

The neurobiology of rewards and values in social decision making

Christian C. Ruff and Ernst Fehr

Abstract | How does our brain choose the best course of action? Choices between material goods are thought to be steered by neural value signals that encode the rewarding properties of the choice options. Social decisions, by contrast, are traditionally thought to rely on neural representations of the self and others. However, recent studies show that many types of social decisions may also involve neural value computations. This suggests a unified mechanism for motivational control of behaviour that may incorporate both social and non-social factors. In this Review, we outline a theoretical framework that may help to identify possible overlaps and differences between the neural processes that guide social and non-social decision making.

Of all the decisions in your life, which were the most important? Chances are high that many of them had a social component, be it when deciding about others (for example, whether to marry someone), for others (for example, which school to send your child to) or jointly with others (for example, whether to buy a house with your partner). Pathological changes specifically in the social aspects of decision making are pervasive in many psychiatric¹ and neurological² disorders and can have devastating consequences for individual and collective welfare. A good understanding of the neural basis of social decision making is thus an important aim of scientists in many disciplines, from basic neuroscience to medicine, psychology and economics.

The neural mechanisms underlying social and non-social types of choices have traditionally been examined from different theoretical angles. In studies of non-social choices (for example, purchasing decisions), researchers have often focused on neural value computations associated with the rewarding properties of the choice options^{3,4}. On the basis of this research, several brain structures have been proposed as elements of a brain valuation system (BOX 1). The neural value signals in these structures are thought to represent a ‘common currency’ for assessing the motivational relevance of all possible stimuli or actions, which is essential for guiding decision making across varying contexts^{5,6}. By contrast, researchers studying social decisions have traditionally paid less attention to motivational processes, such as reward and value coding; instead, they have focused on identifying neurocognitive processes that may have uniquely evolved to guide social behaviour⁷. This emphasis on

‘social-specific cognition’ has identified, for example, that different visual areas seem to be functionally specialized for the perception of faces or bodies⁸ and that the dorsomedial prefrontal cortex (dmPFC) and the temporoparietal junction (TPJ) may be specifically involved in representing the intentions, emotions or actions of other people^{9–11}. The idea of ‘social-specific cognition’ thus implies that the control of social and non-social behaviour should fundamentally differ in terms of neural architecture and information processing demands¹². However, this view is challenged by recent findings that social decisions may also draw on value-related brain processes that strongly resemble those seen during non-social decisions. This apparent overlap in neural representations of motivational relevance for social and non-social choices is often taken to implicitly extend the ‘common currency’ idea to decisions based on social factors. This proposal — that choices in both social and non-social situations are steered by identical neural value computations — therefore questions whether the motivational control of social behaviour requires dedicated neural processes.

In this Review, we propose a framework for studying the possible neurobiological links in the motivational control of social and non-social decision making. We first outline two competing schemas of the value signals that drive both types of decisions; these schemas represent different conceptual combinations of the ideas of a ‘common currency’ and of ‘social-specific cognition’. We then discuss the congruency of these two schemas with empirical results. A major challenge in this context is the sheer diversity of social situations and neural effects that are being investigated under the umbrella terms

Laboratory for Social and Neural Systems Research, Department of Economics, University of Zurich, Bluemlisalpstrasse 10, CH-8006 Zurich, Switzerland.
e-mails: christian.ruff@econ.uzh.ch; ernst.fehr@econ.uzh.ch
doi:10.1038/nrn3776
Published online 2 July 2014

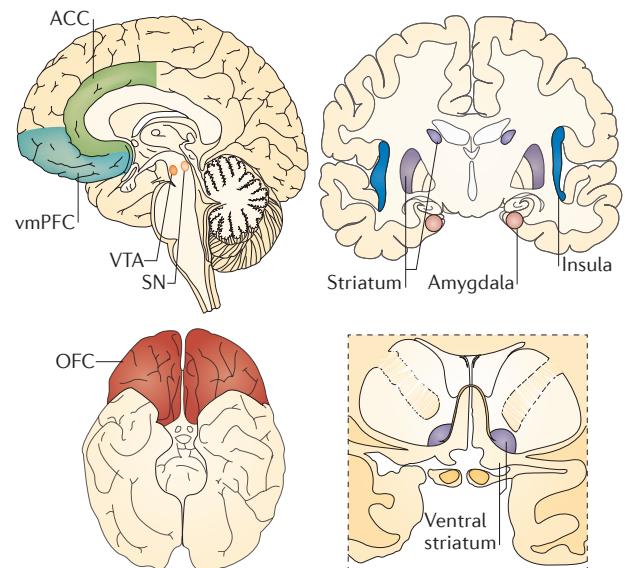
Box 1 | Neural representations of value

Value-based decisions involve at least three distinct stages that require different value signals: choice of an appropriate action, evaluation of the choice outcome and learning from the outcome. Empirical studies of non-social decision making and learning have indeed identified three different types of neural value representations that correspond to these stages and that are associated with activity changes in different neural structures (see the figure for a schematic representation).

Experienced-value signals. The receipt and consumption of rewards (for example, a choice outcome) are associated with neural activity in the orbitofrontal cortex (OFC)^{122,123}. Other structures that are often found to be activated during reward experiences include the amygdala¹²⁴, the anterior insula and the anterior cingulate cortex (ACC)¹²⁵; the latter two structures are also often activated in ‘punishing’ events such as pain¹²⁶.

Anticipated-value signals. Signals associated with the anticipation of rewards or value learning have been identified in dopaminergic neurons⁵¹ in the ventral tegmental area (VTA) and substantia nigra (SN). During learning, these neurons encode the ‘reward-prediction error’ (that is, the difference between the experienced value and the anticipated value of rewards)¹²⁷. Prediction errors are presumed to signal the need to update anticipated values (for example, expectancies of reward) and have also been consistently observed in the dopaminergically innervated ventral striatum in human functional MRI studies (but see REF. 128). The striatum and other dopaminergically innervated structures (for example, the OFC and amygdala) are thus often thought to constitute some basic reward circuitry that dynamically encodes the anticipated value of stimuli¹²⁹.

Decision-value signals. These signals are thought to integrate all anticipated values and costs associated with the different options into a single quantity that is interpreted to make the choice. Much recent functional MRI research in humans has implicated the ventromedial prefrontal cortex (vmPFC; comprising the medial OFC and portions of the mPFC) in computations of these signals at the time of the decision⁸. It can be difficult to dissociate an anticipated value from a decision value, but vmPFC activity has been proposed to reflect valuation during decision making rather than valuation during subsequent reward administration and associated reward prediction¹³⁰. Debate is ongoing about whether choices regarding different courses of action rely on decision-value computations in more motor-related structures, such as the ACC¹³¹.



of ‘social decisions’, ‘rewards’ and ‘values’. We therefore group the discussion of empirical findings about social decision making by the type of social situation and the neural value process under consideration. We finish the Review with suggestions for future work that may clarify how neural processes devoted to value coding and reward coding can account for both social and non-social factors in influencing choices.

Social valuation: a conceptual framework

There are two schemas that illustrate how neural value representations underlying social decisions could relate to those driving non-social choices (FIG. 1). These two schemas have been implicitly alluded to in the literature but have rarely been explicitly compared. Although they are not mutually exclusive, the two schemas make contradicting predictions.

The ‘extended common currency schema’ (FIG. 1 a) assumes that identical neural processes assign motivational relevance to social and non-social factors. However, these value-related processes may incorporate input from different brain areas that compute cognitive information that is relevant for the social or non-social choice. This schema therefore predicts that both types

of choices induce similar activity in the brain value system but that these shared value representations change their functional connectivity with other, domain-specific brain regions in social versus non-social choice contexts.

By contrast, the ‘social-valuation-specific schema’ (FIG. 1 b) proposes that social rewards and values are processed in a dedicated neural circuitry that evolved specifically to deal with interactions with others. In this schema, the neurons in this circuitry derive values based on social information by using types of neural value computations similar to those used by the neurons that encode non-social value representations. The two types of neurons may be located in close spatial proximity (perhaps owing to the need for mutual interactions) or in different brain regions. This schema therefore predicts that social and non-social neural value signals are implemented in different spatial patterns of brain activity but nevertheless follow similar computational principles.

Classes of social valuation. Before we discuss empirical support for these two schemas, it is important to note that studies of social decision making address a wide range of contexts and choices that — at first glance — mainly seem to be linked by the fact that more than one

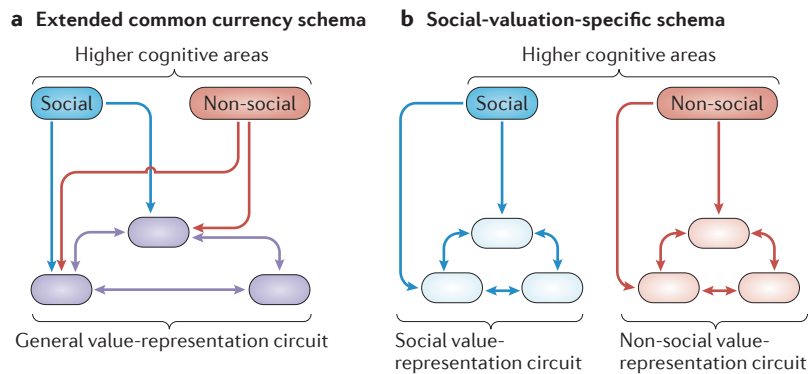


Figure 1 | Two schemas for neural value computation in social versus non-social contexts. The figure illustrates two competing schemas for how the brain may determine the value of social and non-social factors during decision making. Please note that this illustration does not show specific brain areas and connections but rather depicts abstract principles of how brain areas and their interactions could implement these computations. **a** | The ‘extended common currency schema’ assumes that a single neural circuit (shown in purple) determines the motivational significance of both social and non-social events. The activity of this circuit represents the integrated value of all factors that are relevant for a choice; however, the perceptual and cognitive information that is relevant for these unified value computations may differ between social and non-social choices and may be provided by distinct domain-specific brain areas (shown in blue and red, respectively). Social and non-social decisions should thus result in identical activity in reward-related brain areas (if their motivational and hedonic properties are matched) but may differ in the functional connectivity of these shared valuation areas with other brain regions. **b** | The ‘social-valuation-specific schema’ is based on the social brain hypothesis, which proposes that social aspects of the environment are processed in a neural circuitry that evolved specifically to deal with these demands. In this schema, neural valuation of social and non-social factors engages neural processes that follow similar computational principles but are implemented in distinct neuronal populations that are specialized for each type of information. These specialized neurons may be located in different areas (shown in light blue and light red, respectively) or in close proximity within a value-processing brain region. In either case, the two types of valuation neurons (social versus non-social) will predominantly show functional interactions with other areas that are specialized for either social or non-social cognitive functions, respectively.

person is involved. Viewed from this general perspective, it can be difficult to understand which aspects of these choice situations may be rewarding (or punishing) for the observer. We therefore propose three classes of situations in which the brain engages in valuation of a specifically social nature, which differ in terms of either the object or the frame of reference for valuation. The first of these classes includes all situations in which an agent assesses how specific other individuals and their behaviour affect his or her own well-being (that is, the object of valuation is another person and his or her actions). The second class concerns situations in which an agent’s brain values choice options and outcomes vicariously for others (that is, value signals are computed in the frame of reference of another person). The third class comprises situations in which an agent guides his or her behaviour to comply with normative social principles. These principles provide an abstract social frame of reference for assessing the value of actions and outcomes. BOX 2 explains these three classes in more detail.

