nature neuroscience

REVIEW

Choice, uncertainty and value in prefrontal and cingulate cortex

Matthew F S Rushworth & Timothy E J Behrens

Reinforcement learning models that focus on the striatum and dopamine can predict the choices of animals and people. Representations of reward expectation and of reward prediction errors that are pertinent to decision making, however, are not confined to these regions but are also found in prefrontal and cingulate cortex. Moreover, decisions are not guided solely by the magnitude of the reward that is expected. Uncertainty in the estimate of the reward expectation, the value of information that might be gained by taking a course of action and the cost of an action all influence the manner in which decisions are made through prefrontal and cingulate cortex.

Reward-maximizing behavior has been assumed to rely predominantly on the ability to estimate the value of different stimuli in the environment and of different actions an animal might take. Reinforcement learning models have begun to provide accounts of reward-guided behavior and suggest that neural activity in the striatum and the dopamine system represents the parameters described by reinforcement learning models. Statistics, economics and machine learning, however, suggest that decision making can be substantially improved by considering a richer representation of the reward environment. Here we review and contrast some of the components of such a representation, their implications for behavior and evidence supporting their dependence on prefrontal and cingulate cortex.

Reinforcement-guided decision making

Reinforcement learning theory proposes strategies for choosing the best course of action when the only guiding cues are previous experiences of reinforcement. One critical insight deriving from these ideas is that such behavior can be divided into two components. The first addresses the problem of learning (or tracking) relationships among stimuli, actions and reinforcements. The most studied strategy for understanding such relationships proposes that stimuli or actions induce predictions or expectations of the amount or probability of reward that might follow them. Such predictions are often called 'reward expectations'. In the event of a discrepancy between a reward expectation elicited by a stimulus or action and the subsequent outcome, the reward expectation should be modified to ensure that it is more accurate the next time the stimulus or action occurs[1,2]. The discrepancy between the reward expectation and the outcome is represented quantitatively as the 'prediction error' (δ), and the extent to which expectations are updated is taken to be the product of the prediction error and the 'learning rate' (α). The new revised 'expectation' of reward on the next learning trial, V_{t+1}, is thus a function of the expectation on the current trial, V_t, and the product of the prediction error and learning rate:

$$V_{t+1} = V_t + \delta\alpha \quad (1)$$

The learning rate represents the expected value of information available on the current trial, and depends on the animal's current level of understanding of the environment. In situations where the animal is uncertain (**Fig. 1**) about the environment, new information is more valuable[3–5]. Hence, new outcomes have a large impact on future expectations either because they are surprising (inducing a large prediction error) or because of uncertainty about current expectations (inducing a large learning rate).

The dopamine system has provided the richest source of data for neuroscientists seeking experimental validation of such theoretical ideas. Single-neuron recordings from dopamine neurons across different species show several features of their activity that mimic theoretical learning parameters. When an animal is presented with a conditioned stimulus that predicts a later reward, phasic responses are induced in dopamine neurons that encode the expectation at the time of the cue and encode a quantitative prediction error when the outcome is observed[6,7]. Such signals are encoded quantitatively, whether the magnitude[8,9] or the probability[10] of the outcome is manipulated. When magnitude and probability are manipulated together, dopamine neurons respond to their combination, consistent with a representation of the expected value or utility of the outcome[8]. The striatum is one of the main projection regions of the midbrain dopaminergic nuclei, and the activity of its neurons encodes the expected value of potential actions and the difference in the expected values of potential actions[11]. Several functional magnetic resonance imaging (fMRI) studies similarly show activation in the vicinity of the midbrain dopaminergic nuclei and in the striatum that is correlated with reward expectation[12–14] and with prediction error[15–17].

