


## REVIEW ARTICLE OPEN

## Social learning through prediction error in the brain

Jessica Joiner<sup>1</sup>, Matthew Piva<sup>2,3</sup>, Courtney Turrin<sup>1</sup> and Steve W. C. Chang <sup>1,2,3</sup>

Learning about the world is critical to survival and success. In social animals, learning about others is a necessary component of navigating the social world, ultimately contributing to increasing evolutionary fitness. How humans and nonhuman animals represent the internal states and experiences of others has long been a subject of intense interest in the developmental psychology tradition, and, more recently, in studies of learning and decision making involving self and other. In this review, we explore how psychology conceptualizes the process of representing others, and how neuroscience has uncovered correlates of reinforcement learning signals to explore the neural mechanisms underlying social learning from the perspective of representing reward-related information about self and other. In particular, we discuss self-referenced and other-referenced types of reward prediction errors across multiple brain structures that effectively allow reinforcement learning algorithms to mediate social learning. Prediction-based computational principles in the brain may be strikingly conserved between self-referenced and other-referenced information.

*npj Science of Learning* (2017)2:8; doi:10.1038/s41539-017-0009-2

**HISTORICAL PERSPECTIVES ON REPRESENTING OTHER**

Learning about the world and making adaptive decisions is a critical feature of cognition. This important link allows human and nonhuman animals to manipulate their environment and survive. Decision-making takes on more complex dynamics when an animal is not solitary, but lives in a community with other members of its own species. We know much about how human and nonhuman animals learn from their own actions and outcomes, and where such self-referenced information is represented in the brain. However, much less is known about the computations underlying how we learn about others. In this review, we examine the presence of other-referenced prediction errors in the brain that represent other's actions and reward outcomes.

One of the first academic disciplines to attempt to understand how we develop a concept of others is developmental psychology, in which researchers often explore how babies come to understand the world.<sup>1,2</sup> One viewpoint, that of theory–theory, is that, like small scientists testing causal relations,<sup>3</sup> children are constantly gathering data from the world and testing the predictions they make using the collected data. Other people could be regarded as stimuli to be learned about based on observation and direct and vicarious experience.

In contrast, simulation theory assumes that we develop our understanding of others through self-referencing, in which we use the machinery that we use for our own mental processes, and project that knowledge onto other's behaviors.<sup>4</sup> Later, this notion of simulating others became associated with the “mirroring” neuronal activity observed in individual cortical motor neurons when macaque monkeys observe an action and perform the same action.<sup>5</sup>

Notably, these theories have different predictions about how other-referenced information is represented in the brain.<sup>6</sup> In a simulationist account, the notion of “other” is derived from one's

sense of the self, that is, egocentrically. Ideas about others would originate from and rely on the self-referenced, egocentric mechanisms. However, in theory–theory, information about others would be processed and evaluated like any other information from the environment, perhaps engaging allocentric systems. Historically, these two ideas capture a central question about how others are represented in the brain.

**OBSERVATIONAL AND SOCIAL LEARNING**

Both human and nonhuman animals rely on observation to navigate the world. Rats,<sup>7</sup> birds,<sup>8</sup> and chimpanzees<sup>9</sup> observe others to learn about their behaviors in a given environmental or social context. One of the earliest forms of observational learning occurs in imitation. Imitative learning typically involves a young organism copying the motor behaviors of a social exemplar. Infant humans<sup>10</sup> and infant monkeys<sup>11</sup> imitate gross facial expressions made by a caregiver early on in development, sticking out their tongue reflexively when an adult human demonstrates, likely an example of a simple motor response prepared by the brain for social development. Imitation in human children is most famously exemplified in observational learning or, more broadly, social learning studies.<sup>12</sup>

Social learning occurs when the learner watches another agent act. Notably, without any practice or direct primary reinforcement, the learner can perform the previously observed behavior. This suggests that the learner is able to acquire new knowledge or skills from watching other's experienced outcomes, possibly through vicarious reinforcement.<sup>13</sup> The efficacy of social learning depends on several social variables. For example, similarity between the observer and the observed increases the efficacy of learning.<sup>14</sup> Furthermore, there is a close link between empathy and social learning. Empathy is sensitive to learned information about the traits of the other person, such as how fair they are<sup>15</sup> or

<sup>1</sup>Department of Psychology, Yale University, New Haven, CT 06511, USA; <sup>2</sup>Department of Neuroscience, Yale School of Medicine, New Haven, CT 06520, USA and

<sup>3</sup>Interdepartmental Neuroscience Program, Yale University School of Medicine, New Haven, CT 06510, USA

Correspondence: Jessica Joiner (jessica.joiner@yale.edu)

Received: 30 October 2016 Revised: 17 May 2017 Accepted: 25 May 2017

Published online: 16 June 2017

whether or not the other person is considered in-group or out-group relative to the observer.<sup>16</sup> In addition, social status directly affects learning based on others in primates, in which high-status individuals are more likely to be imitated.<sup>17</sup>

In humans, observational learning may be at the core of establishing social and cultural norms.<sup>12</sup> In Bandura's classic study<sup>13</sup> on behavioral modeling, children that saw an adult model aggressive behavior toward a large doll later performed the same aggressive behaviors when given the opportunity to interact with the same doll. Observational learning has an important role in development as well as in later social interactions and social cognition. A critical question in social learning is how self and other are represented across the brain structures involved in learning and whether the learning-related signals referenced to self and other are engaging similar or distinct neural computations.

Social learning as we define it in this review focuses on this observation-based learning, in which a subject learns about another through observing their actions as well as their reward outcomes. However, the aspects of social learning are as multitudinous as the facets of a social interaction itself. One could learn about different aspects of the other, such as their personality or their mental state. Social learning could also reflect learning from others about one's own reward outcomes (e.g., a teacher providing feedback on a student's essay and grade).

### HIGHER-LEVEL SOCIAL COGNITION

Learning about others allows us to model the internal states of other entities. The ability to model another's beliefs is called Theory of Mind (ToM). ToM is arguably the most complex form of understanding other individuals, heavily engaging other-referenced processing. The fact that human babies can model the beliefs of others speaks to how complex and rich their representation of the world is from the beginning. Not surprisingly, understanding the mechanisms behind ToM has long been of great interest, with competing ideas about whether or not ToM represents a separate social process or the convergence of many generalized processes.<sup>18</sup> ToM is often measured using a false belief task,<sup>19</sup> which tests if a participant can understand whether a social model has an incorrect belief about an object's location. Notably, very young children, even as young as 11-month-old infants, are capable of modeling the internal beliefs of others and "pass" a false belief test,<sup>20</sup> indicating that other-referenced processing in the brain emerges very early in the human ontogeny.

Studying ToM in nonhuman animals, however, has led to more mixed results. For example, monkeys fail the same false belief tasks infants can pass.<sup>21</sup> Nevertheless, nonhuman primates have been shown to engage in other forms of understanding or at least representing others. Monkeys display joint attention via gaze following.<sup>22–24</sup> Monkeys will typically follow the gaze of another entity toward an object or direction, indicating either that they can understand something about the perspective of the other entity or that other's gaze is reflexively allocating one's attention through hard-wired neural mechanisms evolved to deal with the association between other's gaze angle and something of interest and value. Similarly, monkeys<sup>25</sup> and chimpanzees<sup>26</sup> have been shown to comprehend what visual information is available to a separate agent—for example, when given an opportunity to steal food, they prefer to do so from someone who does not have visual access to them at the moment of theft. This indicates that primates understand other entities have a different perspective, even if they do not necessarily model the beliefs of the other entity.

Taken together, both humans and nonhuman animals are capable of complex social cognition, but the level of sophistication is what might differentiate them in the evolutionary history.