Below, we discuss findings from studies of social decision making in these three classes of situations. For each class, we consider how the observed behaviour and brain

activity provide evidence for or against the extended common currency schema and the social-valuation-specific schema of social choice.

Neural valuation of other people

Experiencing rewards or punishments from specific others. Humans frequently interact with others, and many aspects of these direct interactions may be rewarding (or punishing). A basic example is the mere perception of an attractive person, which can elicit neural activity in several areas of the brain’s reward circuitry, including the orbitofrontal cortex (OFC)^{13,14}, ventral striatum¹⁵ and amygdala^{16,17}. Similar effects have been found when participants viewed erotic photos^{18,19} or faces with positive affect^{20–22}. Although this neural activity may arguably reflect low-level visual differences between different types of pictures, several studies have linked the neural activations to some measures of behavioural preference^{6,17–19,22–24}. For example, one study found that the increase in the blood-oxygen-level-dependent (BOLD) signal in the OFC during viewing more versus less attractive faces predicted the amount of money participants were willing to spend to view the more attractive faces²⁴. Moreover, choices between viewing different attractive faces²⁵ or erotic photos¹⁸ follow similar economic principles (for example, delay discounting) as those established for choices between non-social rewards. These studies thus suggest that merely seeing others who are attractive, in a positive mood or in erotic poses already constitutes an experienced value that is encoded by similar activity in the brain reward system as for typical non-social values.

Regarding more abstract aspects of social interactions, it has been proposed that social gestures (for example, verbal approval) also have rewarding properties that may be processed in brain networks similar to those used in the processing of non-social rewards. In one functional MRI (fMRI) study²⁶, participants completed a non-social gambling task (in which monetary rewards were the outcome) and a social reward task (in which others’ positive evaluations of the participant’s personality were shown). Despite differences in the task and stimuli, social and monetary reward outcomes led to comparable BOLD activations in the ventral striatum. Similar activations have also been found in the ventral striatum and ventromedial PFC (vmPFC) when participants read that other depicted individuals like^{27–30}, understand³¹ or want to meet them³². Finally, ventral striatum activity has been observed during outcomes that indicate changes in the social status of the participants in the social interactions³³. Thus, neural experienced values during social interactions can encode both immediate hedonic aspects and longer-lasting social consequences of basic social gestures (for example, changes in reputation or status).

Conceptually related findings have been reported for social punishments. Social exclusion, as simulated in the context of a game paradigm, led to BOLD signal increases in the anterior cingulate cortex (ACC)³⁴. These rejection-related BOLD responses have been proposed to reflect processes that are related to (socially induced) pain, as they correlated with social stress-related inflammatory responses³⁵, were reduced by administration of

Delay discounting

The systematic decrease in a given item’s value with increasing delay until it will be received. Excessive delay discounting is often thought to underlie impulsive choices.

Box 2 | Three classes of situations requiring social valuation

To simplify the interpretation of the wide range of functional MRI findings on social choices, one can divide social situations into three classes that involve different targets and reference frames for neural valuation processes.

In the first of these classes, agents assess how specific other individuals and their behaviour affect the agent's own well-being. This class comprises all situations in which the agent assesses whether the social interaction with a specific other person is rewarding or punishing to him- or herself, which may require neural signals akin to value signals for non-social goods.

The second class of situation concerns vicarious valuation of choice options and outcomes for others, even though the agent does not directly benefit from the observed choices and outcomes. The 'social' aspect of the valuation processes in this class of situation consists of the fact that the agent values outcomes in the frame of reference of another person rather than in his or her own.

The third class comprises situations in which choice options and outcomes are valued in a reference frame that is specified by normative social principles. These social principles can run counter to the agent's self-interest and can guide behaviour even during fully anonymous social interactions. Thus, in this class of situation, the agent has to consider whether abstract social constellations (for example, outcome distributions) are congruent with normative social principles.

Studies of decision making have examined distinct stages of the choice process that involve different types of value signals (experienced value, anticipated value and decision value; see BOX 1). We therefore propose that neural value representations during social decisions in the three classes of situations can be organized as shown in this table (which includes an example of choice, outcome and learning for each class of situation).

Decision-making stage (value type)	Target (and reference frame) of valuation		
	Value of other people (for oneself)	Value of other people's experiences (for them)	Value of social constellation (for normative social principle)
Choice (decision value and anticipated value)	Deciding whether to marry someone	Choosing a school for your child	Altruistic punishment of norm violations
Outcome (experienced value)	Being smiled at	Empathy with someone's pain	Enjoying fair distributions
Learning (anticipated value)	Learning about someone's trustworthiness	Learning about someone else's preferences	Changing your opinion to increase social conformity

a conventional painkiller³⁶ and overlapped with BOLD responses to physical pain³⁷. Moreover, in a direct comparison of punishments and rewards, receiving social feedback that a depicted partner had rejected the participant's romantic interest activated the ACC, whereas receiving feedback indicating a match in interest activated the vmPFC³².

Thus, several aspects of direct social interactions (for example, perceiving others, receiving interest or approval and being excluded or rejected) have been linked to neural activity in brain structures that are also involved in encoding non-social rewards. Moreover, the rewarding properties of social stimuli can also relate to the strength of functional connectivity of these valuation areas with regions involved in social cognition, such as the TPJ³⁸. These results therefore provide support for the extended common currency schema. However, only a few studies have directly compared activity patterns elicited by social outcomes with those due to non-social financial rewards^{19–21,24,26}. Although these studies have found areas of overlapping activity, they have also identified activity differences. For example, one study indicated that financial rewards and erotic pictures elicit activity in distinct portions of the OFC¹⁹, whereas other studies showed that either type of reward uniquely activates voxels in the ventral striatum²⁶ or the vmPFC²⁴. These results may therefore also be consistent with predictions from the social-valuation-specific schema. As few studies have parametrically varied the degree of reward and measured

the hedonic and behavioural impact of the rewards, it is difficult to determine whether slightly different activation patterns for social versus non-social rewards and punishments indeed reflect fundamentally different neural computations or simply show differences in the reward value of the different types of feedback.

Decisions involving specific other individuals. All of the studies discussed above investigated neural expected values at the time of the (rewarding or non-rewarding) outcome. Different types of neural value representation (namely, decision values and anticipated values; see BOX 1) might be needed during actual choices that require participants to rely on subjective value estimates in the absence of any explicit reward.

Several of the studies mentioned above have provided evidence for such value representations. For example, value-related activity can already be observed in the ventral striatum and vmPFC during the mere anticipation of social gestures or face stimuli with rewarding^{20,21,28,29,39} or punishing⁴⁰ properties. Interestingly, studies have shown that the presence of observers during decision making — and thus the possibility of future approval and reputation gains — changed value-related responses in the ventral striatum during risky choices⁴¹ and altruistic choices³⁹. The degree of change in ventral striatum activity during social observation correlated with changes in overt behaviour^{39,41}, suggesting that the mere anticipation

of peer feedback may introduce a neural ‘anticipated value’ that changes decision-value computations and, therefore, choice.

A slightly different line of research has examined neural decision values during choices about voluntary giving to specific others. Several studies suggest that humans make such decisions by integrating the monetary costs and social benefits of the decisions in value-related brain areas^{42–45}. Voluntary giving may therefore resemble a social form of purchase decision (the ‘purchased item’ is the specific other’s welfare). Giving money to liked others has indeed been found to elicit a neural ‘experienced value’: sharing a financial reward with a friend elicited more subjective happiness and more ventral striatum activity than sharing the same reward with a confederate or computer⁴⁶. During decisions to transfer a portion of one’s endowment to a specific charity, BOLD activity in the ventral striatum is stronger than during mandatory donations⁴³ or during decisions about similar payments to oneself⁴². These BOLD activity increases in the ventral striatum may therefore reflect valuation of anticipated hedonic feelings⁴³ or social consequences (for example, gains in reputation or social status)³⁹ of the donation compared with the short-term monetary consequences. In line with this interpretation, a study found that BOLD activity in the OFC co-varied with the freely chosen amount of money given to a specific charity⁴⁵. As the OFC showed functional connectivity with the anterior insula and the posterior superior temporal sulcus during these decisions, it was argued that social valuation computations in the vmPFC may depend on input from specialized regions that compute the deservingness of each charity.

The findings of these studies are therefore broadly consistent with predictions from the extended common currency schema, as they indicate that decisions about payments to a specific other — in these cases, mostly charitable organizations — include valuation processes implemented in the same neural structures as those involved in decisions about non-social goods. In addition, these neural structures also seem to receive specialized inputs from other brain areas that are involved in social cognition.

Learning about others. Interactions with a specific person often involve repeated contacts, and information about the other person strongly shapes behaviour in these interactive contexts. A series of studies using the trust game have suggested that learning about the trustworthiness of others is associated with activity in the brain’s reward circuitry that mimics activity seen during learning about (non-social) reward values of objects^{47–52}. For instance, information about the trustee’s moral character⁴⁷ or experience with a trustee’s behaviour⁵⁰ changed both neural activity in the striatum during the outcome phase of each round and subsequent investment behaviour in the game, and these effects changed over the course of the experiment. Specifically, striatal responses tied to trust decisions were expressed during the outcome phase of trials early in the experiment (when opponents’ reputations were not yet known) and then shifted forward to the decision phase of each trial,

possibly now indicating learned reward predictions for each opponent⁴⁸. Moreover, these trust-learning effects are strongly modulated by prior expectations (as is the case for non-social reward-learning effects⁵¹): unexpected reciprocity in any trial induced stronger ventral striatum activity and higher investments in subsequent encounters compared with expected reciprocity⁵³, suggesting that this activity may reflect a reward prediction error similar to that observed during non-social forms of reward learning. However, striatal prediction errors and behavioural learning during the outcome of trust decisions can also be suppressed when information about the opponents’ trustworthiness is presented⁵⁴, suggesting that strong priors can overrule the importance of new information gathered during single trials.

The outcomes of the trust game used in these studies were purely financial, and one may therefore wonder whether the learning effects described above are really uniquely social or simply reflect that the other people predict a financial outcome. The latter interpretation is contradicted by studies of non-financial social rewards: faces with positive affect²² or symbolic gestures⁵⁵ that are fully unrelated to monetary outcomes elicit behavioural effects and prediction-error signalling in the ventral striatum similar to those seen for monetary rewards. Thus, stimuli and actions that are uniquely encountered in social interactions can reinforce behaviour through neural mechanisms that are similar to those underlying non-social reinforcement with money (FIG. 2a). However, social-specific effects were observed for (punishment) learning by trust violations in the context of a pharmacological fMRI study⁴⁹ that involved oxytocin, a neuropeptide that plays a central part in the ability to form social attachments in mammals and that can enhance trusting behaviour⁵⁶. In the placebo participants taking part in this study, receiving information that an opponent had breached their trust decreased trusting behaviour and increased BOLD levels in the amygdala and the caudate. As these behavioural and neural effects were not seen in a lottery game that matched this experiment in terms of financial investments and outcomes, they were specifically tied to the social dimension of breaches of trust rather than their financial consequences. Crucially, these effects were strongly reduced in participants who had been given oxytocin⁴⁹ (FIG. 2b). This suggests that oxytocin influences trusting behaviour by blocking social experience-related changes in neural activity in the amygdala and caudate. At a more general level, this finding illustrates how experimental modulations of neural function (for example, pharmacological interventions) can provide evidence that social decision making is causally steered by neural valuation processes (BOX 3).

