schultz, W. (2012). Prediction of economic choice by primate amygdala neurons. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(46), 18950–18955. https://doi.org/10.1073/pnas.1212706109.

Hamilton, A. F.d. C. (2013). Reflecting on the mirror neuron system in autism: A systematic review of current theories. *Developmental Cognitive Neuroscience*, *3*, 91–105. https://doi.org/10.1016/j.dcn.2012.09.008.

Haroush, K., & Williams, Z. M. (2015). Neuronal prediction of opponent's behavior during cooperative social interchange in primates. *Cell*, *160*(6), 1233–1245. https://doi.org/10.1016/j.cell.2015.01.045.

Hayashi, T., Akikawa, R., Kawasaki, K., Egawa, J., Minamimoto, T., Kobayashi, K., et al. (2020). Macaques exhibit implicit gaze bias anticipating others' false-belief-driven actions via medial prefrontal cortex. *Cell Reports*, *30*(13), 4433–4444.e5. https://doi.org/10.1016/j.celrep.2020.03.013.

Hayden, B. Y., & Platt, M. L. (2010). Neurons in anterior cingulate cortex multiplex information about reward and action. *Journal of Neuroscience*, *30*(9), 3339–3346. https://doi.org/10.1523/JNEUROSCI.4874-09.2010.

Heyes, C. M., Jaldow, E., Nokes, T., & Dawson, G. R. (1994). Imitation in rats (Rattus norvegicus): The role of demonstrator action. *Behavioural Processes*, *32*(2), 173–182. https://doi.org/10.1016/0376-6357(94)90074-4.

Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience*, *21*(7), 1229–1243. https://doi.org/10.1162/jocn.2009.21189.

Hill, M. R., Boorman, E. D., & Fried, I. (2016). Observational learning computations in neurons of the human anterior cingulate cortex. *Nature Communications*, 7(1), 12722. https://doi.org/10.1038/ncomms12722.

Horschler, D. J., MacLean, E. L., & Santos, L. R. (2020). Do non-human primates really represent others' beliefs? *Trends in Cognitive Sciences*, *24*(8), 594–605. https://doi.org/10.1016/j.tics.2020.05.009.

Hosokawa, T., & Watanabe, M. (2012). Prefrontal neurons represent winning and losing during competitive video shooting games between monkeys. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(22), 7662–7671. https://doi.org/10.1523/JNEUROSCI.6479-11.2012.

Isoda, M., & Noritake, A. (2013). What makes the dorsomedial frontal cortex active during reading the mental states of others? *Frontiers in Neuroscience*, 7. https://doi.org/10.3389/fnins.2013.00232.

Janowski, V., Camerer, C., & Rangel, A. (2013). Empathic choice involves vmPFC value signals that are modulated by social processing implemented in IPL. *Social Cognitive and Affective Neuroscience*, *8*(2), 201–208. https://doi.org/10.1093/scan/nsr086.

Jenkins, A. C., Macrae, C. N., & Mitchell, J. P. (2008). Repetition suppression of ventromedial prefrontal activity during judgments of self and others. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(11), 4507–4512. https://doi.org/10.1073/pnas.0708785105.

Joiner, J., Piva, M., Turrin, C., & Chang, S. W. C. (2017). Social learning through prediction error in the brain. *NPJ Science of Learning*, *2*(1), 1–9. https://doi.org/10.1038/s41539-017-0009-2.

Kashtelyan, V., Lichtenberg, N. T., Chen, M. L., Cheer, J. F., & Roesch, M. R. (2014). Observation of reward delivery to a conspecific modulates dopamine release in ventral striatum. *Current Biology*, *24*(21), 2564–2568. https://doi.org/10.1016/j.cub.2014.09.016.

Kennerley, S. W., Behrens, T. E. J., & Wallis, J. D. (2011). Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nature Neuroscience*, *14*(12), 1581–1589. https://doi.org/10.1038/nn.2961.

Klavir, O., Genud-Gabai, R., & Paz, R. (2013). Functional connectivity between amygdala and cingulate cortex for adaptive aversive learning. *Neuron*, *80*(5), 1290–1300. https://doi.org/10.1016/j.neuron.2013.09.035.

Kostorz, K., Flanagin, V. L., & Glasauer, S. (2020). Synchronization between instructor and observer when learning a complex bimanual skill. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2020.116659, 116659.

Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2016). Great apes anticipate that other individuals will act according to false beliefs. *Science*, *354*(6308), 110–114. https://doi.org/10.1126/science.aaf8110.

Kuraoka, K., & Nakamura, K. (2007). Responses of single neurons in monkey amygdala to facial and vocal emotions. *Journal of Neurophysiology*, *97*(2), 1379–1387. https://doi.org/10.1152/jn.00464.2006.

Lee, D., & Seo, H. (2007). Mechanisms of reinforcement learning and decision making in the primate dorsolateral prefrontal cortex. *Annals of the New York Academy of Sciences*, *1104*(1), 108–122. https://doi.org/10.1196/annals.1390.007.

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. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(7), 2691–2700. https://doi.org/10.1073/pnas.1810890116.

Livneh, U., & Paz, R. (2012). Amygdala-prefrontal synchronization underlies resistance to extinction of aversive memories. *Neuron*, *75*(1), 133–142. https://doi.org/10.1016/j.neuron.2012.05.016.

Lockwood, P. L., Apps, M. A. J., & Chang, S. W. C. (2020). Is there a "social" brain? Implementations and algorithms. *Trends in Cognitive Sciences*, *24*(10), 802–813. https://doi.org/10.1016/j.tics.2020.06.011.

Lockwood, P. L., Apps, M. A. J., Roiser, J. P., & Viding, E. (2015). Encoding of vicarious reward prediction in anterior cingulate cortex and relationship with trait empathy. *Journal of Neuroscience*, *35*(40), 13720–13727. https://doi.org/10.1523/JNEUROSCI.1703-15.2015.

Lockwood, P. L., Wittmann, M. K., Apps, M. A. J., Klein-Flügge, M. C., Crockett, M. J., Humphreys, G. W., et al. (2018). Neural mechanisms for learning self and other ownership. *Nature Communications*, *9*(1), 4747. https://doi.org/10.1038/s41467-018-07231-9.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York, NY: Henry Holt and Co Inc. 2(4.2).

Miller, E. K., Erickson, C. A., & Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *Journal of Neuroscience*, *16*(16), 5154–5167. https://doi.org/10.1523/JNEUROSCI.16-16-05154.1996.

Minxha, J., Mosher, C., Morrow, J. K., Mamelak, A. N., Adolphs, R., Gothard, K. M., et al. (2017). Fixations gate species-specific responses to free viewing of faces in the human and macaque amygdala. *Cell Reports*, *18*(4), 878–891. https://doi.org/10.1016/j.celrep.2016.12.083.

Monosov, I. E., Leopold, D. A., & Hikosaka, O. (2015). Neurons in the primate medial basal forebrain signal combined information about reward uncertainty, value, and punishment anticipation. *Journal of Neuroscience*, *35*(19), 7443–7459. https://doi.org/10.1523/JNEUROSCI.0051-15.2015.

Morecraft, R. J., Stilwell-Morecraft, K. S., Cipolloni, P. B., Ge, J., McNeal, D. W., & Pandya, D. N. (2012). Cytoarchitecture and cortical connections of the anterior cingulate and adjacent somatomotor fields in the rhesus monkey. *Brain Research Bulletin*, *87*(4–5), 457–497. https://doi.org/10.1016/j.brainresbull.2011.12.005.

Morecraft, R. J., & Van Hoesen, G. W. (1998). Convergence of limbic input to the cingulate motor cortex in the rhesus monkey. *Brain Research Bulletin*, *45*(2), 209–232. https://doi.org/10.1016/S0361-9230(97)00344-4.

Mosher, C. P., Zimmerman, P. E., & Gothard, K. M. (2014). Neurons in the monkey amygdala detect eye contact during naturalistic social interactions. *Current Biology: CB*, *24*(20), 2459–2464. https://doi.org/10.1016/j.cub.2014.08.063.

Munuera, J., Rigotti, M., & Salzman, C. D. (2018). Shared neural coding for social hierarchy and reward value in primate amygdala. *Nature Neuroscience*, *21*(3), 415–423. https://doi.org/10.1038/s41593-018-0082-8.

Noritake, A., Ninomiya, T., & Isoda, M. (2018). Social reward monitoring and valuation in the macaque brain. *Nature Neuroscience*, *21*(10), 1452–1462. https://doi.org/10.1038/s41593-018-0229-7.