Understanding the computations of other-referenced information and representations of self and other will further inform how the brain was evolved to enrich what is often referred to as higher-level social cognition.

### REINFORCEMENT LEARNING FRAMEWORK

Reinforcement learning (RL) is perhaps the most influential framework developed to describe how an agent learns by interacting with its environment. RL is derived from the behaviorist view of animal behavior, in which an organism's knowledge of the world is exclusively modeled based on its behavior. Crucially, RL theories focus on mechanistic accounts for behaviors based on several learning-related parameters established from empirical sources.

Both humans and nonhuman animals are excellent models for a variety of learning and decision-making tasks that are grounded on RL theories. Describing learning and learned outcomes through mathematical models is a powerful way to make explicit and testable predictions about how an organism will behave in a particular context and how they will make decisions that take into account internal states, such as motivation and subjective value.<sup>27</sup> The RL framework can capture seemingly complex behaviors with relatively simple yet elegant rules, as in the famous Rescorla–Wagner model.<sup>28</sup> Although various RL models differ in how they describe different cognitive phenomena, they share several core elements, such as the rate of learning or the salience of stimuli, to fit the specifics of learning and decision-making processes.

RL has its roots and applications in both engineering and psychology. RL has its core foundations in the work of Richard Bellman, most famous for developing the Bellman optimality equation and dynamic programming. The more widely appreciated root of RL is conceptualizing how organisms gather information from their environment to learn and make decisions. RL requires an agent that moves through different states, or contexts, in a given environment. Other necessary components include a reward signal, a value function, and a policy. Reward outcome is central to all forms of RL and consists of a quantity the agent gets as a result of its actions within the environment. The agent then computes a value function using that reward outcome that calculates the expected value of certain states/contexts as well as the conjunction of specific states and actions. The agent uses these value functions to develop a set of preferred actions, known as a policy. A model of the environment is an optional component of RL that can provide the organism with guidance on how to move from state to state.

In dynamic programming, developed by Bellman for engineering applications, a complete model of the environment is required. This idea requires the action of an agent to be guided by the expected payoff of the action in addition to the total expected payoff of potential actions in hypothetical future states.<sup>29</sup> The same principle applies to temporal discounting (TD) models, the predominant form of RL model applied in psychological studies of humans and other animals.<sup>30</sup> TD learning notably differs from dynamic programming, as it does not require any model of the environment. Instead, learning is accomplished by comparing expected reward to actual reward after a certain transition in time. This difference is the reward prediction error. This prediction error is used to update the value function and, ultimately, the policy of an agent interacting with its environment. Prediction error signaling is indeed the fundamental attribute of the original models of learning.<sup>28</sup> In simple terms, a prediction error calculates the difference between what the animal expects to have happen and what actually happens to the animal on a given event or trial.<sup>31</sup> This can also be described as an error signal.<sup>32</sup>

## PREDICTIVE CODING AND REINFORCEMENT LEARNING IN THE BRAIN

Prediction errors are effectively used as the signal that drives self-referenced learning. Organisms update their behavior on a trial by trial basis to account for new information provided by this discrepancy in expectation and outcome. In particular, the reward prediction error, which calculates the difference between expected payoff and received payoff, has been established as the striking correlate of a mathematical learning rule in neurobiology.<sup>33</sup>

A classic type of reward prediction error encoded in the brain is consistent with the type required for TD learning.<sup>34</sup> Owing to the essential nature of reward to adaptive behavior, areas that encode reward are some of the best studied regions of the brain besides the regions of the brain involved in sensorimotor transformations. Classically, the dopaminergic substantia nigra<sup>35</sup> and the ventral tegmentum<sup>36</sup> as well as the dorsal and ventral striatum<sup>37</sup> have been shown to be primary areas that process reward receipt and valuation, with dopamine's relationship to reward now known as one of the most iconic behavior to neurotransmitter associations.<sup>38</sup> As one might expect, these areas provide strong examples of encoding reward prediction errors.<sup>39, 40</sup>

Reward prediction error signals have also been found elsewhere in the brain. Primate lateral habenula neurons encode reciprocal information about reward outcomes to the previously described dopamine neurons in the midbrain.<sup>41</sup> Notably, the activity of the lateral habenula neurons precede the activity of the dopamine neurons, suggesting that the lateral habenula neurons serve as an input for the prediction error signal detected in the midbrain.<sup>41</sup> Furthermore, functional magnetic resonance imaging (fMRI) in humans has revealed the presence of multiple kinds of prediction errors and other learning-related signals across many reward-related structures in the cerebral cortex,<sup>42–44</sup> indicating that the prediction error signaling is a widely generalized mechanism linking learning and decision-making. Applying these models to conceptualizing behavior and neural activity has proved fruitful in the study of learning and decision-making, perhaps most famously in the finding that midbrain dopamine neurons represent the TD reward prediction error.<sup>33</sup>

At least two important branches of research into RL in the neurosciences continue today. The first involves the potential balance between neural substrates of model-free (basic TD learning) vs. model-based (akin to dynamic programming) learning.<sup>44</sup> These studies have collectively identified neural substrates of the model-based state transition error,<sup>45</sup> representation of model-based in addition to model-free prediction error in the striatum and ventromedial prefrontal cortex,<sup>46</sup> as well as brain areas that act as arbiters between model-free and model-based approaches.<sup>47</sup> The second branch is vicarious reinforcement, which can also be modeled in a RL framework to account for how other's behaviors could be used to update our own learning and decision-making processes using vicarious classes of prediction errors.<sup>48</sup> RL can potentially be implemented in social learning about the actions and rewards of others.<sup>48–50</sup>

Such vicarious reinforcement in an RL framework would intuitively have to be performed in a model-based manner, as it is unclear how a model-free RL system could possibly learn about another agent without creating and updating a model of the other agent's potential thoughts and future actions. Accordingly, research into how humans may use RL mechanisms to learn and make inferences about others have used a modified Q learning framework that involves a simulated other.<sup>50</sup> Still, although RL constitutes a strong opportunity to explain and conceptualize social learning, there exist other computational frameworks that may be applied to social cognition. For example, some have argued that the putative TD reward prediction error forming the basis of RL theories may instead be interpreted in terms of

expectation violation or even salience, especially in relation to activity in cortical regions.<sup>51, 52</sup> Other models specifically designed to elucidate mentalizing via game theoretic approaches have been highly successful in exploring social behaviors in the relative absence of an explicit RL framework. These mainly consist of algorithms that produce iterative representations of other agents recursively ad infinitum.<sup>53, 54</sup> Such approaches have not only explained typical human behavior in a stag-hunt game, but have also identified specific deficits in recursive social cognition in patients with autism spectrum disorders.<sup>55</sup>

Prediction error signals can occur for a variety of different events to be learned about, like action values, reward value, and reward timing.<sup>40, 56, 57</sup> Furthermore, prediction errors are not limited to the reward domain. Evidence of prediction error calculations are even present in sensorimotor areas of the brain that deal with fine tuning actions like the cerebellum and the frontal eye fields<sup>31</sup> (see Table 1 for types of prediction errors and associated brain regions). Therefore, a critic signal is responsible

**Table 1.** Representative list of brain areas in which signals that can be described as prediction errors have been found from either primate electrophysiology or human neuroimaging studies