The findings on social learning by trust violation reviewed above seem to be consistent with the social-valuation-specific schema, as they demonstrate neural and behavioural learning effects that are unique to social information. A similar point is made by several recent demonstrations that, during social interactions, prediction-error computations can be observed in neural structures outside the classic reward system if the updated predictions concern specific features of

Trust game

An experimental paradigm to measure trust between anonymous players. Player 1 decides how much of his or her money to transfer to player 2. The transferred amount is multiplied by a factor > 1 before player 2 decides how much money to give back to player 1, thus honouring the trust evident in the initial transfer.

Reciprocity

A behavioural strategy whereby an individual responds to another agent’s kind (or hostile) action with a kind (or hostile) action. Reciprocity lies at the heart of ‘tit-for-tat’ and is often thought to be an important motivation for altruism.

Priors

Beliefs about a state of the world that can strongly influence the interpretation of new information. The concept is taken from Bayesian statistics, where incoming evidence is multiplied with a prior estimate to determine an event’s posterior likelihood (which is the new prior before the next information arrives).

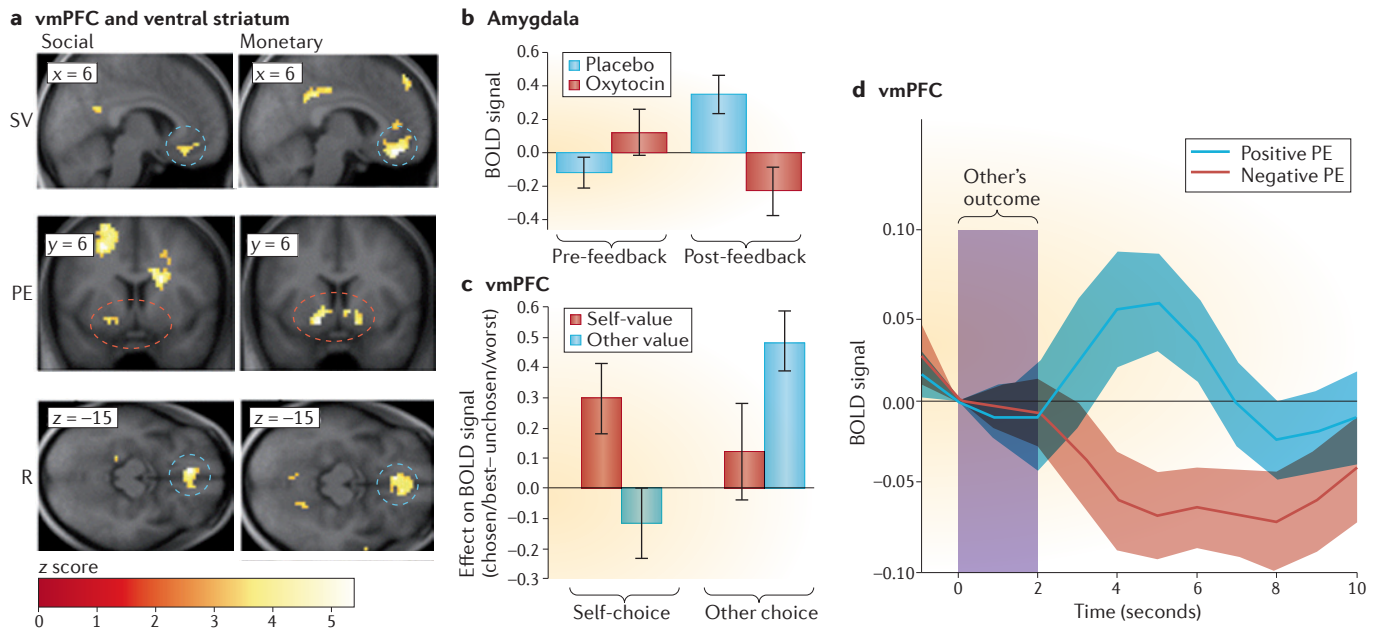


Figure 2 | Social influences on brain activity during reward processing. **a** | ‘Social’ outcomes (emotional faces) and ‘monetary’ outcomes (financial wins and losses) in a value-learning task elicit similar activation in both the ventromedial prefrontal cortex (vmPFC; circled in blue) and the ventral striatum (circled in orange). Specifically, the two types of task induce similar subjective value (SV) signals during choices as well as reward (R) and prediction-error (PE) signals during outcomes. **b** | Oxytocin blocks learning-related activity changes in the amygdala during decisions related to trust. Feedback about breaches of trust lead to decreases in trust and trust-correlated blood-oxygen-level-dependent (BOLD) signal increases in placebo-treated individuals but not in oxytocin-treated individuals. **c** | Decision-value coding in the vmPFC flexibly adapts to different social frames of reference. When participants make choices for themselves (‘self-choice’), then vmPFC BOLD signals reflect the decision values of the relevant options (‘self-value’). By contrast, when choices are made for another person (‘other choice’), vmPFC BOLD signals reflect the decision values of the options for this person (‘other value’). **d** | The BOLD signal in the vmPFC reflects a vicarious prediction error during observation of another person’s choices. The plot shows the time course of BOLD activity evoked by observing the other’s outcome (time 0) when the outcome was better (blue) or worse (red) than expected. Part **a** is adapted with permission from REF. 22, Oxford University Press. Part **b** is adapted with permission from REF. 49, Cell Press/Elsevier. Part **c** is adapted with permission from REF. 70, Cell Press/Elsevier. Part **d** is adapted with permission from REF. 74, National Academy of Sciences.

the social interaction or the partner (rather than non-social rewards). For instance, the posterior temporal sulcus can encode prediction errors related to a mismatch in romantic interest from an interaction partner³², the credibility of a confederate’s advice⁵⁷ or the influence of one’s own action on an opponent’s chosen strategy⁵⁸. Importantly, in the latter two studies, these social-reward prediction errors occurred in parallel to financial-reward prediction errors expressed in the ventral striatum. Whether these two prediction-error representations are integrated by functional connectivity between these areas⁵⁸ or by their joint influence on the vmPFC⁵⁷ is an open question.

Vicarious neural valuation

Vicarious experience of rewards and punishment.

Humans and other species have the capability to learn by observation. Current debate centres on whether this learning depends on neural mechanisms similar to those for learning based on personal reward experience. Several studies have shown that merely observing others receiving rewards and punishments can trigger neural experienced value representations similar to those elicited by a personal experience of these stimuli. One study showed that

the reported pleasure of watching others win monetary rewards was associated with BOLD activity increases in the ventral striatum that overlapped with (but were weaker than) activations elicited when participants won these rewards themselves⁵⁹. In another study, monetary wins for the participants or for a friend elicited ventral striatum activity that was higher than activity evoked by comparable wins for an unrelated partner⁶⁰. Similarly, vicarious punishments, such as watching others receive painful shocks⁶¹ or smelling disgusting odours⁶², can trigger neural responses in the ACC and the insula that are similar to those induced by the personal experience of these shocks or smells. Such vicarious activation may even be induced by very abstract punishments, such as embarrassing social situations⁶³. Importantly, these vicarious neural experienced values can predict later choice behaviour: activation of the anterior insula (or ventral striatum) in subjects who were viewing other people receiving painful shocks correlated with their subsequent decisions to help (or not to help) the observed person by choosing to endure some of these painful shocks themselves⁶⁴. These latter neural and behavioural effects were strongly modulated by whether the observed person was perceived as a member of the subject’s own social group or a rival

Box 3 | The causal role of social rewards and values in social behaviour

Many of the proposed links between social decision making (and social learning) and neural reward processing come from neuroimaging studies in healthy participants and thus rely on correlations between task parameters and measured blood flow throughout the brain. This raises the question of whether the observed neural responses are merely correlated with the observed behaviour or whether they play a causal part. Studies in clinical populations with pathological social behaviour and experiments using methods to influence neural activity suggest such a causal role of reward-related brain processes for many aspects of social behaviour.

In various psychiatric disorders, abnormal social behaviour in trust games or the prisoner's dilemma game is associated with distinct patterns of neural activity in the ventral striatum¹³², cingulate cortex¹³³, the insula¹³⁴ or the orbitofrontal cortex and amygdala¹³⁵; such altered patterns of activation may even provide biomarkers for the diagnosis and treatment of these disorders¹. Moreover, patients with autism are insensitive to reputation concerns during altruistic choices¹³⁶, pointing to a lack of a specifically social motivation. In neurology, stroke-related lesions of reward-related brain areas have been found to strongly affect social aspects of behaviour. For example, a patient with a limbic system lesion encompassing the orbitofrontal cortex and amygdala showed specific difficulties in reasoning about social exchange but not about formally similar non-social problems¹³⁷. Moreover, damage to the ventromedial prefrontal cortex was found to lead to low offers¹³⁸ and acceptance rates^{139,140} in the ultimatum game; it has been proposed that patients with such lesions may be less sensitive to inequality in several economic games¹³⁹ and may fail to integrate social and non-social signals into a decision value¹⁴⁰.

Direct manipulation of neural activity in subcortical and medial reward-related brain areas is difficult, as current non-invasive brain stimulation methods available for use in humans mainly target brain areas at the cortical surface. However, studies have shown that manipulation of activity in various prefrontal areas using transcranial magnetic or direct-current stimulation can lead to lower acceptance rates in the ultimatum game¹⁴¹, to lower investments in one's reputation as a trustee in the trust game¹⁴² and to more or less pronounced compliance with social norms¹⁴³. Whether these effects on specifically social decisions reflect remote influences from the prefrontal cortex on reward-processing neural structures could be investigated using combinations of brain stimulation methods with neuroimaging^{113,119}.

group, which suggests that vicarious neural representations of rewards and punishments may be an important motivational force contributing to parochialism.

Thus, several studies support the assumption of the extended common currency schema, in which shared representations in reward-related structures underlie the first-hand and vicarious experience of rewards and punishments. Importantly, vicarious value-related neural activity has been found to change with the perceived characteristics of the observed person, such as similarity to oneself⁵⁹, previous fair behaviour⁶⁵, friendship⁶⁰ or ingroup or outgroup membership⁶⁴. This suggests that vicarious experiences of rewards and punishments may not reflect automatic and context-invariant responses — as is sometimes argued for non-social valuation processes⁶ — but instead may flexibly incorporate social classification processes. The extent of these social-context effects on vicarious experienced values may depend on distinct patterns of functional connectivity between valuation regions and areas involved in social cognition — a possibility that is a topic for future studies.

Vicarious decision making. Decisions are rarely made in social isolation, and individuals often benefit from information about how others decided in similar situations. One study⁶⁶ suggested that the integration of vicarious and personal reward experience may occur at the level

of value-related neural processing in the ventral striatum. Here, participants made stock purchase decisions based on non-social information (previous outcomes) and social information about each stock (by passively observing other individuals' decisions and outcomes). Both types of information had distinct, additive effects on choice and ventral striatum activity. These findings of fully independent social and non-social influences on neural value coding may provide a possible neural basis for puzzling social influences on financial decisions that seem to contradict non-social evidence (for example, irrational herding behaviour during market bubbles; see also REF. 67).