Even if action-outcome associations are well learned, reinforcement learning theory considers a second subcomponent of behavior: how to select actions based on current expectations[2]. One might assume that this is easy: that the animal should simply use the action with the highest expected value. Such a strategy maximizes the animal's

Department of Experimental Psychology and Centre for Functional MRI of the Brain, John Radcliffe Hospital, University of Oxford, South Parks Road, Oxford OX1 3UD, UK. Correspondence should be addressed to M.F.S.R. (matthew.rushworth@psy.ox.ac.uk) or T.E.J.B. (behrens@fmrib.ox.lac.uk).

Published online 26 March 2008; doi:10.1038/nn2066

Figure 1 Probability, risk and uncertainty. Three related but very different concepts have come to the fore in recent descriptions of decisions and learning. Here we refer to them as probability, risk and uncertainty. The probability of an outcome derives from the underlying stochasticity of the system the animal is trying to predict. The animal cannot influence this probability, but can maximize its income by acquiring an accurate estimate of this probability and using this estimate in the computation of the value of the action. The risk of an outcome is defined as the outcome variance (for example, ref. 90) and measures the unpredictability of the outcome. Similarly, it derives from underlying stochasticity and cannot be influenced by the animal. Unlike probability, risk has its maximum at a reward rate of 0.5, when the outcome is least predictable. In economic models, risk is often assumed to be a cost that should be weighed up against expected value in making a decision. (**a**) (from ref. 90) Outcome probability (black) and risk (red) as a function of the true underlying reward rate. Uncertainty is fundamentally unlike probability and risk in that it is not a measure of the stochasticity of the underlying system, but rather a measure of the quality of the animal's own estimate. The animal can resolve uncertainty in its estimate by acquiring more data; that is, by sampling the action on more occasions. (**b**,**c**) Illustration. An animal is asked to estimate the reward rate of an action by repetitive sampling of the action, The true underlying reward rate is 0.6. In each figure, the gray shaded areas show probability density functions that represent the most the animal can know about the true reward rate after witnessing different outcomes. By chance, after 10 trials (**b**), the animal has experienced only 4 rewards. The gray distribution therefore has a mean of 0.4 but it also has a large variance; it is this variance that defined the animal's uncertainty. It is still possible that the underlying reward rate is 0.6 or even 0.8 even though the first 10 trials have resulted in only 4 rewards. After 50 trials (**c**), the animal has received 30 rewards. The mean of the distribution is now 0.6 but, crucially, the variance is much reduced. The animal is much more confident in its estimate of the reward rate. Such uncertainty is important in both learning and decision making. In learning, animals should give more weight to new data, and therefore learn faster, when they are uncertain in their current estimate. In decision making, animals should be prepared to pay to acquire new information that will reduce this uncertainty, and therefore allow them to make more accurate future predictions.

expected utility in the short term but often forces the animal to make only few actions repeatedly. By forgoing the opportunity to explore the other available options, the animal risks missing a high-value action in the future. Although there is no easy solution to this problem, it is clear that exploratory behavior should be modulated by at least two factors[18,19]: the expected immediate cost and return on that action and the expected degree to which the information obtained from making the action might influence future predictions. This second factor is the value of the information available from each action, which was also crucial for dictating the learning rate. In situations when animals will learn more from each new outcome, they should be prepared to forego more immediate rewards to obtain new information.

Prediction errors in the cingulate cortex

Reward expectation and prediction error signals may be especially prominent aspects of dopamine neuron activity, but similar information is found in prefrontal and anterior cingulate cortex (ACC), even if it is not always discussed using terms derived from computational theory. The clearest examples come from recordings in ACC, a prominent target for dopamine projections[20]. Similarly to dopamine neurons, ACC neurons encode a reward prediction that combines information about reward magnitude and reward probability[21,22]. Furthermore ACC neurons also encode a quantitative reward prediction error at the time of the outcome[23–25] (**Fig. 2**).