Prediction error computed	Correlated brain area	
Egocentric		
Self action (action executed)—(action intended)	SC <sup>31</sup> Cerebellum <sup>31</sup> FEF <sup>31</sup> LIP <sup>31</sup> dlPFC <sup>31</sup>	OFC <sup>31</sup> ACC <sup>31</sup> MCC <sup>85, 86</sup>
Self reward outcome (actual reward outcome)—(expected reward outcome)	VTA <sup>33, 34, 36, 57</sup> VS/other striatum <sup>31, 56, 84</sup> SN <sup>35, 56</sup>	LHb <sup>41</sup> ACC <sup>86</sup> dlPFC <sup>31</sup>
Self reward value (actual value of reward)—(expected value of reward)	vmPFC <sup>40</sup> VTA <sup>40</sup> SN <sup>40</sup>	
Self reward timing (actual timing of reward)—(expected timing of reward)	VTA <sup>35, 57</sup> SN <sup>35, 57</sup>	
Alloentric		
Other action (other's actual action)—(other's expected action)	dlPFC <sup>84</sup> dmPFC <sup>91</sup>	VS/other striatum <sup>63–68</sup>
Other reward outcome (other's actual reward)—(other's expected reward)	ACCg <sup>48</sup> vmPFC <sup>84</sup> dmPFC <sup>42</sup>	MTG <sup>42</sup> STS <sup>42</sup> TPJ <sup>42</sup>
Other motivation (other's actual motivation)—(other's expected motivation)	ACCg <sup>48</sup>	
This is not a comprehensive list but rather a list to highlight the presence of predictive coding in the brain. Note that the list for the action-related error signals is mostly adapted from Schultz and Dickinson <sup>31</sup> review		
ACC anterior cingulate cortex, ACCg anterior cingulate gyrus, dlPFC dorsolateral prefrontal cortex, dmPFC dorsomedial prefrontal cortex, FEF frontal eye field, LHb lateral habenula, LIP lateral intraparietal area, MCC middle cingulate cortex, MTG medial temporal gyrus, OFC orbitofrontal cortex, SC superior colliculus, SN substantia nigra, STS superior temporal sulcus, TPJ temporoparietal junction, vmPFC ventromedial prefrontal cortex, VS ventral striatum, VTA ventral tegmental area		

for correcting behavioral outputs and cognitive representations across a variety of functional domains of the brain, endorsing the notion that predictive coding is a key feature of the brain.

As strides in describing increasingly complex human behaviors have been made, attempts to carry the study of learning and decision-making for the self into learning and decision-making that takes into account the behavior of others is now a subject of intense interest. Reacting appropriately to conspecifics and correctly anticipating their behavior is a necessity for social organisms, requiring them to rely on understanding each other just as much as they rely on understanding where to forage for food to survive. As expected, learning about others and representation of self and other are mediated by several reward-related brain structures.

### NEURAL BASIS OF SELF-REFERENCED AND OTHER-REFERENCED REINFORCEMENT SIGNALS

In this section, we discuss selected research findings that have provided novel insights into how the brain signals self-referenced and other-referenced information in the domain of reinforcement learning and decision-making. When applicable, we focus on other-referenced prediction error signals regarding actions and reward outcomes relevant to reward-guided social learning.

#### Striatum

Recent advances in the field of neuroscience have elegantly provided various supports for the use of RL mechanisms of learning about others. Although the striatum has long been a center of focus for self-referenced reward information and prediction error in the brain, the role of striatum in learning is not restricted to self-referential processing. In a study examining observational learning and vicarious reinforcement with respect to dopamine release, observer rats vocalized more and experienced significantly more dopamine release in the ventral striatum when seeing another rat receive reward compared to when reward was delivered to an empty box.<sup>58</sup> These results extend the role of dopamine release in associations with prediction error signaling to the social domain, implicating the involvement of the similar RL mechanisms for signaling other's reward outcome. Notably, the degree of dopamine release for other's reward outcome was still substantially weaker compared to one's own reward, suggesting that similar mechanisms are utilized but in ways that could be differentiated for self and other.<sup>58</sup> In monkeys engaged in a task environment involving actions from and reward outcomes for self and other, neurons in the striatum signal one's received reward but not the reward received by others while signaling the actions performed by others,<sup>59</sup> indicating that there may be specializations for signaling self-referenced and other-referenced information in the striatum, and this differentiation may further depend on the encoding of action and reward outcome of another individual.

There is also evidence that the striatum represents other-referenced reward and prediction errors from human fMRI studies. When socially evaluated by peers, previous positive social interaction with an individual led to that individual being associated with positive outcomes, which correlated with activity in the striatum as well as the orbitofrontal cortex. This suggests that social interaction can similarly activate brain regions that typically signal reinforcing values of primary reinforcers.<sup>60</sup> The striatum also appears to be involved in the relative valuation of reward where other's performance is compared to one's own performance.<sup>61, 62</sup> In an ultimatum game where subjects give money to a partner and receive a proportion of it back, activation of the striatum was also correlated with prediction errors that reflect the difference between the offer the subject received from the partner and what they expected the partner to give, but not

between how the subject expected to feel and how they actually felt, which appears to be reflected in ventromedial prefrontal cortex (vmPFC) and the posterior cingulate cortex.<sup>63</sup>

Furthermore, RL-like prediction errors regarding expectations formed about how others viewed the subjects were correlated with activity in the striatum, OFC, rACC, and anterior insula.<sup>60</sup> A variety of economic-game style tasks that require learning about other's actions and outcomes and/or modeling the internal states of others have reported that the striatum is implicated in these processes. For example, the observed actions of others influence one's own economic decisions and this is reflected in striatal BOLD response.<sup>64</sup> Furthermore, if added payoff for social learning is removed, so that only pure observation of others is necessary for the task, an interpersonal prediction error still occurs in the striatum.<sup>65</sup> Similarly, there is evidence from a reciprocity game that learning to trust or not trust others based on their behavior is mediated by a prediction error signal in the caudate nucleus.<sup>66</sup>

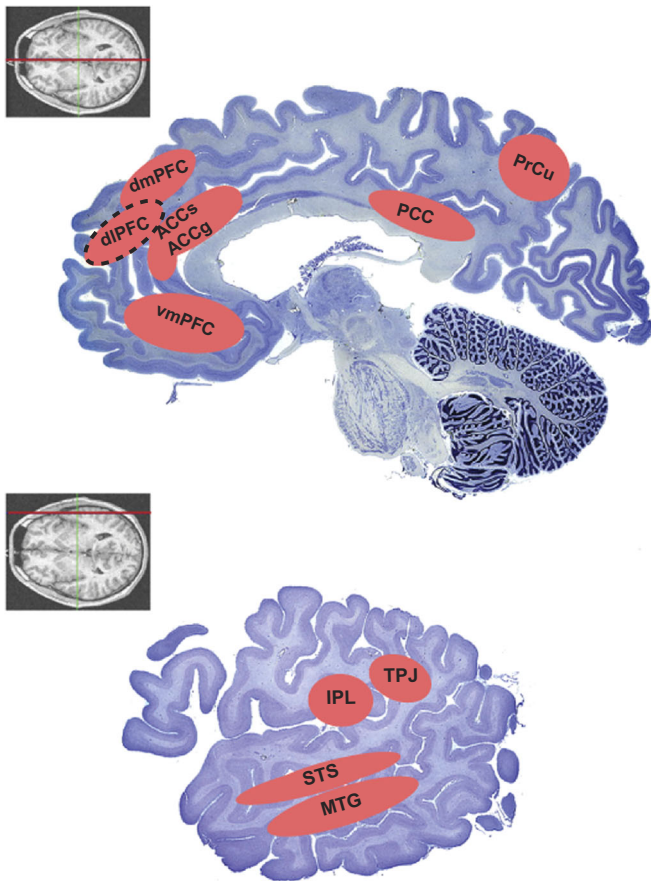
Interestingly, these other-referenced prediction errors in the striatum may even be associated with social norms, given their activation in economic games that rely on feedback from others. A prediction error type signal associated with going against group opinion also has been shown to correlate with how subjects changed their behavior to conform with the group on subsequent judgements.<sup>67</sup>

In a trust game where an investor gives money to a trustee who can return a proportion of the money, the difference between the trustee's repayment ratio expected by the actor and what the trustee actually repaid resulted in a prediction error in the striatum in the subjects who relied on the behavior of the partner for learning.<sup>68</sup> In addition, in the same study, the difference between the investment ratio and the investor's model of the other's model of what the investor will do formed a second order prediction error. Notably, the study found that a subject who failed to deeply model the mind of the partner experienced more striatal correlates of the first type of prediction error (i.e., relying more on the action of the other), whereas the more a subject modeled the mind of the partner, the more likely they were to activate the striatum for the second order prediction error (i.e., relying more on the mental representation of the other).