Humans not only decide for themselves but can also make purely vicarious choices for others that have no direct impact on their own outcome. Three recent studies suggested that such vicarious decisions may involve neural valuation processes that are similar to those involved in choices that have consequences for oneself^{68–70}. In one of these studies, BOLD activity in the vmPFC related to the participant's estimates of how much a pictured other person would contribute to a public good⁶⁸. No vmPFC activation was observed if the participant estimated how much the depicted person would invest in a personal saving context. A second study showed that vmPFC BOLD signals correlated with the number of monetary bids for specific movie DVDs independently of whether the outcome (and cost) of the bids affected the participant or another depicted person⁶⁹. The third study showed that when participants alternated between choosing for themselves and for another person, the vmPFC flexibly encoded the decision-value signal relevant for the current choice, independently of whether it was the participant or the other person who received the resulting outcome⁷⁰ (FIG. 2c). By contrast, a region in the dmPFC always encoded the corresponding value signal for the other, currently choice-irrelevant reference frame⁷⁰.

These studies thus suggest that decision-value signals in the vmPFC can flexibly encode the value of both self-related and vicarious choices. Moreover, the same OFC region showed distinct functional connectivity with parietal regions during vicarious (compared to self-related) choices⁶⁹. These results therefore support the common currency schema, both in terms of the overlap of neural activity in value regions and the distinct functional connectivity of these regions during social choice.

Vicarious learning. Observing others may not only influence the current choice but may also help humans and animals to learn the value of stimuli and actions for future behaviour. For instance, observing that others receive painful shocks following a visual stimulus can lead to conditioned fear responses and amygdala activation to this visual stimulus, even though participants never experienced the shocks themselves⁷¹. For positive values, several further studies also suggested that vicarious learning may follow computational principles that are similar to those for learning through personal experience^{72–75}.

For example, one recent study induced prediction errors at the outcome stage of altruistic choices by unexpectedly changing the outcome, either for the

participant or for a charity benefiting from the choice. Both types of outcome changes (participant or charity) led to reward-prediction errors in the ventral striatum⁷², thus demonstrating vicarious learning of outcome probabilities for another agent. Another study showed that the dorsal striatum encodes comparable reward prediction errors when participants make a choice and benefit from the rewards or when they simply observe another person's choices and outcomes⁷³. These studies thus suggest that both experience-based and vicarious learning of anticipated values have the same neural substrate (which supports the common currency schema). However, two other studies suggest that vicarious social learning also draws on distinct prediction-error computations in different neural structures^{74,75}. These studies showed both vicarious reward prediction errors in the vmPFC (FIG. 2d) and action-prediction errors in either the dorsolateral PFC (dlPFC)⁷⁴ or the dmPFC⁷⁵ that encoded violations of expectations about which action the other participant would choose. These signals occurred in addition to BOLD signals in the ventral striatum that reflected the prediction-error signal for the participant's own expected rewards⁷⁴. Thus, in analogy to learning about specific others, vicarious learning may not only involve updating of anticipated values in reward-coding regions but also spatially and functionally distinct prediction errors that nevertheless follow similar computational principles (which supports the social-valuation-specific schema).

Neural valuation of normative social principles

Humans base their decisions not only on their self-interest but also take account of normative social principles (for example, one should be fair, one should be polite, and so on) that are often at odds with maximizing personal payout. Behaviour in line with these principles can be observed in a whole range of fully anonymous situations and cultural contexts and is thus unlikely to merely reflect properties of interactions with other specific individuals (discussed above). Exactly what motivates people to display this behaviour is a topic of debate, but it is often assumed that individual behaviour reflecting these principles may have played a crucial part in the evolution of human social behaviour⁷⁶. Several studies now propose that one proximate mechanism for maintaining this behaviour may involve neural valuation processes that overlap with those underlying non-social rewards.

Social principles affect reward experience. Neural valuation of financial outcomes can be strongly modulated by observing the simultaneous outcomes for other people in the same situation, even though these outcomes should not matter when viewed from the perspective of material self-interest. In one study, two participants simultaneously played a game in which the payoffs for a correct performance were sometimes divided unevenly between both participants (that is, one participant received more than the other for a similar performance). Here, BOLD activity in a given participant's ventral striatum co-varied with the payoff to this player when the other player did not receive any money, but

it reflected relative payoff (compared with the other's gain) when both players received money⁷⁷. Another study found that this relative reward coding did not occur if the observed outcome was not actually paid to another person⁷⁸, underlining that the modulation of ventral striatum activity was due to 'social comparison' of one's own outcome to that of others (rather than merely to the numerical discrepancy between two outcomes).

Such comparison-based modulation of neural reward processing may lie at the heart of the phenomenon of inequality aversion⁷⁹. Recent findings suggest that the behavioural enjoyment and BOLD activity in the ventral striatum and OFC that are elicited by financial payouts to one of two players strongly reflect differences in the initial capital of both players. Participants with less initial capital showed greater enjoyment ratings and BOLD activity in response to payments to themselves than to comparable payments to other players, whereas participants with a higher starting capital showed the opposite pattern⁸⁰ (FIG. 3a). These neural signals may thus reflect experienced values in the ventral striatum and OFC that track changes in the outcome difference between two individuals rather than changes in each individual's outcome. That differences in such neural value coding may lie at the heart of individual differences in inequality aversion was suggested by a study that used behavioural measures to identify participants with either dislike ('prosocials') or indifference ('individualists') to inequality. In this study, prosocials, but not individualists, rated unequal payments to two participants as less attractive than equal payments, and they showed a higher BOLD signal in the amygdala during these outcomes⁸¹ (FIG. 3b). Thus, neural activity in the brain's valuation circuitry in response to financial rewards can directly incorporate information about the outcomes of others. This supports the idea that social principles can influence the very neural representations that encode one's own financial rewards (as per the extended common currency schema). Such modulatory social influences on neural processing may underlie behavioural findings that salary satisfaction depends more on comparisons with average income in society than on absolute salary magnitude⁸².

Another example of modulatory effects on neural expected-value representations in line with social principles concerns the phenomenon of cooperation. It is widely thought that cooperative behaviour has evolved in many species through the beneficial effects it may have on reproductive fitness⁸³, but it is unclear how it may be instantiated and maintained in individuals. Several neuroimaging studies have found that cooperative behaviour leads to BOLD increases in reward-related brain areas, such as the ventral striatum and OFC^{84–87}. In one such study, two participants played a repeated prisoner's dilemma game, where the outcome revealed whether both had chosen the same strategy — and hence had cooperated — or had defected⁸⁴. Both the ventral striatum and the OFC showed higher BOLD activity during outcomes that indicated mutual cooperation. Crucially, these activations exceeded those

Proximate mechanism

Biological mechanism that enables an organism to behave in the way it does. These mechanisms develop in response to the ultimate evolutionary causes for a given behaviour.

Inequality aversion

Distaste for inequality in the distribution of outcomes. This concept — also referred to as inequality aversion — may motivate altruistic or fairness-related choices.

Prisoner's dilemma game

An experimental paradigm to measure social cooperation. Two anonymous players decide whether to cooperate or to defect. For each individual player, unilateral cooperation yields the worst material payoff, whereas unilateral defection leads to the best material payoff. However, bilateral cooperation leads to a better payoff for each of the players than bilateral defection.

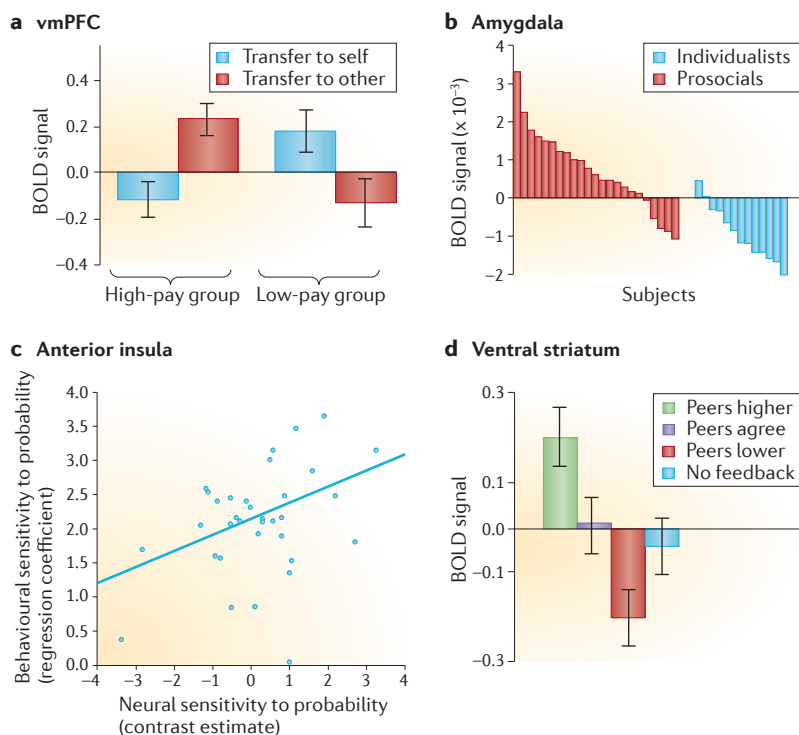


Figure 3 | Social influences on brain activity during choice behaviour and learning. **a** | Inequality modulates value-related blood-oxygen-level-dependent (BOLD) signals in the ventromedial prefrontal cortex (vmPFC) and striatum. Participants were randomly given a higher endowment (high-pay group) or lower endowment (low-pay group) than an interaction partner. Identical monetary transfers from the experimenter to both players led to different neural responses, with higher BOLD activity for transfers to the player who had less starting capital. **b** | Social value orientation is associated with BOLD activation in the amygdala in response to unequal financial payoffs to self and other. Participants with a prosocial value orientation show increased amygdala activity for large differences between payoffs to themselves or the other, whereas participants with an individualistic value orientation show the opposite. Each bar represents the strength of this relation for one participant. **c** | The risk associated with non-monetary social decisions (that is, the probability in a fictitious moral dilemma that humans will die as a consequence of the participant's choice) elicits BOLD activity in the anterior insula; this neural measure predicted how strongly participants adapted their choices to avoid this risk. **d** | BOLD signals in the ventral striatum during face attractiveness ratings change if the ratings of other people disagree, indicating influences of social conformity on ventral striatum responses. Faces originally matched in attractiveness elicit higher (or lower) ratings and ventral striatum BOLD activity after feedback that peers rate these faces as higher (or lower) than oneself. Part **a** is from REF. 80, Nature Publishing Group. Part **b** is from REF. 81, Nature Publishing Group. Part **c** is adapted with permission from REF. 98, Cell Press/Elsevier. Part **d** is adapted with permission from REF. 102, © 2011 by SAGE Publications.

during non-social control conditions (in which the identical financial outcome or a choice outcome was determined by a computer). These findings were interpreted as indicating an experienced value of cooperation per se that may bias individuals to display cooperative behaviour⁸⁴.