One might argue that the activity recorded from a small sample of neurons does not represent the function of the area as a whole, but the involvement of the ACC in the monitoring of errors and other outcomes is confirmed both by lesion studies and by techniques such as fMRI that reflect the activity of populations of neurons. Functional MRI activity in the ACC increases when experimental subjects make errors[26–28], but it also increases when positive feedback is provided, if the feedback indicates that the estimate of an action's value should be increased[29]. A negative-going event-related potential, the error- or feedback-related negativity (ERN or FRN), which has a dipole source in the ACC, can be recorded over the frontal midline scalp when errors are made[30]. Like the ACC blood oxygenation level–dependent (BOLD) signal, the ERN may reflect not just errors but also prediction errors[31] that are used to guide subsequent action selection. The ERN increases in response to positive feedback when negative feedback is expected[32]. In a probabilistic gambling task, the ERN is larger on error trials if those error trials are followed by a switch in action selection[19]. Such switches are more likely to occur when the estimates of the actions' values have been revised downward to a greater extent.

There are, however, key differences in coding between ACC and dopaminergic cells. In the dopamine system, the same cells encode a positive and negative prediction error by a phasic increase in firing and an absence of tonic firing, respectively[6,33,34]. By contrast, in the ACC, different populations of cells encode positive and negative prediction errors, and both types of error result in an increase in firing in their respective populations[25]. Although the size of positive prediction errors is encoded by linear increases in dopamine neuron firing rate, negative prediction errors are all encoded by the same response—an absence of activity[33]. This may reflect limitations in the dynamic range of dopamine neurons.

Furthermore, reward expectation and prediction error–related activity in the ACC may be particularly closely tied to action selection. Reward expectation and prediction error–related dopaminergic activity are reported both during pavlovian learning[35,36], in which the association is only between cue and outcome, and in instrumental learning, in which the outcome depends on an action[33,34]. However, ACC cells encode the expected value of actions better than they encode the expected value of stimuli[25], and similarly they encode the prediction error preferentially in instrumental situations when the learned information will guide future actions[25]; the same neurons do not respond to an unexpected visual stimulus. Although such studies suggest that the ACC has a special role in encoding the value of actions, other researchers disagree; some ACC neurons represent less information about the direction of previous eye movement responses than do