Olsson, A., Knapska, E., & Lindström, B. (2020). The neural and computational systems of social learning. *Nature Reviews Neuroscience*, *21*(4), 197–212. https://doi.org/10.1038/s41583-020-0276-4.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science (New York, N.Y.)*, *308*(5719), 255–258. https://doi.org/10.1126/science.1107621.

Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature*, *441*(7090), 223–226. https://doi.org/10.1038/nature04676.

Pagnotta, M., Laland, K. N., & Coco, M. I. (2020). Attentional coordination in demonstrator-observer dyads facilitates learning and predicts performance in a novel manual task. *Cognition*, *201*. https://doi.org/10.1016/j.cognition.2020.104314, 104314.

Pärnamets, P., Espinosa, L., & Olsson, A. (2020). Physiological synchrony predicts observational threat learning in humans. *Proceedings of the Biological Sciences*, *287*(1927), 20192779. https://doi.org/10.1098/rspb.2019.2779.

Paton, J. J., Belova, M. A., Morrison, S. E., & Salzman, C. D. (2006). The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature*, *439*(7078), 865–870. https://doi.org/10.1038/nature04490.

Piva, M., Velnoskey, K., Jia, R., Nair, A., Levy, I., & Chang, S. W. (2019). The dorsomedial prefrontal cortex computes task-invariant relative subjective value for self and other. *eLife*, *8*. https://doi.org/10.7554/eLife.44939.

Porrino, L. J., Crane, A. M., & Goldman-Rakic, P. S. (1981). Direct and indirect pathways from the amygdala to the frontal lobe in rhesus monkeys. *Journal of Comparative Neurology*, *198*(1), 121–136. https://doi.org/10.1002/cne.901980111.

Putnam, P. T., & Gothard, K. M. (2019). Multidimensional neural selectivity in the primate amygdala. *ENeuro*, *6*(5). https://doi.org/10.1523/ENEURO.0153-19.2019.

Redcay, E., & Schilbach, L. (2019). Using second-person neuroscience to elucidate the mechanisms of social interaction. *Nature Reviews. Neuroscience*, *20*(8), 495–505. https://doi.org/10.1038/s41583-019-0179-4.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*(1), 169–192. https://doi.org/10.1146/annurev.neuro.27.070203.144230.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, *3*(2), 131–141. https://doi.org/10.1016/0926-6410(95)00038-0.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews. Neuroscience*, *2*(9), 661–670. https://doi.org/10.1038/35090060.

Rogers, R. D., Ramnani, N., Mackay, C., Wilson, J. L., Jezzard, P., Carter, C. S., et al. (2004). Distinct portions of anterior cingulate cortex and medial prefrontal cortex are activated by reward processing in separable phases of decision-making cognition. *Biological Psychiatry*, *55*(6), 594–602. https://doi.org/10.1016/j.biopsych.2003.11.012.

Rudebeck, P. H., Buckley, M. J., Walton, M. E., & Rushworth, M. F. S. (2006). A role for the macaque anterior cingulate gyrus in social valuation. *Science (New York, N.Y.)*, *313*(5791), 1310–1312. https://doi.org/10.1126/science.1128197.

Rudebeck, P. H., Mitz, A. R., Chacko, R. V., & Murray, E. A. (2013). Effects of amygdala lesions on reward-value coding in orbital and medial prefrontal cortex. *Neuron*, *80*(6), 1519–1531. https://doi.org/10.1016/j.neuron.2013.09.036.

Rudebeck, P. H., Ripple, J. A., Mitz, A. R., Averbeck, B. B., & Murray, E. A. (2017). Amygdala contributions to stimulus-reward encoding in the macaque medial and orbital frontal cortex during learning. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *37*(8), 2186–2202. https://doi.org/10.1523/JNEUROSCI.0933-16.2017.

Rutishauser, U., Tudusciuc, O., Neumann, D., Mamelak, A. N., Heller, A. C., Ross, I. B., et al. (2011). Single-unit responses selective for whole faces in the human amygdala. *Current Biology: CB*, *21*(19), 1654–1660. https://doi.org/10.1016/j.cub.2011.08.035.

Saver, J. L., & Damasio, A. R. (1991). Preserved access and processing of social knowledge in a patient with acquired sociopathy due to ventromedial frontal damage. *Neuropsychologia*, *29*(12), 1241–1249. https://doi.org/10.1016/0028-3932(91)90037-9.