#### Anterior cingulate cortex

The anterior cingulate cortex (ACC) is implicated in a variety of behaviors and cognitive states,<sup>48, 69-71</sup> and could be summarized as an integrative area that relates to motivation and initiating reward-guided or goal-directed behaviors. Seen in this light, ACC may be a core locus of integrating different streams of self-referenced and other-referenced information for generating an adaptive action plan (see Fig. 1 for visualization of other-referenced reward and action areas of the brain). This is bolstered by the considerable evidence that ACC is engaged during social decision-making, with its neuronal signals reflecting information processing about self, other, or both.<sup>48, 72-77</sup> In the domain of observational learning, the ability of mice to learn by observing the shock of conspecifics can be effectively abolished by ACC-specific deletion of calcium currents.<sup>78</sup> Relatedly, observational aspects of pain have been a major focus for investigating empathy in the human brain. Observing a demonstration of another person being injured and experiencing pain elicits empathic concerns and actively engages a specific portion of the ACC that is also similarly activated when experiencing pain.<sup>79</sup> Such shared mechanisms support that observation-driven vicarious pain processing was co-opted or repurposed from processing one's own pain.

ACC may represent a critical junction in the cortical pathway of representing and differentiating self and other through processing motivation from the perspectives of self and others.<sup>48</sup> Monitoring the spiking activity of individual ACC neurons while monkeys played a social reward allocation task, in which an actor animal has



**Fig. 1** Key brain regions involved in representing information with respect to another individual. These areas are often implicated in mentalizing, detecting the beliefs of others, or signaling decision variables concerning another individual. See texts for how these areas are implicated in representing information with respect to another individual. The insets with coronal magnetic resonance images indicate the sections (red line) that correspond to the Nissl-stained sagittal slices. The dotted outline around an area indicates that this area is projected medially from the lateral surface for the purpose of including the area on a more medial aspect of the brain. Adapted with permission from <http://www.brains.rad.msu.edu>, <http://brainmuseum.org>, supported by the US National Science Foundation and the National Institutes of health. ACCg anterior cingulate gyrus, ACCs anterior cingulate sulcus, dlPFC dorsolateral prefrontal cortex, dmPFC dorsomedial prefrontal cortex, IPL inferior parietal lobule, MTG medial temporal gyrus, PCC posterior cingulate cortex, PrCu precuneus, STS superior temporal sulcus, TPJ temporoparietal junction, vmPFC ventromedial prefrontal cortex

an option to deliver or withhold juice rewards to and from the recipient,<sup>80</sup> showed that there are specializations with respect to signaling the reward outcome of self and other. More specifically, in the gyrus of ACC (ACCg), some neurons exclusively encode self reward, whereas others exclusively encode other's reward, and still some encoded the reward outcomes of self and other.<sup>81</sup> Notably, lesioning ACCg, but not the ACC sulcus (ACCs), abolishes social valuation in monkeys,<sup>82</sup> indicating a causal contribution of ACCg in social cognition. Similarly, in the human brain, the rostral ACC neurons, overlapping with the ACCg neurons mentioned above, signal reward outcomes from others during a card game requiring observational learning.<sup>77</sup>

Furthermore, neurons in ACC have been shown to mediate collective reward-guided actions when monkeys play a prisoner's dilemma game,<sup>74</sup> providing strong evidence that self and other

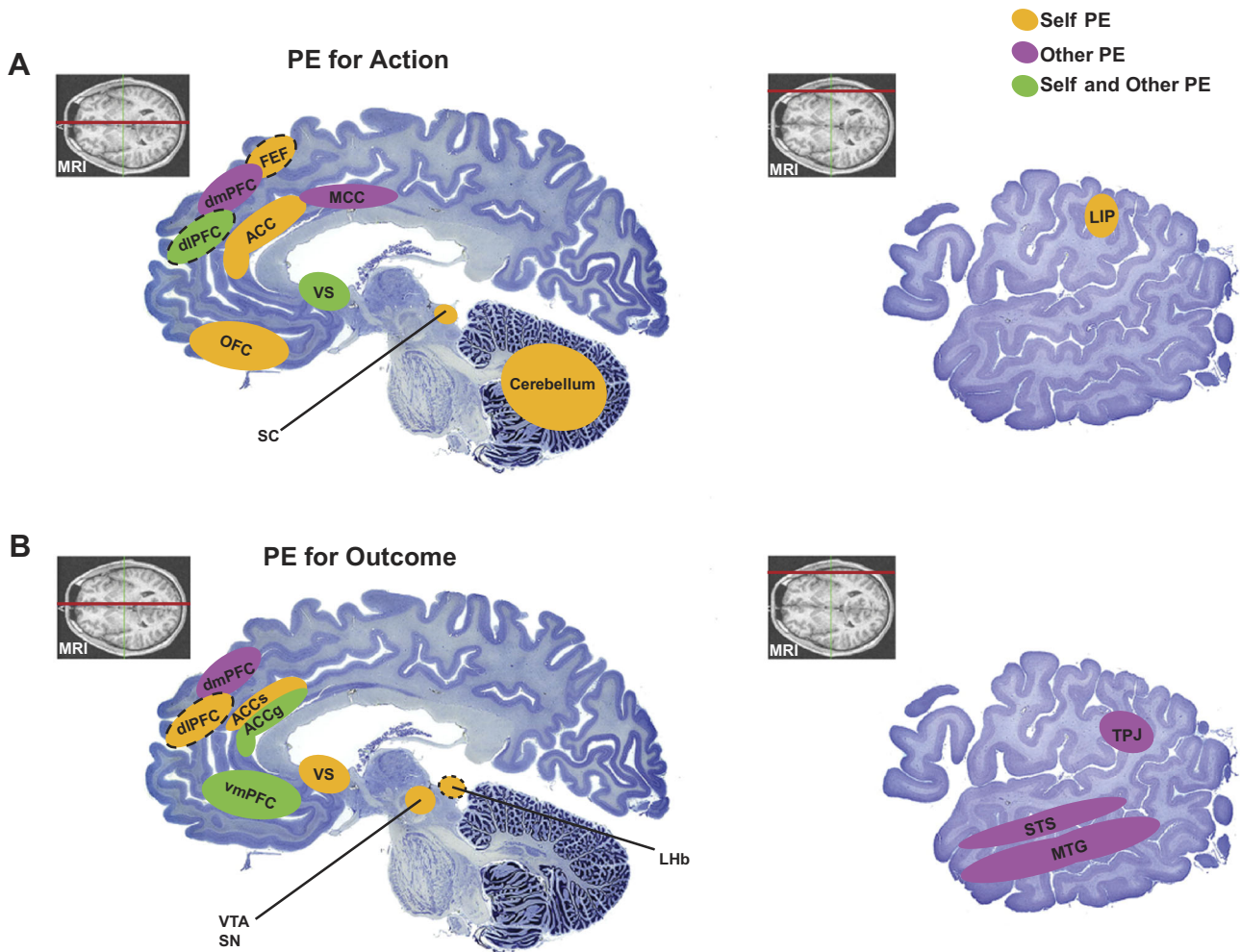
processes are integrated in ACC. The evidence of self and other integration in ACC is also supported by the presence of an anatomical gradient along the human cingulum mapping self and other in a trust game that is absent without a responding partner.<sup>83</sup> Moreover, it has been postulated that the ACCs and the ACCg represent distinct streams of information.<sup>18, 48, 72, 82</sup>