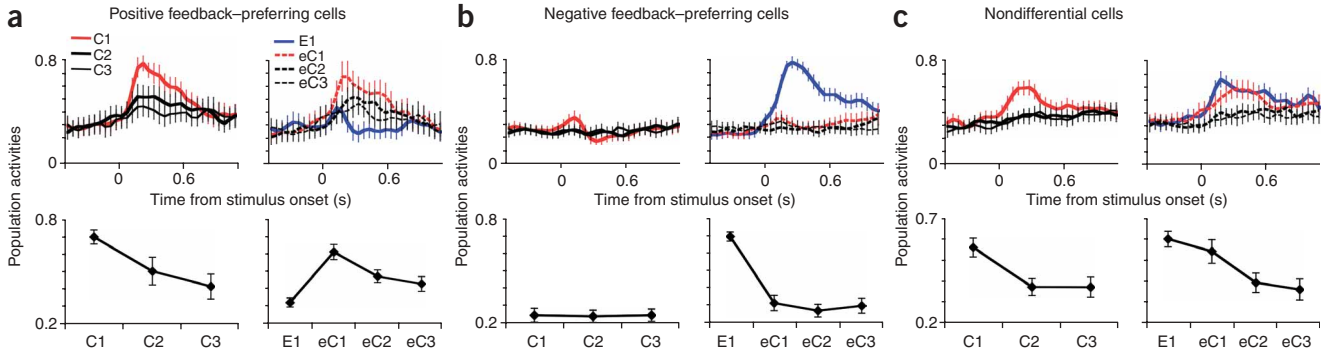

Figure 2 Prediction error encoding in the ACC. Changes of activity in a population of medial frontal cortical neurons centered on the ACC sulcus during the course of learning which action was rewarded with a secondary reinforcer. (**a**–**c**) Averaged responses of 16 positive feedback–preferring cells (**a**), 32 negative feedback–preferring cells (**b**) and 34 nondifferential cells (**c**). Bin width in upper graphs in each section, 50 ms. The activity of each cell was normalized by its peak activity and then averaged across cells. Each graph shows activity across three trials of a typical problem set. On the first trial, monkeys did not know which was the correct action to choose. On half of trials (left column, C1) the monkeys guessed correctly and chose the action associated with a positive secondary reinforcer. Usually the monkeys continued to choose correctly on the subsequent trials (C2 and C3) on these blocks. In the other half of blocks, the monkeys' first choices were incorrect (right column, E1). The monkeys usually corrected their choices on the subsequent three trials (eC1, eC2 and eC3). Positive feedback–preferring neurons and nondifferential neurons were active in relation to the positive prediction error when the first choice was made correctly but subsequently decreased their activity once the correct choice was known. Negative feedback–preferring neurons and nondifferential neurons were active in relation to the negative prediction error when the first choice was made incorrectly but subsequently decreased their activity once the correct choice was known. (Reprinted with permission from Matsumoto *et al.*[25] (*Nature Neuroscience*)).

neurons in lateral prefrontal cortex[23]. Nevertheless, the ACC region in which most recordings are made is in a position to influence and be influenced by action-selection processes: it is anatomically interconnected with the adjacent rostral cingulate motor area (CMAr), which has direct projections to the spinal cord and the primary motor cortex[37,38]. CMAr neurons show activity that is time-locked to the execution of movements[39,40]. Although CMAr neurons encode comparatively little information about the spatial features of an action, their firing rates do contain information about other aspects of plans, such as the sequential order in which actions are made[39,40]. By contrast, whereas dopamine neurons encode the expected value of the action that will be taken, or in some cases the expected value of the best action that could be taken in a given context[7], their activity does not seem to contribute to the selection of the action itself in a simple way[34], and the presence of prediction error activity is not always related to subsequent adjustments in action selection[33]. In reinforcement learning models, such a situation could occur in the presence of a low learning rate.

Parameters encoded in cingulate cortex

In addition to the reward expectation and the prediction error, the learning rate and the statistical parameters of the environment that determine the learning rate are also encoded in ACC activity when new information is observed. When subjects are asked to adjust their behavior in response to new outcomes while the statistics of the reward environment change, subjects flexibly adapt their learning rates such