Social principles influence decisions. Social principles not only influence reward experience but can also bias decisions and even lead participants to accept a personal cost for implementing a particular principle. Several studies have shown that these puzzling behaviours may relate to influences on neural decision-value

computations. For instance, humans frequently sacrifice some of their own resources to punish violators of social norms, a behavioural tendency that may have evolved to stabilize cooperation and social-group cohesion. Such altruistic punishment is associated with feelings of satisfaction⁸⁸ and with neural activity in the ventral striatum and the OFC^{88,89}, even though it resulted in a financial loss for the participant. Other examples for influences of fairness-related social principles on decision making come from fMRI studies of the ultimatum game^{90–93}. In these studies, rejection of unfair outcome distributions elicited activation of the anterior insula⁹⁰, whereas acceptance of fair offers elicited BOLD activity in the ventral striatum, vmPFC and amygdala⁹¹. These neural responses were not restricted to situations in which the participants' decisions reflected costly punishments for unfair offers: BOLD activity in the vmPFC and anterior insula were also modulated by fairness if participants decided about the outcome distributions themselves^{94,95}. The (un)fairness of choices resulting in outcome distributions is therefore encoded in neural structures that overlap with those often presumed to encode anticipated values of non-social choices (which is in accordance with the extended common currency schema). However, in close similarity to the effects of vicarious reward experience (discussed above), fairness-related neural responses and the associated behaviour can strongly depend on contextual factors, such as the proposer's intentions⁹², the relation of a specific offer's distribution to the average in a series of offers⁹³ or whether the opponent belongs to the player's social group or a rival group⁹⁶. This contradicts a direct, automatic mapping of the objective 'fairness' of others' decisions to neural responses in the reward system and raises questions about the origins of these value-related responses. One possible mechanism by which social principles may influence value-coding regions is through functional connectivity with areas involved in social cognition (such as the TPJ and dmPFC⁹⁶), as has been observed for basic social reward experience³⁸ and for altruistic⁴⁵ and vicarious⁶⁹ choices.

Valuation-related neural brain structures may also be implicated in abstract social decisions that do not involve any personal gains or losses for the participant. For instance, decisions about the distribution of money between different groups of African children elicited neural activity related to both equity (how fairly is the money split) and efficiency (how much money in total is given)⁹⁷. Inequity correlated with BOLD activity in the insula, whereas efficiency correlated with BOLD activity in the striatum, suggesting separate neural coding of both of these putatively decision-relevant social principles in the brain's reward circuitry. Finally, it seems that even social decisions that do not involve money may involve similar neural processes as those involved in classic forms of value-based decision making. For example, in one study, moral judgements about courses of actions that resulted in varying hypothetical numbers of deaths activated the OFC and the ventral striatum, and activity in both brain structures correlated with

the expected number of lives lost⁹⁸. In another parallel to financial decision making⁹⁹, activity in the anterior insula encoded the risk (that is, the outcome probability) of deaths associated with each decision⁹⁸ (FIG. 3c).

Together, these studies suggest that decisions taken to implement normative social principles, often at a personal cost, mainly involve neural computations in brain regions that are also associated with value coding during non-social decision making. These findings imply that both normative social principles and non-social factors (for example, the monetary payoffs to the participant) can be integrated into a common decision-value computation underlying choice (which is in accordance with the extended common currency schema).

Value learning in line with social principles. Although numerous studies have demonstrated that normative social principles can influence neural value computations during choices, it is relatively unknown how these principles influence value-based learning processes. Some insight has been provided by studies on social conformity — the general tendency to align behaviour and attitudes with those of others, which may facilitate the creation and maintenance of social groups and information transmission within a population. These studies have shown that neural mechanisms that underlie the influences of conformity on an individual's valuation of objects are similar to the neural mechanisms involved in non-social reward-based learning. For instance, when the attractiveness ratings of faces¹⁰⁰ or music choices¹⁰¹ made by other people do not match those of the participant, then BOLD activity increases in the rostral ACC and decreases in the ventral striatum (thus reflecting higher ventral striatum activity during conformity) of the participant. These neural effects of disagreement are more pronounced in participants who subsequently change their ratings¹⁰¹ and for those face stimuli that the participant subsequently rates differently¹⁰⁰ to conform with others' opinion. This suggests that these neural responses in the rostral ACC and ventral striatum reflect conformity-related prediction-error signals associated with updating the estimated value of the stimuli. Importantly, being confronted with the discrepant opinion of others also influences the neural valuation of the same items during subsequent encounters: face stimuli that were initially rated as more (or less) attractive by a peer group than by the participant yielded higher (or lower) attractiveness ratings and ventral striatum and OFC activity when the participant rated them a second time¹⁰² (FIG. 3d).

These studies thus suggest that conformity-related changes in behaviour are brought about by similar prediction-error learning signals to those involved in non-social forms of learning. However, no studies have directly compared these two types of learning, so it is currently unclear whether these results provide support for truly domain-general value processing (that is, for the extended common currency schema) or for common principles of neural processing that are implemented in different neural structures (that is, for the social-valuation-specific schema).

Is social neural valuation distinct?

Most of the evidence discussed above supports the notion that social behaviour is controlled by reward- and motivation-related neural processes in the ventral striatum, vmPFC, amygdala and insula, which strongly resemble the neural value signals that were originally identified in the context of non-social decision making. These neural processes reflect experienced value coding of social rewards at the time of consumption, computation of the anticipated or decision value associated with a given social choice and prediction-error-like neural signals during social learning. As discussed throughout the article, all of these signals have been observed during different classes of situations that reflect either social interactions with specific others, vicarious decision making and learning or influences of abstract social principles on valuation and behaviour. As the only exception to this, there is so far little evidence that learning-related value signals are involved in the control of behaviour that conforms to social principles, possibly because such behaviour usually does not change rapidly from trial to trial in the context of behavioural experiments.

The findings discussed above are clearly at odds with the notion that all social aspects of our environment — including their motivational and hedonic value — are processed in fully separate, dedicated neural circuitries. Numerous findings indeed suggest that BOLD activity in the ventral striatum, vmPFC, amygdala and/or insula encodes a 'common neural currency' that assigns motivational properties to all stimuli and situations, irrespective of whether they relate to specific others, to abstract social principles or to material objects. This conclusion, however, may be premature, as some aspects of the findings discussed above suggest that there are also differences between social and non-social neural valuation.

For example, it is unclear whether the BOLD changes in overlapping areas for social and non-social decisions involve the same neurons. fMRI may not provide adequate spatial resolution to resolve whether different types of decisions recruit distinct sub-areas or neural populations of a given brain area¹⁰³. Single-unit recording studies in non-human primates are beginning to identify different types of neurons in the striatum, OFC and ACC that may selectively encode social versus non-social aspects of rewards. For example, different types of neurons in the striatum have been found to respond to social (images of conspecifics) or non-social (juice) rewards¹⁰⁴ or to be involved in signalling either when a reward was given or when this reward was due to the action of the monkey or a conspecific¹⁰⁵. Another study found that neurons in the macaque OFC mainly encoded rewards given to the monkey himself, whereas neurons in a subportion of the ACC encoded rewards received by another monkey¹⁰⁶. These recent studies seem to be consistent with the social-valuation-specific schema and illustrate that invasive recordings (and possibly brain stimulation methods) in non-human primates may offer important evidence for arbitrating between the two competing schemas of social reward processing, provided that all aspects of social decision making reviewed here can be identified and reliably measured in these species (BOX 4).

Ultimatum game

An experimental paradigm to measure fairness preferences. Two anonymous players decide how to split a given sum of money. Player 1 can make an offer on how to split the money between the two players. Player 2 can accept or reject the offer. In case of acceptance, the decision will be implemented, but in case of rejection both players receive nothing. By rejecting positive but unfair offers, player 2 can punish player 1 at the cost of receiving nothing.

Box 4 | Evolutionary aspects of social rewards and valuation

The rewarding properties of social behaviour may have evolved to facilitate group cohesion and cooperation and may thus also be observable in non-human primates, our evolutionary cousins. A few studies have provided evidence that social behaviour is closely tied to reward processing in non-human primates. For example, macaques are willing to trade access to food for the possibility of viewing pictures of socially relevant others¹⁴⁴, indicating that they value these social stimuli more than they value the food. This valuation of social information seems to be dependent on the integrity of the anterior cingulate gyrus¹⁴⁵. Macaques can also learn to represent an opponent's actions in reward-based competitive interactive games through distinct neural signals in the prefrontal cortex¹⁴⁶. Finally, monkeys can show vicarious reinforcement when their own actions lead to rewards for other monkeys¹⁴⁷, a behavioural tendency that is amplified by oxytocin administration¹⁴⁸ and that apparently involves distinct types of neurons in the anterior cingulate cortex¹⁰⁶.

Many other social aspects of reward-based learning and decision making have only been demonstrated behaviourally in non-human primates. For example, brown capuchin monkeys can learn foraging strategies purely vicariously, by observing conspecifics' rewarded choice¹⁴⁹. There is also considerable debate regarding the degree to which non-human primates display behaviour that is in line with normative social principles, such as altruistic food sharing^{150–152}, altruistic helping¹⁵³, inequity aversion^{154–157} and fairness-related behaviour¹⁵⁸. Whether these behaviours engage reward- and value-related neural processes as they do in humans remains an interesting question for further research. These comparative studies may require methodological developments to enable parallel studies in monkeys and humans (for example, methods to image cerebral blood flow in chimpanzees^{159,160}), but they may also provide converging evidence on social rewards and values using methods that are not readily available for human studies, such as single-unit recording^{106,146} and experimental lesions¹⁴⁵.

If different types of neurons underlie valuation of social versus non-social information, then even neural activity at the level of BOLD signals may show crucial dissociations in a given brain area. Such dissociations are less apparent in the literature, as relatively few studies have directly compared social and non-social rewards or decisions in the same participants; many conclusions about possible neural overlaps have therefore been based on reverse inference¹⁰⁷ or on comparisons between the results of different studies. When social and non-social decisions are directly compared, it can be extremely difficult to match different classes of stimuli so as to provide similar subjective values, which is a prerequisite for an objective test of whether an identical scale assigns value in both situations. These concerns notwithstanding, a growing number of studies have reported apparent functional specialization or domain-specific activity within regions of the reward circuitry for social versus non-social aspects of decision-relevant information^{19,24,43,57,58,74,108}. This may be most apparent for tasks that require value learning: here, similar prediction-error computations are reflected in activity of different areas of the neural valuation circuitry, depending on whether these computations involved information about social context or information about non-social rewards. Moreover, some pharmacological studies also suggest that there are differences in the precise neuronal underpinnings of social and non-social value coding. For example, the effects of oxytocin on social behaviour include neural modulations in value-related regions without concomitant dopamine release¹⁰⁹, whereas BOLD responses in these same regions for non-social decisions are clearly affected by pharmacological modulation of the dopamine system¹¹⁰. All of this

suggests that social and non-social decision making may rely on parallel neural computations that follow similar principles but that are nevertheless instantiated in adjacent or overlapping neural circuits that are specialized for processing one type of information (in accordance with the social-valuation-specific schema). Evidence for such distinct neural populations within a given cortical region may perhaps be obtained using methods such as high-resolution fMRI, repetition suppression paradigms¹¹¹ or multivariate pattern analyses¹¹².

Finally, even shared neural processes in some regions may nevertheless depend on input from, and interactions with, distinct brain networks that encode either social or non-social aspects of the environment. In this case, the main difference in neural processing during social and non-social decisions may not lie in local value computations but rather in the remote neural regions that may provide the information on which these computations are based. Consistent with this view, several types of social decisions reviewed above have been associated with responses in regions outside the classic valuation circuitry (such as the dlPFC¹¹³, TPJ^{11,114} and dmPFC^{70,75}) that may in principle provide input that is relevant for the construction of uniquely social values. This possibility has been formally tested with connectivity analyses^{38,45,57,58,67,69,74,113,115}, which have revealed that different types of social decision making and learning involve functional coupling between BOLD responses in 'classic' valuation regions and BOLD responses in regions outside the reward circuitry (for example, in the TPJ or dmPFC). However, the specificity of these effects in social contexts remains to be established by direct comparisons of how social versus non-social decisions may change patterns of connectivity, and how the strength of this connectivity relates to behavioural performance. These tests may also be useful for studying the mechanisms by which social value-related brain responses change with characteristics of the target of valuation^{59,60,64,65} or choice situation^{92,93}.