Accurate social learning requires multiple types of prediction error signals with respect to others (see Fig. 2 for representation of self-referencing and other-referencing prediction errors in the brain). For example, observational action prediction errors signal the difference between the actual action of the other and the expected action, whereas vicarious outcome prediction errors signal the difference between the actual and the predicted outcome of the other.<sup>48, 84</sup> Furthermore, to estimate the motivation of others, vicarious dynamic prediction errors signal the difference between the actual and the estimated movement kinematics of others during their actions.<sup>48</sup> Prediction errors for self-referenced action values have been reported in ACC,<sup>85, 86</sup> and both the sulcus and gyrus portions of the ACC are implicated in self-reward valuation and decision-making.<sup>87, 88</sup> The ACCs is most well-studied for involved in multitudinous functions, from error detection and motivation to cognitive control and response selection.<sup>85</sup> Recently, there is an extensive debate over whether or not ACCs are involved in computing value-guided behavioral adaptation or cognitive control.<sup>69-71</sup>

Notably, there seem to be functional dissociations for signaling self-referenced and other-referenced information between the gyrus and sulcus. For example, prediction errors related to the choices made by another person are found in the ACCg but not in the ACCs.<sup>48, 89</sup> Furthermore, the ACCs neurons encode reward outcomes in self-referenced manner in a social decision-making task, whereas a sub-group of the ACCg neurons do so in an other-referenced manner.<sup>72</sup> Similarly, in a competitive game, self-referenced reward outcome prediction errors correlate with activity in the ventral striatum, but, critically, belief-based prediction errors about the competitive partners action are encoded in rostral ACC (rACC).<sup>90</sup> Furthermore, in a social decision-making task involving utilizing advice from another person, learning rates for self and other are differentially computed by the ACCs and the ACCg, respectively.<sup>49</sup> Overall, although social signals have been detected in ACC, the ACCg is most clearly linked to other-referenced information processing based on accumulating evidence spanning whole-brain neuroimaging, electrophysiological recording, and anatomical specializations.<sup>48</sup>

#### Prefrontal cortex

The prefrontal cortex has many subsections, and is often thought of as the locus of higher level cognitive processes related to decision making. It is intuitive, then, many parts of the prefrontal cortex process other-referenced information. When observing erroneous choices by another individual informs an association between a specific target and a possible reward during a turn-taking decision-making task in pairs of monkeys, neurons in the dorsomedial frontal cortex encode the errors made by the partner monkey, serving a social error monitoring function,<sup>91</sup> which relies on other-referenced information. Similarly, the vmPFC encodes in humans the value of observing another person's behavior in a reward seeking task, and correlates with that individual's move towards conforming to social norms.<sup>92</sup> Other types of prediction errors are also found in the prefrontal cortex. When participants learn the contingencies between stimulus and reward outcome through direct experience or observing the action and outcome of another person, different reward-related prefrontal structures signal learning-related events for self and other. In such scenarios, the ventral striatum signals self prediction errors, the dorsolateral



**Fig. 2** Key brain regions that have been shown correlates of self-referenced prediction errors (in yellow) or other-referenced prediction errors (in purple), or both kinds of prediction errors (in green) in the domain of actions (a) and reward/value outcomes (b). For motor learning-related errors, we only provide representative areas as they are beyond the scope of this review. It is worthwhile to note that the distributions of these areas for their involvements in self-referenced or other-referenced prediction error signaling are naturally constrained by the amount of research examining different types of prediction errors. The insets with coronal magnetic resonance images indicate the sections (red line) that correspond to the Nissl-stained sagittal slices. The dotted outline around an area indicates that this area is projected medially from the lateral surface for the purpose of including the area on a more medial aspect of the brain. Adapted with permission from <http://www.brains.rad.msu.edu>, supported by the US National Science Foundation and the National Institutes of Health. ACC anterior cingulate cortex, ACCg anterior cingulate gyrus, ACCs anterior cingulate sulcus, dIPFC dorsolateral prefrontal cortex, dmPFC dorsomedial prefrontal cortex, Lhb lateral habenula, LIP lateral intraparietal area, MTG medial temporal gyrus, OFC orbitofrontal cortex, SC superior colliculus, SN substantia nigra, STS superior temporal sulcus, TPJ temporoparietal junction, vmPFC ventromedial prefrontal cortex, VS ventral striatum, VTA ventral tegmental area

prefrontal cortex (dIPFC) signals other's action prediction errors, and vmPFC signal other's outcome prediction errors.<sup>84</sup>

Furthermore, Suzuki et al.<sup>50</sup> examined the neural correlates of learning stimulus-reward outcome contingencies when participants learned the association directly and when the participants predicted which stimulus another person will likely choose, encouraging the participants to model or mentally simulate the other individual. This manipulation necessitated the use of an other-referenced prediction error, one that calculates the discrepancy between what the other person does and what the participant thought the other person would do. Again, different parts of the prefrontal cortex are engaged as a function of self-referenced and other-referenced computation. The vmPFC tracked the simulated other's prediction error in a similar manner to self, whereas the simulated other's action prediction error was signaled by the dorsomedial prefrontal cortex (dmPFC) and dIPFC.<sup>50</sup> Notably, neuronal activity in the monkey dmPFC has been shown to closely reflect the strategy of an opponent in a competitive

reward-based task, further strengthening the specialized role of dmPFC in simulating others.<sup>93</sup> This is consistent with the findings of Behrens et al.,<sup>49</sup> in which separable reward signals were computed in ACCg and ACCs for other-referenced and self-referenced reward information, respectively, and that these signals were integrated in vmPFC.

Finally, the orbitofrontal cortex (OFC) is a key cortical region for signaling reward value<sup>94</sup> and is also associated with showing value prediction error signaling.<sup>95</sup> Although OFC neurons are sensitive to social reward context involving self and other,<sup>96</sup> the reward outcome encoding of these neurons seems to be self-referenced,<sup>72</sup> suggesting that OFC may be more restricted to mediating behavioral adaptations, including adjusting to social context, in a self-referenced framework.

Encoding of various prediction errors regarding others is a signature of many reward-related regions of the brain, suggesting a tight biological link between self learning and learning about or from others. In particular, these findings endorse the notion that

comprehending and learning from the experience of another person is processed under shared predictive coding principles with particular regional specializations for the self and other domain.

#### Temporal parietal junction and mentalizing networks

When ToM is engaged, demanding the modeling of another individual, precuneus (PrCu), posterior cingulate cortex (PCC), as well as superior temporal sulcus (STS), temporal parietal junction (TPJ), and the medial prefrontal cortex (mPFC) are particularly activated over others.<sup>97</sup> STS and TPJ have long been considered as the neural hotspots for higher-level cognitions like ToM and modeling the minds of others. TPJ, in particular, has been regarded as a uniquely social cognition-focused area,<sup>98</sup> with evidence that TPJ is necessary for representing the belief of others.<sup>99</sup> A meta-analysis of ToM-related areas determined that the most reliably implicated areas are TPJ and mPFC, with activations in PrCu and STS being sensitive to the types of ToM engaged in the context of various ToM measures.<sup>100</sup> Notably, researchers have found a close link between self-referential thoughts and mentalizing of others in mPFC,<sup>101</sup> indicating how self-referenced and other-referenced information is associated with one another in one of the key regions of the mentalizing network.