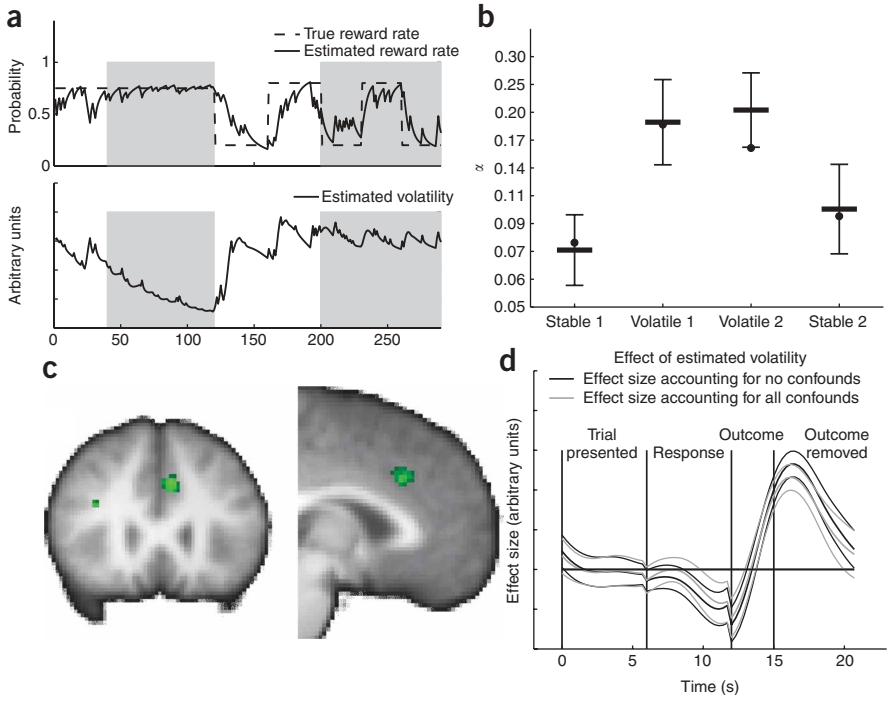

Figure 3 ACC, volatility and the learning rate. Learning rates are flexibly adapted to best suit environmental statistics, and this effect is mediated by the ACC. (**a**) Subjects underwent a decision making task in which the reward rate changed. Crucially, this changing reward rate went through periods of stability and periods of volatility (top panel). Optimal behavior requires that the subjects estimate this volatility (bottom panel) and adjust their learning rate accordingly. (**b**) Subject learning rates (α) during the stable and volatile phases of the experiment. Bars, mean ± s.e.m. for human subjects. Dots, optimal learning rate. (**c**) A circumscribed region in the ACC correlates with the volatility estimate (or the related uncertainty). (**d**) Time course of the effect size in ACC. BOLD signal is related to the estimated volatility only when the outcome is observed. (Reprinted with permission from Behrens *et al.*[5] (*Nature Neuroscience*)).