Seo, H., Barraclough, D. J., & Lee, D. (2007). Dynamic signals related to choices and outcomes in the dorsolateral prefrontal cortex. *Cerebral Cortex (New York, N.Y.: 1991)*, *17*(Suppl. 1), i110–i117. https://doi.org/10.1093/cercor/bhm064.

Seo, H., Cai, X., Donahue, C. H., & Lee, D. (2014). Neural correlates of strategic reasoning during competitive games. *Science*, *346*(6207), 340–343. https://doi.org/10.1126/science.1256254.

Shepherd, S. V., Klein, J. T., Deaner, R. O., & Platt, M. L. (2009). Mirroring of attention by neurons in macaque parietal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(23), 9489–9494. https://doi.org/10.1073/pnas.0900419106.

Sliwa, J., & Freiwald, W. A. (2017). A dedicated network for social interaction processing in the primate brain. *Science (New York, N.Y.)*, *356*(6339), 745–749. https://doi.org/10.1126/science.aam6383.

Snyder, L. H. (2000). Coordinate transformations for eye and arm movements in the brain. *Current Opinion in Neurobiology*, *10*(6), 747–754. https://doi.org/10.1016/s0959-4388(00)00152-5.

Snyder, L. H., Batista, A. P., & Andersen, R. A. (2000). Intention-related activity in the posterior parietal cortex: A review. *Vision Research*, *40*(10), 1433–1441. https://doi.org/10.1016/S0042-6989(00)00052-3.

Spezio, M. L., Huang, P.-Y. S., Castelli, F., & Adolphs, R. (2007). Amygdala damage impairs eye contact during conversations with real people. *The Journal of Neuroscience*, *27*(15), 3994–3997. https://doi.org/10.1523/JNEUROSCI.3789-06.2007.

Suzuki, W., Banno, T., Miyakawa, N., Abe, H., Goda, N., & Ichinohe, N. (2015). Mirror neurons in a new world monkey, common marmoset. *Frontiers in Neuroscience*, *9*. https://doi.org/10.3389/fnins.2015.00459.

Taubert, J., Flessert, M., Wardle, S. G., Basile, B. M., Murphy, A. P., Murray, E. A., et al. (2018). Amygdala lesions eliminate viewing preferences for faces in rhesus monkeys. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(31), 8043–8048. https://doi.org/10.1073/pnas.1807245115.

Tomasello, M., Davis-Dasilva, M., Camak, L., & Bard, K. (1987). Observational learning of tool-use by young chimpanzees. *Human Evolution*, *2*(2), 175–183. https://doi.org/10.1007/BF02436405.

Wang, S., Yu, R., Tyszka, J. M., Zhen, S., Kovach, C., Sun, S., et al. (2017). The human amygdala parametrically encodes the intensity of specific facial emotions and their categorical ambiguity. *Nature Communications*, *8*, 14821. https://doi.org/10.1038/ncomms14821.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1), 103–128. https://doi.org/10.1016/0010-0277(83)90004-5.

Wise, S. P. (2008). Forward frontal fields: Phylogeny and fundamental function. *Trends in Neurosciences*, *31*(12), 599–608. https://doi.org/10.1016/j.tins.2008.08.008.

Yim, M. Y., Cai, X., & Wang, X.-J. (2019). Transforming the choice outcome to an action plan in monkey lateral prefrontal cortex: A neural circuit model. *Neuron*, *103*(3), 520–532.e5. https://doi.org/10.1016/j.neuron.2019.05.032.

Yoshida, K., Saito, N., Iriki, A., & Isoda, M. (2011). Representation of others' action by neurons in monkey medial frontal cortex. *Current Biology: CB*, *21*(3), 249–253. https://doi.org/10.1016/j.cub.2011.01.004.

Yoshida, K., Saito, N., Iriki, A., & Isoda, M. (2012). Social error monitoring in macaque frontal cortex. *Nature Neuroscience*, *15*(9), 1307–1312. https://doi.org/10.1038/nn.3180.

Zhou, T., Zhu, H., Fan, Z., Wang, F., Chen, Y., Liang, H., et al. (2017). History of winning remodels thalamo-PFC circuit to reinforce social dominance. *Science*, *357*(6347), 162–168. https://doi.org/10.1126/science.aak9726.