Outlook

We hope to have illustrated many interesting parallels that are emerging in the neuroscientific study of social and individual decision making, most of which relate to reward coding and neural value computations. These findings increase our understanding of how social aspects of our environment can be integrated with non-social information in order to control our behaviour. Most importantly, however, these parallels lead to interesting questions that may guide the neurobiological study of social decision making in coming years. In closing, we outline some of these questions.

The apparent overlap in neural processing for social and non-social decisions suggests that some phenomena discovered in both domains may have a common neural origin. For example, some modulatory effects of social comparisons on reward-related striatal BOLD activity^{77,78,80,81} show conceptual similarities to findings on adaptive coding of reward value in the non-social domain¹¹⁶, perhaps indicating that social comparisons may provide just another 'frame of reference' for neural

Reverse inference

A scientific strategy of inferring the presence of a specific mental process when observing brain activity that has been correlated in previous studies with this process. This strategy can lead to invalid conclusions when used carelessly (because activity in a brain region may often be triggered by several different mental processes) and should be applied with caution.

Repetition suppression

Phenomenon that repetition of the same stimulus elicits reduced activity in neurons specialized for this stimulus, possibly reflecting adaptation.

Multivariate pattern analyses

Neuroimaging analyses that do not examine activity changes in each voxel in isolation but rather identify patterns of activity changes across spatially distributed voxels.

Adaptive coding

Neural computations that are not fixed but adapt to the environment. For example, the same range of neural activity can encode different value ranges in different settings.

coding of reward magnitude. The degree to which many basic principles of neural reward processing really differ between decisions in the non-social domain and social decisions remains to be established. Conversely, similarities between social and non-social reward processing and valuation are often proposed based exclusively on reverse inference¹⁰⁷ — that is, on the observation of similar brain activity (and occasionally self-reports of hedonic values) for both types of stimuli. It may therefore be crucial to bolster such conclusions with evidence that the two types of rewards actually change choice behaviour and learning in a similar manner. This could elucidate the degree to which the brain really treats both social and non-social factors as rewards in the classic biological sense.

The finding that social principles (such as fairness) influence neural valuation processes in subcortical brain regions such as the ventral striatum and the amygdala suggests that some of these principles may not reflect cognitive concepts that originated late in human civilization but may instead depend on basic neural reward mechanisms that may have helped to shape the evolution of human societies (see also BOX 4). However, different cultures differ in the value they place on social principles, and neuroimaging studies may elucidate how these cultural differences are reflected in basic responses of the brain's reward circuitry¹¹⁷. Within a specific culture, the neural processes found to be influenced by these principles may be used as measures for individual differences in social value orientations⁸¹, particularly in situations

in which overt behavioural responses are problematic or prone to bias. These measures may also be practically used to answer a question of utmost importance to society: namely, whether and how social value orientations are acquired during brain development and how they can change by subsequent learning.

Finally, a major open question is the degree to which any observed neural responses are causally necessary for (rather than just correlate with) social decision making and learning. Resolving this issue will require direct experimental manipulation of these neural signals, which is currently difficult to achieve (BOX 4). Nevertheless, promising avenues include combinations of pharmacological interventions with imaging, which focus on how blocking or increasing neuromodulatory neurotransmitters or hormones may affect value-related processing regions^{110,118} during social decision making. Another putative approach rests on combinations of brain stimulation methods and imaging¹¹⁹. Reward-related brain areas are likely to interact with other areas during social decision making, offering putative indirect pathways by which subcortical regions could be targeted¹²⁰. Such multimethod approaches, in combination with mathematical models of functional connectivity¹²¹, may shift the focus from responses of single regions in the reward circuitry to functional interactions within brain networks. These patterns of network interactions may ultimately provide the main difference in how social and individual decisions are controlled by the human brain.

1. Kishida, K. T., King-Casas, B. & Montague, P. R. Neuroeconomic approaches to mental disorders. *Neuron* **67**, 543–554 (2010).
2. Damasio, A. R., Tranel, D. & Damasio, H. Individuals with sociopathic behavior caused by frontal damage fail to respond autonomically to social stimuli. *Behav. Brain Res.* **41**, 81–94 (1990).
3. Rangel, A., Camerer, C. & Montague, P. R. A framework for studying the neurobiology of value-based decision making. *Nature Rev. Neurosci.* **9**, 545–556 (2008).
4. Kable, J. W. & Glimcher, P. W. The neurobiology of decision: consensus and controversy. *Neuron* **63**, 733–745 (2009).
5. Chib, V. S., Rangel, A., Shimojo, S. & O'Doherty, J. P. Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *J. Neurosci.* **29**, 12315–12320 (2009).
6. Lebreton, M., Jorge, S., Michel, V., Thirion, B. & Pessiglione, M. An automatic valuation system in the human brain: evidence from functional neuroimaging. *Neuron* **64**, 431–439 (2009).
7. Dunbar, R. I. M. & Shultz, S. Evolution in the social brain. *Science* **317**, 1344–1347 (2007).
8. Peelen, M. V. & Downing, P. E. The neural basis of visual body perception. *Nature Rev. Neurosci.* **8**, 636–648 (2007).
9. Amodio, D. M. & Frith, C. D. Meeting of minds: the medial frontal cortex and social cognition. *Nature Rev. Neurosci.* **7**, 268–277 (2006).
10. Saxe, R. Uniquely human social cognition. *Curr. Opin. Neurobiol.* **16**, 235–239 (2006).
11. Carter, M. C., Bowling, D. L., Rieck, C. & Huettel, S. A. A distinct role of the temporo-parietal junction in predicting socially guided decisions. *Science* **337**, 109–111 (2012).
12. Adolphs, R. Conceptual challenges and directions for social neuroscience. *Neuron* **65**, 752–767 (2010).
13. Aharon, I. *et al.* Beautiful faces have variable reward value: fMRI and behavioral evidence. *Neuron* **32**, 537–551 (2001).
14. O'Doherty, J. P. *et al.* Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia* **41**, 147–155 (2003).

15. Kampe, K. K. W., Frith, C. D., Dolan, R. J. & Frith, U. Psychology: reward value of attractiveness and gaze. *Nature* **413**, 589 (2001).
16. Winston, J. S., O'Doherty, J., Kilner, J. M., Perrett, D. I. & Dolan, R. J. Brain systems for assessing facial attractiveness. *Neuropsychologia* **45**, 195–206 (2007).
17. Todorov, A., Said, C. P., Oosterhof, N. N. & Engell, A. D. Task-invariant brain responses to the social value of faces. *J. Cogn. Neurosci.* **23**, 2766–2781 (2011).
18. Prévost, C., Pessiglione, M., Météreau, E., Cléry-Melin, M.-L. & Dreher, J.-C. Separate valuation subsystems for delay and effort decision costs. *J. Neurosci.* **30**, 14080–14090 (2010).
19. Sescousse, G., Redouté, J. & Dreher, J.-C. The architecture of reward value coding in the human orbitofrontal cortex. *J. Neurosci.* **30**, 13095–13104 (2010).
20. Spreckelmeyer, K. N. *et al.* Anticipation of monetary and social reward differently activates mesolimbic brain structures in men and women. *Soc. Cogn. Affect. Neurosci.* **4**, 158–165 (2009).
21. Rademacher, L. *et al.* Dissociation of neural networks for anticipation and consumption of monetary and social rewards. *Neuroimage* **49**, 3276–3285 (2010).
22. Lin, A., Adolphs, R. & Rangel, A. Social and monetary reward learning engage overlapping neural substrates. *Soc. Cogn. Affect. Neurosci.* **7**, 274–281 (2012).
23. Kim, H., Adolphs, R., O'Doherty, J. P. & Shimojo, S. Temporal isolation of neural processes underlying face preference decisions. *Proc. Natl Acad. Sci. USA* **104**, 18253–18258 (2007).
24. Smith, D. V. *et al.* Distinct value signals in anterior and posterior ventromedial prefrontal cortex. *J. Neurosci.* **30**, 2490–2495 (2010).
25. Hayden, B. Y., Piarik, P. C., Deaner, R. O. & Platt, M. L. Economic principles motivating social attention in humans. *Proc. Biol. Sci.* **274**, 1751–1756 (2007).
26. Izuma, K., Saito, D. N. & Sadato, N. Processing of social and monetary rewards in the human striatum. *Neuron* **58**, 284–294 (2008).

This study shows that monetary gains and observed social gestures elicit largely overlapping patterns of BOLD activity in the striatum, therefore suggesting that both types of outcomes trigger related value computations.

27. Davey, C. G., Allen, N. B., Harrison, B. J., Dwyer, D. B. & Yücel, M. Being liked activates primary reward and midline self-related brain regions. *Hum. Brain Mapp.* **31**, 660–668 (2010).
28. Gunther Moor, B., van Leijenhorst, L., Rombouts, S. A. R. B., Crone, E. A. & Van der Molen, M. W. Do you like me? Neural correlates of social evaluation and developmental trajectories. *Soc. Neurosci.* **5**, 461–482 (2010).
29. Powers, K. E., Somerville, L. H., Kelley, W. M. & Heatherton, T. F. Rejection sensitivity polarizes striatal-medial prefrontal activity when anticipating social feedback. *J. Cogn. Neurosci.* **25**, 1887–1895 (2013).
30. Meshi, D., Morawetz, C. & Heekeren, H. R. Nucleus accumbens response to gains in reputation for the self relative to gains for others predicts social media use. *Front. Hum. Neurosci.* **7**, 439 (2013).
31. Morelli, S. A., Torre, J. B. & Eisenberger, N. I. The neural bases of feeling understood and not understood. *Soc. Cogn. Affect. Neurosci.* <http://dx.doi.org/10.1093/scan/nst191> (2014).
32. Cooper, J. C., Dunne, S., Furey, T. & O'Doherty, J. P. The role of the posterior temporal and medial prefrontal cortices in mediating learning from romantic interest and rejection. *Cereb. Cortex* <http://dx.doi.org/10.1093/cercor/bht102> (2013).
33. Zink, C. F. *et al.* Know your place: neural processing of social hierarchy in humans. *Neuron* **58**, 273–283 (2008).
34. Eisenberger, N. I., Lieberman, M. D. & Williams, K. D. Does rejection hurt? An fMRI study of social exclusion. *Science* **302**, 290–292 (2003).
35. Slavich, G. M., Way, B. M., Eisenberger, N. I. & Taylor, S. E. Neural sensitivity to social rejection is associated with inflammatory responses to social stress. *Proc. Natl Acad. Sci. USA* **107**, 14817–14822 (2010).
36. Dewall, C. N. *et al.* Acetaminophen reduces social pain: behavioral and neural evidence. *Psychol. Sci.* **21**, 931–937 (2010).