Notably, recruitments of TPJ and STS are not specific to tasks designed to measure ToM. TPJ and STS are also activated in situations when considering other's information to guide one's actions, suggesting their involvements in broadly defined other-referenced computation. When participants take into account the advice of another person to make a decision about obtaining potential rewards, dmPFC, middle temporal gyrus (MTG), STS, and TPJ activations signal social prediction error.<sup>49</sup> Furthermore, when playing a simplified poker game against a human opponent and a computer algorithm, TPJ emerges as a unique region for predicting social decisions that are behaviorally relevant.<sup>102</sup> In addition, STS has been well known for their roles in social perception from visual cues.<sup>103</sup> Therefore, tracking and interpreting socially relevant information may be the fundamental building blocks of these areas constituting a so-called mentalizing network. Recently, an elegant proposal was put forward suggesting that TPJ is a computational hub in which distinct cognitive processes, like attention, memory, sensory perception, and language all converge together to generate a representation of behaviorally relevant social context.<sup>104</sup>

Corresponding to this idea, many of the nodes in this proposed mentalizing network have been observed to perform additional functions potentially relevant to other aspects of social behavior. For example, PCC has been proposed to compute subjective value<sup>105</sup> as well as other-related social processes including person perception, person updating, and first impressions.<sup>106–108</sup>

#### CONCLUDING REMARKS

Hale and Saxe<sup>109</sup> have proposed that mentalizing may be a fundamentally predictive process. Although our current understanding of how the brain implements processes described in theory–theory or simulation theory is not complete, the fact that other-referenced prediction errors appear to be represented neurally suggests that there are shared prediction-based learning mechanisms for social learning and reinforcement learning. The neural mechanisms underlying other-referenced learning may be co-opted from the predictive mechanisms used to learn for the self, one of which is prediction error signaling. Connecting the terminology of reinforcement learning and decision making to the social domain can enhance the development of ideas and methods in studying how we think about others.<sup>110</sup>

There are many additional dimensions of other-referenced learning that remain to be explored. As experimenters continue to push the limits of studying social learning, interaction, and valuation; we may find ourselves brushing up against the limits of how the brain operationalizes what is “social” and “nonsocial”. Beyond other-referenced representations in the brain, social processing can also refer to the comparison of social agents and nonsocial, yet interactive, agents. Although divergent brain areas may apply similar computations to account for self and other, the neural processes underlying social information processing may not be categorically distinct from other types of information, but rather lie on a continuum. For example, when human participants play a game with other individuals or slot machine partners programmed to display varying levels of generosity, activations in TPJ, PCC, PrCu, vmPFC, and several other regions reflected the prediction errors for generosity similarly for both human and slot machine partners.<sup>111</sup>

This finding and many others that have observed modulatory differences in brain activations between social and nonsocial information may suggest that the brain may not in fact differentiate the two categorically but processes information as a function of implementing algorithms demanded by specific behavioral constraints. Perhaps social functions could be regarded as repurposed ancestral functions of the brain evolved to deal with an organism's social environment.<sup>112</sup> Then the notion of the “social brain” should be concerned with how specific sets of commonly used computational algorithms are utilized to guide adaptive behaviors.

#### ACKNOWLEDGEMENTS

We would like to thank A.N., L.S., and D.L. for thoughtful advice on the manuscript. This work was supported by the National Institute for Mental Health (S.W.C.C., R00-MH099093, R01-MH110750, R21-MH107853), Alfred P. Sloan Foundation (S.W.C.C., FG-2015-66028), and the Natural Sciences and Engineering Research Council of Canada PGSD Fellowship (M.P., 471313).

#### AUTHOR CONTRIBUTIONS

J.A.J., M.P., C.T., and S.W.C.C. wrote the paper.

#### ADDITIONAL INFORMATION

**Competing interests:** The authors declare no competing financial interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

#### REFERENCES

- Bloom, P. *How Children Learn the Meanings of Words* (MIT Press, 2000).
- Wynn, K. Infants possess a system of numerical knowledge. *Curr. Dir. Psychol. Sci.* **4**, 172–177 (1995).
- Keil, F. C. Science starts early. *Science* **331**, 1022–1023 (2011).
- Goldman, A. I. In defense of the simulation theory. *Mind Lang.* **7**, 104–119 (1992).
- Rizzolatti, G. & Craighero, L. The mirror-neuron system. *Annu. Rev. Neurosci.* **27**, 169–192 (2004).
- Apperly, I. A. Beyond simulation–theory and theory–theory: Why social cognitive neuroscience should use its own concepts to study “theory of mind”. *Cognition* **107**, 266–283 (2008).
- Heyes, C., Jaldow, E., Nokes, T. & Dawson, G. Imitation in rats (*Rattus norvegicus*): the role of demonstrator action. *Behav. Process.* **32**, 173–182 (1994).
- Akins, C. K., Klein, E. D. & Zentall, T. R. Imitative learning in Japanese quail (*Coturnix japonica*) using the bidirectional control procedure. *Anim. Learn. Behav.* **30**, 275–281 (2002).
- Tomasello, M., Davis-Dasilva, M., CamaK, L. & Bard, K. Observational learning of tool-use by young chimpanzees. *Hum. Evol.* **2**, 175–183 (1987).
- Field, T., Woodson, R., Greenberg, R. & Cohen, D. Discrimination and imitation of facial expression by neonates. *Science* **218**, 179–181 (1982).