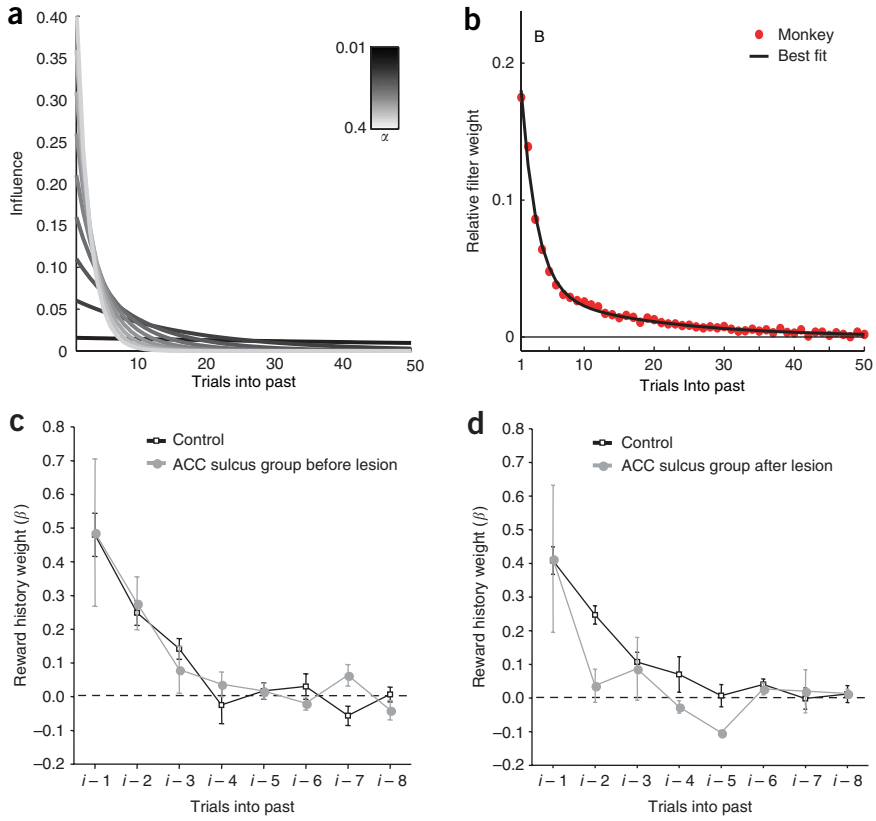

Figure 4 Interrelationships between the learning rate and the reward history. Adaptive nature of reward integration, and its mediation by the ACC. In all cases the curves show the influence of outcomes from different past trials on the current decision. (**a**) Relationship between learning rate (α) and reward integration curves. Higher learning rates (lighter shades) imply more influence of recent trials, and shorter integration curves than lower learning rates (darker curves). (**b**) Empirical integration curve from a macaque monkey when reward conditions are relatively stable. The current decision is influenced by many trials in the past[41,43]. (**c**) Empirical integration curve from macaque monkeys when reward conditions are relatively volatile. The present decision (decision *i*) is influenced by only four trials into the past[42]. (**d**) After a lesion to the ACC sulcus, macaques are no longer able to appropriately combine recent and historical information to guide behavior. Only the most recent trial guides the current decision[42]. Reprinted with permission from Corrado *et al.*[43] (*Journal of Experimental Analysis of Behavior*) and Kennerley *et al.*[42] (*Nature Neuroscience*).

that in stable conditions, new prediction errors have little effect on future actions, in line with theoretical predictions. In fast-changing or volatile situations, however, subjects learn quickly from new outcomes. To achieve such flexible behavior, subjects must do more than simply update the value of each action. They must also track the volatility, a higher-order statistic of the environment. Theoretical accounts suggest that volatility in the environment induces uncertainty in the current estimate of value. In uncertain situations, subjects should give more weight to new outcomes, implying a faster learning rate (**Figs. 1** and **3**)[3–5]. These crucial parameters of volatility and uncertainty correlate with the BOLD response in the ACC sulcus at the time each new outcome is observed—the crucial time for learning (**Fig. 3**). Experimental controls in these studies allowed these signals influencing the learning rate to be identified independently from signals representing the prediction error.

The learning rate dictates the extent to which an action's expected value is determined by its past history of reward—the reward integration curve. Under a fast learning rate, only the most recent outcomes are relevant, whereas under a slow learning rate, even historical outcomes have a bearing on the next decision (**Fig. 4a**). A learning rate that is flexibly adapted to suit the reward environment can therefore explain differences in the length of this curve that have been reported in recent experiments[33,41,42]. For example, outcomes from more than 30 trials ago still had some influence over the values of choice options in one study[41,43] (**Fig. 4b**), whereas a much shorter reward integration period was reported in a similar task also performed by macaques[42] (**Fig. 4c**). In the former experiment, reward contingencies remained stable for hundreds of trials, whereas in the latter experiment, the monkeys experienced a volatile environment that switched approximately every 25 trials. The importance of the ACC in mediating the influence of the reward environment on the learning rate and therefore on future actions is underlined by the finding that, after an ACC lesion, only the outcome on the most recent trial exerts any influence over subsequent decisions[42] (**Fig. 4d**).

There is further evidence that the ACC mediates the degree to which an outcome will guide learning and future behavior. In one study, human subjects were scanned while they moved around a virtual maze that they had previously explored outside the scanner[44]. The authors derived estimates of subjects' beliefs about their current maze positions on the basis of the observations that the subjects were able to make at each point in the virtual maze. The authors also estimated when the observations the subjects made would have been discrepant with previously held beliefs. On such trials, the feedback was especially valuable and enabled subjects to most significantly revise their beliefs about their positions in the maze. These same trials induced increases in ACC activity.

Finally, a study discussed above[25] also found a class of neuron with activity that appeared to be independent of whether the response was correct, but that diminished throughout the four trials of each block as less and less could be learned from each new outcome. Such neurons might be expected if the ACC were encoding the value of each new outcome for learning. Taken together with the prediction error signals, these findings suggest that ACC activity when new information is observed reflects the extent to which the current outcome should dictate future actions, or the value of information attained from the current outcome.