37. Kross, E., Berman, M. G., Mischel, W., Smith, E. E. & Wager, T. D. Social rejection shares somatosensory representations with physical pain. *Proc. Natl Acad. Sci. USA* **108**, 6270–6275 (2011).
38. Smith, D. V., Clithero, J. A., Boltuck, S. & Huettel, S. A. Functional connectivity with ventromedial prefrontal cortex reflects subjective value for social rewards. *Soc. Cogn. Affect. Neurosci.* <http://dx.doi.org/10.1093/scan/nsu005> (2014).
39. Izuma, K., Saito, D. N. & Sadato, N. Processing of the incentive for social approval in the ventral striatum during charitable donation. *J. Cogn. Neurosci.* **22**, 621–631 (2010).
40. Kohls, G. *et al.* The nucleus accumbens is involved in both the pursuit of social reward and the avoidance of social punishment. *Neuropsychologia* **51**, 2062–2069 (2013).
41. Chein, J., Albert, D., O'Brien, L., Uckert, K. & Steinberg, L. Peers increase adolescent risk taking by enhancing activity in the brain's reward circuitry. *Dev. Sci.* **14**, F1–F10 (2011).
42. Moll, J. *et al.* Human fronto-mesolimbic networks guide decisions about charitable donation. *Proc. Natl Acad. Sci. USA* **103**, 15623–15628 (2006).
43. Harbaugh, W. T., Mayr, U. & Burghart, D. R. Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science* **316**, 1622–1625 (2007).
44. Izuma, K., Saito, D. N. & Sadato, N. The roles of the medial prefrontal cortex and striatum in reputation processing. *Soc. Neurosci.* **5**, 133–147 (2010).
45. Hare, T. A., Camerer, C. F., Knoepfle, D. T. & Rangel, A. Value computations in ventral medial prefrontal cortex during charitable decision making incorporate input from regions involved in social cognition. *J. Neurosci.* **30**, 583–590 (2010).
This study shows that charitable donations involve increased functional connectivity between the vmPFC and TPJ, suggesting that the TPJ provides information for the vmPFC value computations.
46. Fareri, D. S., Niznikiewicz, M. A., Lee, V. K. & Delgado, M. R. Social network modulation of reward-related signals. *J. Neurosci.* **32**, 9045–9052 (2012).
47. Delgado, M. R., Frank, R. H. & Phelps, E. A. Perceptions of moral character modulate the neural systems of reward during the trust game. *Nature Neurosci.* **8**, 1611–1618 (2005).
48. King-Casas, B. *et al.* Getting to know you: reputation and trust in a two-person economic exchange. *Science* **308**, 78–83 (2005).
This study shows that repeated social interactions involve prediction-error computations in the striatum that resemble similar computations during non-social choices.
49. Baumgartner, T., Heinrichs, M., Vonlanthen, A., Fischbacher, U. & Fehr, E. Oxytocin shapes the neural circuitry of trust and trust adaptation in humans. *Neuron* **58**, 639–650 (2008).
This study demonstrates that oxytocin has specific influences on the behavioural and neural sensitivity to violations of trust during social interactions.
50. Phan, K. L., Sripada, C. S., Angstadt, M. & McCabe, K. Reputation for reciprocity engages the brain reward center. *Proc. Natl Acad. Sci. USA* **107**, 13099–13104 (2010).
51. Schultz, W. Multiple dopamine functions at different time courses. *Annu. Rev. Neurosci.* **30**, 259–288 (2007).
52. Smith-Collins, A. P. R. *et al.* Specific neural correlates of successful learning and adaptation during social exchanges. *Soc. Cogn. Affect. Neurosci.* **8**, 887–896 (2013).
53. Baron, S. G., Gobbini, M. I., Engell, A. D. & Todorov, A. Amygdala and dorsomedial prefrontal cortex responses to appearance-based and behavior-based person impressions. *Soc. Cogn. Affect. Neurosci.* **6**, 572–581 (2011).
54. Fouragnan, E. *et al.* Reputational priors magnify striatal responses to violations of trust. *J. Neurosci.* **33**, 3602–3611 (2013).
55. Jones, R. M. *et al.* Behavioral and neural properties of social reinforcement learning. *J. Neurosci.* **31**, 13039–13045 (2011).
This study demonstrates, with various behavioural and neural measures, that social gestures have similar reinforcing properties to those of non-social rewards.
56. Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U. & Fehr, E. Oxytocin increases trust in humans. *Nature* **435**, 673–676 (2005).
57. Behrens, T. E. J., Hunt, L. T., Woolrich, M. W. & Rushworth, M. F. S. Associative learning of social value. *Nature* **456**, 245–249 (2008).
This study shows that social and non-social information relevant for value-based choices is encoded by similar types of neural computations but that these computations are implemented in different neural structures.
58. Hampton, A. N., Bossaerts, P. & O'Doherty, J. P. Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proc. Natl Acad. Sci. USA* **105**, 6741–6746 (2008).
59. Mobbs, D. *et al.* Key role for similarity in vicarious reward. *Science* **324**, 900 (2009).
This study shows that neural responses in value-related areas can be triggered vicariously when observing a positive outcome for another person and that these responses are modulated by the perceived similarity between the observer and the observee.
60. Braams, B. R. *et al.* Reward-related neural responses are dependent on the beneficiary. *Soc. Cogn. Affect. Neurosci.* <http://dx.doi.org/10.1093/scan/nst077> (2013).
61. Singer, T. *et al.* Empathy for pain involves the affective but not sensory components of pain. *Science* **303**, 1157–1162 (2004).
62. Wicker, B. *et al.* Both of us disgusted in *my* insula: the common neural basis of seeing and feeling disgust. *Neuron* **40**, 655–664 (2003).
63. Krach, S. *et al.* Your flaws are my pain: linking empathy to vicarious embarrassment. *PLoS ONE* **6**, e18675 (2011).
64. Hein, G., Silani, G., Preuschhoff, K., Batson, C. D. & Singer, T. Neural responses to ingroup and outgroup members' suffering predict individual differences in costly helping. *Neuron* **68**, 149–160 (2010).
65. Singer, T. *et al.* Empathic neural responses are modulated by the perceived fairness of others. *Nature* **439**, 466–469 (2006).
66. 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).
67. De Martino, B., O'Doherty, J. P., Ray, D., Bossaerts, P. & Camerer, C. In the mind of the market: theory of mind biases value computation during financial bubbles. *Neuron* **79**, 1222–1231 (2013).
68. Cooper, J. C., Kreps, T. A., Wiebe, T., Pirkil, T. & Knutson, B. When giving is good: ventromedial prefrontal cortex activation for others' intentions. *Neuron* **67**, 511–521 (2010).
69. Janowski, V., Camerer, C. & Rangel, A. Empathic choice involves vmPFC value signals that are modulated by social processing implemented in IPL. *Soc. Cogn. Affect. Neurosci.* **8**, 201–208 (2012).
70. Nicolle, A. *et al.* An agent independent axis for executed and modeled choice in medial prefrontal cortex. *Neuron* **75**, 1114–1121 (2012).
This study shows that neural value signals in the vmPFC (versus those in the dmPFC) do not generically encode outcomes for oneself (versus another person) but rather represent values for choices that the agent executes him- or herself (versus only simulates cognitively).
71. Olsson, A. & Phelps, E. A. Social learning of fear. *Nature Neurosci.* **10**, 1095–1102 (2007).
72. Kuss, K. *et al.* A reward prediction error for charitable donations reveals outcome orientation of donors. *Soc. Cogn. Affect. Neurosci.* **8**, 216–223 (2011).
73. Cooper, J. C., Dunne, S., Furey, T. & O'Doherty, J. P. Human dorsal striatum encodes prediction errors during observational learning of instrumental actions. *J. Cogn. Neurosci.* **24**, 106–118 (2012).
74. Burke, C. J., Tobler, P. N., Baddeley, M. & Schultz, W. Neural mechanisms of observational learning. *Proc. Natl Acad. Sci. USA* **107**, 14431–14436 (2010).
This study shows that learning by observing another person involves two distinct types of signals in the dlPFC and vmPFC that relate to predictions errors about the other person's actions and outcomes, respectively.
75. Suzuki, S. *et al.* Learning to simulate others' decisions. *Neuron* **74**, 1125–1137 (2012).
76. Fehr, E. & Fischbacher, U. The nature of human altruism. *Nature* **425**, 785–791 (2003).
77. Fliessbach, K. *et al.* Social comparison affects reward-related brain activity in the human ventral striatum. *Science* **318**, 1305–1308 (2007).
78. 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. USA* **108**, 16044–16049 (2011).
79. Fehr, E. & Schmidt, K. M. A theory of fairness, competition, and cooperation. *Q. J. Econ.* **114**, 817–868 (1999).
80. Tricomi, E., Rangel, A., Camerer, C. F. & O'Doherty, J. P. Neural evidence for inequality-averse social preferences. *Nature* **463**, 1089–1091 (2010).
This study demonstrates that social preferences for equitable outcomes are evident in neural value signals in the striatum and vmPFC.
81. Haruno, M. & Frith, C. D. Activity in the amygdala elicited by unfair divisions predicts social value orientation. *Nature Neurosci.* **13**, 160–161 (2010).
82. Boyce, C. J., Brown, G. D. & Moore, S. C. Money and happiness: rank of income, not income, affects life satisfaction. *Psychol. Sci.* **21**, 471–475 (2010).
83. Pennisi, E. On the origin of cooperation. *Science* **325**, 1196–1199 (2009).
84. Rilling, J. K. *et al.* Neural basis for social cooperation. *Neuron* **35**, 395–405 (2002).
85. Decety, J., Jackson, P. L., Sommerville, J. A., Chaminade, T. & Meltzoff, A. N. The neural bases of cooperation and competition: an fMRI investigation. *Neuroimage* **23**, 744–751 (2004).
86. Yoshida, W., Seymour, B., Friston, K. J. & Dolan, R. J. Neural mechanisms of belief inference during cooperative games. *J. Neurosci.* **30**, 10744–10751 (2010).
87. Krill, A. L. & Platek, S. M. Working together may be better: activation of reward centers during a cooperative maze task. *PLoS ONE* **7**, e30613 (2012).
88. de Quervain, D. J.-F. *et al.* The neural basis of altruistic punishment. *Science* **305**, 1254–1258 (2004).
89. White, S. F., Brislin, S. J., Sinclair, S. & Blair, J. R. Punishing unfairness: rewarding or the organization of a reactively aggressive response? *Hum. Brain Mapp.* **35**, 2137–2147 (2013).
90. Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E. & Cohen, J. D. The neural basis of economic decision-making in the ultimatum game. *Science* **300**, 1755–1758 (2003).
91. Tabibnia, G., Satpute, A. B. & Lieberman, M. D. The sunny side of fairness: preference for fairness activates reward circuitry (and disregarding unfairness activates self-control circuitry). *Psychol. Sci.* **19**, 339–347 (2008).
92. Guroglu, B., Bos, W. V. D., Rombouts, S. A. R. B. & Crone, E. A. Unfair? It depends: neural correlates of fairness in social context. *Soc. Cogn. Affect. Neurosci.* **5**, 414–423 (2010).
93. Wright, N. D., Symmonds, M., Fleming, S. M. & Dolan, R. J. Neural segregation of objective and contextual aspects of fairness. *J. Neurosci.* **31**, 5244–5252 (2011).
94. Zaki, J. & Mitchell, J. P. Equitable decision making is associated with neural markers of intrinsic value. *Proc. Natl Acad. Sci. USA* **108**, 19761–19766 (2011).
95. Dawes, C. T. *et al.* Neural basis of egalitarian behavior. *Proc. Natl Acad. Sci. USA* **109**, 6479–6483 (2012).
96. Baumgartner, T., Götze, L., Gügler, R. & Fehr, E. The mentalizing network orchestrates the impact of parochial altruism on social norm enforcement. *Hum. Brain Mapp.* **33**, 1452–1469 (2012).
97. Hsu, M., Anen, C. & Quartz, S. R. The right and the good: distributive justice and neural encoding of equity and efficiency. *Science* **320**, 1092–1095 (2008).
98. Shenhav, A. & Greene, J. D. Moral judgments recruit domain-general valuation mechanisms to integrate representations of probability and magnitude. *Neuron* **67**, 667–677 (2010).
99. Preuschhoff, K., Quartz, S. R. & Bossaerts, P. Human insula activation reflects risk prediction errors as well as risk. *J. Neurosci.* **28**, 2745–2752 (2008).
100. Klucharev, V., Hytönen, K., Rijpkema, M., Smidts, A. & Fernández, G. Reinforcement learning signal predicts social conformity. *Neuron* **61**, 140–151 (2009).
101. Campbell-Meiklejohn, D. K., Bach, D. R., Roepstorff, A., Dolan, R. J. & Frith, C. D. How the opinion of others affects our valuation of objects. *Curr. Biol.* **20**, 1165–1170 (2010).
102. Zaki, J., Schirmer, J. & Mitchell, J. P. Social influence modulates the neural computation of value. *Psychol. Sci.* **22**, 894–900 (2011).
103. Logothetis, N. K. What we can do and what we cannot do with fMRI. *Nature* **453**, 869–878 (2008).
104. Klein, J. T. & Platt, M. L. Social information signaling by neurons in primate striatum. *Curr. Biol.* **23**, 691–696 (2013).
This study shows that social versus non-social rewards are processed by distinct types of neurons in the striatum of the macaque.