11. Ferrari, P. F. et al. Neonatal imitation in *Rhesus macaques*. *PLoS Biol.* **4**, e302 (2006).
12. Bandura, A. & Walters, R. H. *Social Learning Theory* (Prentice-Hall, 1977).
13. Bandura, A., Ross, D. & Ross, S. A. Vicarious reinforcement and imitative learning. *J. Abnorm. Soc. Psychol.* **67**, 601–607 (1963).
14. Manz, C. C. & Sims, H. P. Vicarious learning: the influence of modeling on organizational behavior. *Acad. Manag. Rev.* **6**, 105–113 (1981).
15. Singer, T. et al. Empathic neural responses are modulated by the perceived fairness of others. *Nature* **439**, 466–469 (2006).
16. Xu, X., Zuo, X., Wang, X. & Han, S. Do you feel my pain? Racial group membership modulates empathic neural responses. *J. Neurosci.* **29**, 8525–8529 (2009).
17. Horner, V., Proctor, D., Bonnie, K. E., Whiten, A. & de Waal, F. B. M. Prestige affects cultural learning in chimpanzees. *PLoS ONE* **5**, e10625 (2010).
18. Behrens, T. E. J., Hunt, L. T. & Rushworth, M. F. S. The computation of social behavior. *Science* **324**, 1160–1164 (2009).
19. Wimmer, H. & Perner, J. Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* **13**, 103–128 (1983).
20. Onishi, K. H. & Baillargeon, R. Do 15-month-old infants understand false beliefs? *Science* **308**, 255–258 (2005).
21. Drayton, L. A. & Santos, L. R. A decade of theory of mind research on cayo santiago: insights into rhesus macaque social cognition. *Am. J. Primatol.* **78**, 106–116 (2014).
22. Rosati, A. G., Arre, A. M., Platt, M. L. & Santos, L. R. Rhesus monkeys show human-like changes in gaze following across the lifespan. *Proc. R. Soc. B Biol. Sci.* **283**, 20160376 (2016).
23. Ferrari, P. F., Kohler, E., Fogassi, L. & Gallese, V. The ability to follow eye gaze and its emergence during development in macaque monkeys. *Proc. Natl Acad. Sci.* **97**, 13997–14002 (2000).
24. Shepherd, S. V. Following gaze: gaze-following behavior as a window into social cognition. *Front. Integr. Neurosci.* **4**, 5 (2010).
25. Flombaum, J. I. & Santos, L. R. Rhesus monkeys attribute perceptions to others. *Curr. Biol.* **15**, 447–452 (2005).
26. Hare, B., Call, J., Agnetta, B. & Tomasello, M. Chimpanzees know what conspecifics do and do not see. *Anim. Behav.* **59**, 771–785 (2000).
27. Montague, P. R., Hyman, S. E. & Cohen, J. D. Computational roles for dopamine in behavioural control. *Nature* **431**, 760–767 (2004).
28. Rescorla, R. A. & Wagner, A. R. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classic. Cond. II Curr. Res. Theory* **2**, 64–99 (1972).
29. Bellman, R. Dynamic programming and Lagrange multipliers. *Proc. Natl Acad. Sci.* **42**, 767–769 (1956).
30. Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An Introduction* Vol. 1 (MIT Press, 1998).
31. Schultz, W. & Dickinson, A. Neuronal coding of prediction errors. *Annu. Rev. Neurosci.* **23**, 473–500 (2000).
32. den Ouden, H. E. M., Friston, K. J., Daw, N. D., McIntosh, A. R. & Stephan, K. E. A dual role for prediction error in associative learning. *Cereb. Cortex* **19**, 1175–1185 (2009).
33. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
34. Bayer, H. M. & Glimcher, P. W. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47**, 129–141 (2005).
35. Schultz, W., Apicella, P., Ljungberg, T., Romo, R. & Scarnati, E. Reward-related activity in the monkey striatum and substantia nigra. *Prog. Brain Res.* **99**, 227–235 (1993).
36. Schultz, W. Behavioral theories and the neurophysiology of reward. *Annu. Rev. Psychol.* **57**, 87–115 (2006).
37. Apicella, P., Ljungberg, T., Scarnati, E. & Schultz, W. Responses to reward in monkey dorsal and ventral striatum. *Exp. Brain Res.* **85**, 491–500 (1991).
38. Montague, P., Dayan, P. & Sejnowski, T. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* **16**, 1936–1947 (1996).
39. Schultz, W., Apicella, P. & Ljungberg, T. Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J. Neurosci.* **13**, 900–913 (1993).
40. Tobler, P. N., Fiorillo, C. D. & Schultz, W. Adaptive coding of reward value by dopamine neurons. *Science* **307**, 1642–1645 (2005).
41. Matsumoto, M. & Hikosaka, A. Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature* **447**, 1111–1115 (2007).
42. Behrens, T. E. J., Woolrich, M. W., Walton, M. E. & Rushworth, M. F. S. Learning the value of information in an uncertain world. *Nat. Neurosci.* **10**, 1214–1221 (2007).
43. Niv, Y. Reinforcement learning in the brain. *J. Math. Psychol.* **53**, 139–154 (2009).
44. Daw, N. D. & Tobler, P. N. Value learning through reinforcement: the basics of dopamine and reinforcement learning. In *Neuroeconomics: Decision-Making and the Brain* 2nd Edn (eds Glimcher, P. W. & Fehr, E.) 283–298 (London, Elsevier, 2013).
45. Gläscher, J., Daw, N., Dayan, P. & O'Doherty, J. P. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron* **66**, 585–595 (2010).
46. Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P. & Dolan, R. J. Model-based influences on humans' choices and striatal prediction errors. *Neuron* **69**, 1204–1215 (2011).
47. Lee, S. W., Shimojo, S. & O'Doherty, J. P. Neural computations underlying arbitration between model-based and model-free learning. *Neuron* **81**, 687–699 (2014).
48. Apps, M. A., Rushworth, M. F. & Chang, S. W. The anterior cingulate gyrus and social cognition: tracking the motivation of others. *Neuron* **90**, 692–707 (2016).
49. Behrens, T. E. J., Hunt, L. T., Woolrich, M. W. & Rushworth, M. F. S. Associative learning of social value. *Nature* **456**, 245–249 (2008).
50. Suzuki, S. et al. Learning to simulate others' decisions. *Neuron* **74**, 1125–1137 (2012).
51. Schoenbaum, G. & Roesch, M. Orbitofrontal cortex, associative learning, and expectancies. *Neuron* **47**, 633–636 (2005).
52. Schoenbaum, G., Roesch, M. R., Stalnaker, T. A. & Takahashi, Y. K. A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. *Nat. Rev. Neurosci.* **10**, 885–892 (2009).
53. Yoshida, W., Dolan, R. J. & Friston, K. J. Game theory of mind. *PLoS Comput. Biol.* **4**, e1000254 (2008).
54. Hula, A., Montague, P. R. & Dayan, P. Monte carlo planning method estimates planning horizons during interactive social exchange. *PLoS Comput. Biol.* **11**, e1004254 (2015).
55. Yoshida, W. et al. Cooperation and heterogeneity of the autistic mind. *J. Neurosci.* **30**, 8815–8818 (2010).
56. Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J. & Frith, C. D. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* **442**, 1042–1045 (2006).
57. Hollerman, J. R. & Schultz, W. Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat. Neurosci.* **1**, 304–309 (1998).
58. Kashielyan, V., Lichtenberg, N., Chen, M., Mindy, L., Cheer, J., Roesch, M. & Matthew, R. Observation of reward delivery to a conspecific modulates dopamine release in ventral striatum. *Curr. Biol.* **24**, 2564–2568 (2014).
59. Báez-Mendoza, R., Harris, C. J. & Schultz, W. Activity of striatal neurons reflects social action and own reward. *Proc. Natl Acad. Sci.* **110**, 16634–16639 (2013).
60. Jones, R. M. et al. Behavioral and neural properties of social reinforcement learning. *J. Neurosci.* **31**, 13039–13045 (2011).
61. Fließbach, K. et al. Social comparison affects reward-related brain activity in the human ventral striatum. *Science* **318**, 1305–1308 (2007).
62. Bault, N., Joffily, M., Rustichini, A. & Coricelli, G. Medial prefrontal cortex and striatum mediate the influence of social comparison on the decision process. *Proc. Natl Acad. Sci.* **108**, 16044–16049 (2011).
63. Xiang, T., Lohrenz, T. & Montague, P. R. Computational substrates of norms and their violations during social exchange. *J. Neurosci.* **33**, 1099–1108 (2013).
64. Burke, C. J., Tobler, P. N., Schultz, W. & Baddeley, M. Striatal BOLD response reflects the impact of herd information on financial decisions. *Front. Hum. Neurosci.* **4**, 48 (2010).
65. Lohrenz, T., Bhatt, M., Apple, N. & Montague, P. R. Keeping up with the Joneses: interpersonal prediction errors and the correlation of behavior in a tandem sequential choice task. *PLoS Comput. Biol.* **9**, e1003275 (2013).
66. King-Casas, B. et al. Getting to know you: reputation and trust in a two-person economic exchange. *Science* **308**, 78–83 (2005).
67. Klucharev, V., Hytönen, K., Rijpkema, M., Smidts, A. & Fernández, G. Reinforcement learning signal predicts social conformity. *Neuron* **61**, 140–151 (2009).
68. Xiang, T., Ray, D., Lohrenz, T., Dayan, P. & Montague, P. R. Computational phenotyping of two-person interactions reveals differential neural response to depth-of-thought. *PLoS Comput. Biol.* **8**, e1002841 (2012).
69. Kolling, N. et al. Value, search, persistence and model updating in anterior cingulate cortex. *Nat. Neurosci.* **19**, 1280–1285 (2016).
70. Shenhav, A., Cohen, J. D. & Botvinick, M. M. Dorsal anterior cingulate cortex and the value of control. *Nat. Neurosci.* **19**, 1286–1291 (2016).
71. Ebitz, R. B. & Hayden, B. Y. Dorsal anterior cingulate: a Rorschach test for cognitive neuroscience. *Nat. Neurosci.* **19**, 1278–1279 (2016).
72. Chang, S. W. C., Gariépy, J.-F. & Platt, M. L. Neuronal reference frames for social decisions in primate frontal cortex. *Nat. Neurosci.* **16**, 243–250 (2013).
73. Apps, M. A. & Ramnani, N. The anterior cingulate gyrus signals the net value of others' rewards. *J. Neurosci.* **34**, 6190–6200 (2014).
74. Haroush, K., Williams & Ziv, M. Neuronal prediction of opponent's behavior during cooperative social interchange in primates. *Cell* **160**, 1233–1245 (2015).