105. Báez-Mendoza, R., Harris, C. J. & Schultz, W. Activity of striatal neurons reflects social action and own reward. *Proc. Natl Acad. Sci. USA* **110**, 16634–16639 (2013).
106. Chang, S. W. C., Gariépy, J.-F. & Platt, M. L. Neuronal reference frames for social decisions in primate frontal cortex. *Nature Neurosci.* **16**, 243–250 (2013).
107. Poldrack, R. A. Can cognitive processes be inferred from neuroimaging data? *Trends Cogn. Sci.* **10**, 59–63 (2006).
108. Spitzer, M., Fischbacher, U., Herrnberger, B., Grön, G. & Fehr, E. The neural signature of social norm compliance. *Neuron* **56**, 185–196 (2007).
109. Striepens, N. *et al.* Oxytocin enhances attractiveness of unfamiliar female faces independent of the dopamine reward system. *Psychoneuroendocrinology* **39**, 74–87 (2014).
110. 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).
111. Grill-Spector, K., Henson, R. & Martin, A. Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn. Sci.* **10**, 14–23 (2006).
112. Norman, K. A., Polyn, S. M., Detre, G. J. & Haxby, J. V. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn. Sci.* **10**, 424–430 (2006).
113. Baumgartner, T., Knoch, D., Hotz, P., Eisenegger, C. & Fehr, E. Dorsolateral and ventromedial prefrontal cortex orchestrate normative choice. *Nature Neurosci.* **14**, 1468–1474 (2011).
114. Morishima, Y., Schunk, D., Bruhin, A., Ruff, C. C. & Fehr, E. Linking brain structure and activation in temporoparietal junction to explain the neurobiology of human altruism. *Neuron* **75**, 73–79 (2012). **This study shows that the individual tendency for altruistic giving is systematically related to both the structure and the functional response profile of the TPJ.**
115. van den Bos, W., Talwar, A. & McClure, S. M. Neural correlates of reinforcement learning and social preferences in competitive bidding. *J. Neurosci.* **33**, 2137–2146 (2013).
116. Seymour, B. & McClure, S. M. Anchors, scales and the relative coding of value in the brain. *Curr. Opin. Neurobiol.* **18**, 173–178 (2008).
117. Han, S. & Northoff, G. Culture-sensitive neural substrates of human cognition: a transcultural neuroimaging approach. *Nature Rev. Neurosci.* **9**, 646–654 (2008).
118. Pleger, B. *et al.* Influence of dopaminergically mediated reward on somatosensory decision-making. *PLoS Biol.* **7**, e1000164 (2009).
119. Driver, J., Blankenburg, F., Bestmann, S., Vanduffel, W. & Ruff, C. C. Concurrent brain-stimulation and neuroimaging for studies of cognition. *Trends Cogn. Sci.* **13**, 319–327 (2009).
120. Cho, S. S. & Strafella, A. P. rTMS of the left dorsolateral prefrontal cortex modulates dopamine release in the ipsilateral anterior cingulate cortex and orbitofrontal cortex. *PLoS ONE* **4**, e6725 (2009).
121. Friston, K. J. Functional and effective connectivity: a review. *Brain Connect.* **1**, 13–36 (2011).
122. Berridge, K. C. & Kringelbach, M. L. Affective neuroscience of pleasure: reward in humans and animals. *Psychopharmacology* **199**, 457–480 (2008).
123. Padoa-Schioppa, C. & Assad, J. A. Neurons in the orbitofrontal cortex encode economic value. *Nature* **441**, 223–226 (2006).
124. Morrison, S. E. & Salzman, C. D. Re-valuing the amygdala. *Curr. Opin. Neurobiol.* **20**, 221–230 (2010).
125. Rushworth, M. F. S. & Behrens, T. E. J. Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature Neurosci.* **11**, 389–397 (2008).
126. Leknes, S. & Tracey, I. A common neurobiology for pain and pleasure. *Nature Rev. Neurosci.* **9**, 314–320 (2008).
127. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
128. D'Ardenne, K., McClure, S. M., Nystrom, L. E. & Cohen, J. D. BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science* **319**, 1264–1267 (2008).
129. O'Doherty, J. P. Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr. Opin. Neurobiol.* **14**, 769–776 (2004).
130. Hare, T. A., O'Doherty, J., Camerer, C. F., Schultz, W. & Rangel, A. Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J. Neurosci.* **28**, 5623–5630 (2008).
131. Wallis, J. D. & Kennerley, S. W. Heterogeneous reward signals in prefrontal cortex. *Curr. Opin. Neurobiol.* **20**, 191–198 (2010).
132. Sripada, C., Angstadt, M., Liberzon, I., McCabe, K. & Phan, K. L. Aberrant reward center response to partner reputation during a social exchange game in generalized social phobia. *Depress. Anxiety* **30**, 353–361 (2013).
133. Chiu, P. H. *et al.* Self responses along cingulate cortex reveal quantitative neural phenotype for high-functioning autism. *Neuron* **57**, 463–473 (2008).
134. King-Casas, B. *et al.* The rupture and repair of cooperation in borderline personality disorder. *Science* **321**, 806–810 (2008).
135. Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E. & Cohen, J. D. Opposing BOLD responses to reciprocated and unreciprocated altruism in putative reward pathways. *Neuroreport* **15**, 2539–2543 (2004).
136. Izuma, K., Matsumoto, K., Camerer, C. F. & Adolphs, R. Insensitivity to social reputation in autism. *Proc. Natl Acad. Sci. USA* **108**, 17302–17307 (2011).
137. Stone, V. E., Cosmides, L., Tooby, J., Kroll, N. & Knight, R. T. Selective impairment of reasoning about social exchange in a patient with bilateral limbic system damage. *Proc. Natl Acad. Sci. USA* **99**, 11531–11536 (2002).
138. Krajbich, I., Adolphs, R., Tranel, D., Denburg, N. L. & Camerer, C. F. Economic games quantify diminished sense of guilt in patients with damage to the prefrontal cortex. *J. Neurosci.* **29**, 2188–2192 (2009).
139. Koenigs, M. & Tranel, D. Irrational economic decision-making after ventromedial prefrontal damage: evidence from the ultimatum game. *J. Neurosci.* **27**, 951–956 (2007).
140. Moretti, L., Dragone, D. & Pellegrino, G. D. Reward and social valuation deficits following ventromedial prefrontal damage. *J. Cogn. Neurosci.* **21**, 128–140 (2009).
141. Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V. & Fehr, E. Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science* **314**, 829–832 (2006).
142. Knoch, D., Schneider, F., Schunk, D., Hohmann, M. & Fehr, E. Disrupting the prefrontal cortex diminishes the human ability to build a good reputation. *Proc. Natl Acad. Sci. USA* **106**, 20895–20899 (2009).
143. Ruff, C. C., Ugazio, G. & Fehr, E. Changing social norm compliance with noninvasive brain stimulation. *Science* **342**, 482–484 (2013).
144. Deaner, R. O., Khera, A. V. & Platt, M. L. Monkeys pay per view: adaptive valuation of social images by rhesus macaques. *Curr. Biol.* **15**, 543–548 (2005).
145. 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).
146. Barraclough, D. J., Conroy, M. L. & Lee, D. Prefrontal cortex and decision making in a mixed-strategy game. *Nature Neurosci.* **7**, 404–410 (2004).
147. Chang, S. W., Winecoff, A. A. & Platt, M. L. Vicarious reinforcement in rhesus macaques (*Macaca Mulatta*). *Front. Neurosci.* **5**, 27 (2011).
148. Chang, S. W., Barter, J. W., Ebitz, R. B., Watson, K. K. & Platt, M. L. Inhaled oxytocin amplifies both vicarious reinforcement and self reinforcement in rhesus macaques (*Macaca mulatta*). *Proc. Natl Acad. Sci. USA* **109**, 959–964 (2012).
149. Bonnie, K. E. & de Waal, F. B. M. Copying without rewards: socially influenced foraging decisions among brown capuchin monkeys. *Animal Cogn.* **10**, 283–292 (2007).
150. de Waal, F. B. M., Leimgruber, K. & Greenberg, A. R. Giving is self-rewarding for monkeys. *Proc. Natl Acad. Sci. USA* **105**, 13685–13689 (2008).
151. Hare, B. & Kwetuenda, S. Bonobos voluntarily share their own food with others. *Curr. Biol.* **20**, R230–R231 (2010).
152. Burkart, J. M., Fehr, E., Efferson, C. & Schaik, C. P. V. Other-regarding preferences in a non-human primate: common marmosets provision food altruistically. *Proc. Natl Acad. Sci. USA* **104**, 19762–19766 (2007).
153. Warneken, F. & Tomasello, M. Altruistic helping in human infants and young chimpanzees. *Science* **311**, 1301–1303 (2006).
154. Wolkentin, M. V., Brosnan, S. F. & de Waal, F. B. M. Inequity responses of monkeys modified by effort. *Proc. Natl Acad. Sci. USA* **104**, 18854–18859 (2007).
155. Brosnan, S. F. & de Waal, F. B. M. Monkeys reject unequal pay. *Nature* **425**, 297–299 (2003).
156. Silk, J. B. *et al.* Chimpanzees are indifferent to the welfare of unrelated group members. *Nature* **437**, 1357–1359 (2005).
157. Bräuer, J., Call, J. & Tomasello, M. Are apes really inequity averse? *Proc. Biol. Sci.* **273**, 3123–3128 (2006).
158. Jensen, K., Call, J. & Tomasello, M. Chimpanzees are rational maximizers in an ultimatum game. *Science* **318**, 107–109 (2007).
159. Parr, L. A., Hecht, E., Barks, S. K., Preuss, T. M. & Votaw, J. R. Face processing in the chimpanzee brain. *Curr. Biol.* **19**, 50–53 (2009).
160. Rilling, J. K. *et al.* A comparison of resting-state brain activity in humans and chimpanzees. *Proc. Natl Acad. Sci. USA* **104**, 17146–17151 (2007).

Acknowledgements

Preparation of this article was made possible by funding from the Swiss National Science Foundation (SNSF) to C.C.R. and E.F. E.F. also acknowledges support from the European Research Council Grant on the “Foundations of Economic Preferences”. The authors thank P. Tobler, S. Huettel and two anonymous referees for helpful comments.

Competing interests statement

The authors declare no competing interests.