75. Lockwood, P. L., Apps, M. A., Roiser, J. P. & Viding, E. Encoding of vicarious reward prediction in anterior cingulate cortex and relationship with trait empathy. *J. Neurosci.* **35**, 13720–13727 (2015).
76. Apps, M. A. J., Lesage, E. & Ramnani, N. Vicarious reinforcement learning signals when instructing others. *J. Neurosci.* **35**, 2904–2913 (2015).
77. Hill, M. R., Boorman, E. D. & Fried, I. Observational learning computations in neurons of the human anterior cingulate cortex. *Nat. Commun.* **7**, 12722 (2016).
78. Jeon, D. et al. Observational fear learning involves affective pain system and Cav1.2 Ca<sup>2+</sup> channels in ACC. *Nat. Neurosci.* **13**, 482–488 (2010).
79. Singer, T. et al. Empathy for pain involves the affective but not sensory components of pain. *Science* **303**, 1157–1162 (2004).
80. Chang, S. W. C., Winecoff, A. A. & Platt, M. L. Vicarious reinforcement in rhesus macaques (*Macaca mulatta*). *Front. Neurosci.* **5**, 27 (2011).
81. Chang, S. Coordinate transformation approach to social interactions. *Front. Neurosci.* **7**, 147 (2013).
82. Rudebeck, P. H., Buckley, M. J., Walton, M. E. & Rushworth, M. F. S. A role for the macaque anterior cingulate gyrus in social valuation. *Science* **313**, 1310–1312 (2006).
83. Tomlin, D. et al. Agent-specific responses in the cingulate cortex during economic exchanges. *Science* **312**, 1047–1050 (2006).
84. Burke, C. J., Tobler, P. N., Baddeley, M. & Schultz, W. Neural mechanisms of observational learning. *Proc. Natl Acad. Sci.* **107**, 14431–14436 (2010).
85. Holroyd, C. B. & Coles, M. G. Dorsal anterior cingulate cortex integrates reinforcement history to guide voluntary behavior. *Cortex* **44**, 548–559 (2008).
86. Matsumoto, M., Matsumoto, K., Abe, H. & Tanaka, K. Medial prefrontal cell activity signaling prediction errors of action values. *Nat. Neurosci.* **10**, 647–656 (2007).
87. Kennerley, S. W., Walton, M. E., Behrens, T. E., Buckley, M. J. & Rushworth, M. F. Optimal decision making and the anterior cingulate cortex. *Nat. Neurosci.* **9**, 940–947 (2006).
88. Amemori, K.-i & Graybiel, A. M. Localized microstimulation of primate pregenual cingulate cortex induces negative decision-making. *Nat. Neurosci.* **15**, 776–785 (2012).
89. Apps, M. A. J., Lockwood, P. L. & Balsters, J. H. The role of the midcingulate cortex in monitoring others' decisions. *Front Neurosci.* **7**, 251 (2013).
90. Zhu, L., Mathewson, K. E. & Hsu, M. Dissociable neural representations of reinforcement and belief prediction errors underlie strategic learning. *Proc. Natl Acad. Sci.* **109**, 1419–1424 (2012).
91. Yoshida, K., Saito, N., Iriki, A. & Isoda, M. Social error monitoring in macaque frontal cortex. *Nat. Neurosci.* **15**, 1307–1312 (2012).
92. Chung, D., Christopoulos, G. I., King-Casas, B., Ball, S. B. & Chiu, P. H. Social signals of safety and risk confer utility and have asymmetric effects on observers' choices. *Nat. Neurosci.* **18**, 912–916 (2015).
93. Seo, H., Cai, X., Donahue, C. H. & Lee, D. Neural correlates of strategic reasoning during competitive games. *Science* **346**, 340–343 (2014).
94. Kringelbach, M. L. The human orbitofrontal cortex: linking reward to hedonic experience. *Nat. Rev. Neurosci.* **6**, 691–702 (2005).
95. O'Doherty, J. et al. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* **304**, 452–454 (2004).
96. Azzi, J. C. B., Sirigu, A. & Duhamel, J.-R. Modulation of value representation by social context in the primate orbitofrontal cortex. *Proc. Natl Acad. Sci. USA* **109**, 2126–2131 (2012).
97. Frith, C. D. & Frith, U. The neural basis of mentalizing. *Neuron* **50**, 531–534 (2006).
98. Saxe, R. & Kanwisher, N. People thinking about thinking people: The role of the temporo-parietal junction in "theory of mind". *Neuroimage* **19**, 1835–1842 (2003).
99. Samson, D., Apperly, I. A., Chiavarino, C. & Humphreys, G. W. Left temporoparietal junction is necessary for representing someone else's belief. *Nat. Neurosci.* **7**, 499–500 (2004).
100. Schurz, M., Radua, J., Aichhorn, M., Richlan, F. & Perner, J. Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neurosci Biobehav. Rev.* **42**, 9–34 (2014).
101. Mitchell, J. P. Inferences about mental states. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **364**, 1309–1316 (2009).
102. Carter, R. M., Bowling, D. L., Reeck, C. & Huettel, S. A. A distinct role of the temporal-parietal junction in predicting socially guided decisions. *Science* **337**, 109–111 (2012).
103. Allison, T., Puce, A. & McCarthy, G. Social perception from visual cues: role of the STS region. *Trends Cogn. Sci.* **4**, 267–278 (2000).
104. Carter, R. M. & Huettel, S. A. A nexus model of the temporal-parietal junction. *Trends Cogn. Sci.* **17**, 328–336 (2013).
105. Kable, J. W. & Glimcher, P. W. The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* **10**, 1625–1633 (2007).
106. Mende-Siedlecki, P., Baron, S. G. & Todorov, A. Diagnostic value underlies asymmetric updating of impressions in the morality and ability domains. *J. Neurosci.* **33**, 19406–19415 (2013).
107. Mende-Siedlecki, P. & Todorov, A. Neural dissociations between meaningful and mere inconsistency in impression updating. *Soc. Cogn. Affect Neurosci.* **11**, 1489–1500 (2016).
108. Schiller, D., Freeman, J. B., Mitchell, J. P., Uleman, J. S. & Phelps, E. A. A neural mechanism of first impressions. *Nat. Neurosci.* **12**, 508–514 (2009).
109. Koster-Hale, J. & Saxe, R. Theory of mind: a neural prediction problem. *Neuron* **79**, 836–848 (2013).
110. Beam, E., Appelbaum, L. G., Jack, J., Moody, J. & Huettel, S. A. Mapping the semantic structure of cognitive neuroscience. *J. Cogn. Neurosci.* **26**, 1949–1965 (2014).
111. Hackel, L. M., Doll, B. B. & Amodio, D. M. Instrumental learning of traits versus rewards: dissociable neural correlates and effects on choice. *Nat. Neurosci.* **18**, 1233–1235 (2015).
112. Chang, S. W. et al. Neuroethology of primate social behavior. *Proc. Natl Acad. Sci.* **110**, 10387–10394 (2013).



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017