Cingulate cortex and acquisition of new information

Knowing the value of information is not only essential when considering how much influence new outcomes should have in learning. The value of potential information is another crucial factor determining choice. Animals should be prepared to sacrifice immediate reward for the opportunity to acquire information that will yield greater rewards in the future. Several lines of converging evidence suggest that the ACC is also active before a decision that could potentially yield information. One direct test of the idea comes from situations wherein subjects are

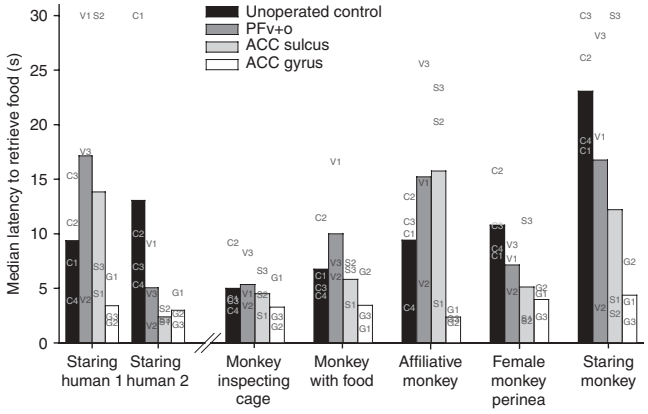

Figure 5 Anterior cingulate cortex and the valuation of social information. The ACC gyrus is necessary for the normal valuation of social information. Control macaques were slow to pick up food rewards in the presence of pictures of conspecifics. Macaques with ACC gyrus lesions did not show the normal pattern of preparedness to sacrifice food in order to acquire information by observing pictures of other animals. The figure shows the median latency to retrieve food in the presence of social stimuli—pictures of unknown human actors (left) or macaques (right) shown to four groups of macaques: unoperated controls, combined ventral and orbital prefrontal lesions (PFv+o), ACC sulcus lesions and ACC gyrus lesions. Bars indicate mean group performance. Letter and number pairs indicate performance of individual macaques from the control group (C1–C4), ventral prefrontal and orbital lesion group (V1–V3), ACC sulcus lesion group (S1–S3) and ACC gyrus lesion group (G1–G3). Control macaques were slow to retrieve food in the presence of social stimuli and, in general, their responses indicated that pictures that one control found interesting, others often found interesting; whereas pictures that one control found uninteresting, others found uninteresting (note the general increase in black bar height moving from left to right in the second part of the figure). This modulation of reaching latency by social evaluation was absent in the animals with ACC gyrus lesions. Reprinted with permission from Rudebeck *et al.*[50] (*Science*).

asked directly to sacrifice income for information. For example, subjects were asked to choose repetitively between computer 'bandits' (slot machines) that paid out at different rates[18]. The payout rate of each bandit was changing, such that the optimal behavior was not to find the best bandit and exploit it continually, but rather to sporadically explore the bandits to discover if the identity of the best bandit had changed. By modeling subject learning, the authors were able identify such sporadic exploration trials (in which the subjects did not choose the option valued most highly at the time). These trials, in which the subject chooses information over immediate reward, show an increase in activation in the ACC and other frontal regions. In this task, subjects clearly have mnemonic loads to juggle as they switch between bandits. When subjects are simply asked to make their own choice of action, ACC activation is also prominent[44–46], especially when they attempt to find the best course of action rather than when they are simply asked to generate actions randomly[29]. In macaques learning, by trial and error, the correct order in which to touch three targets to obtain a food reward, ACC neurons that are active during the execution of responses are more active when the macaques are searching for the correct order, but they become less active once the correct order had been learned and the macaque merely had to repeat it[47].

The other clear situation in which subjects are asked to sacrifice income for information comes from studies of social decision making. Here the critical information that an individual needs to know concerns the predispositions and actions of other agents, and, again, a subregion of the ACC may be critical for the evaluation of information obtained in a social context.

Male macaques value the opportunity to acquire information about other macaques, particularly dominant males and females[48,49]—individuals that would normally have the highest potential impact on the macaque's own evolutionary fitness. Under normal circumstances, male macaques will forego food to observe dominant males and females, but this effect is abolished after a lesion of the ACC gyrus[50] (**Fig. 5**). Lesions of the ACC sulcus, the region most important for mediating the effect of reward history on decision making, or of the OFC do not have the same effect. ACC, but not OFC, lesions in rodents similarly devalue the acquisition of social information[51].

Activation changes can also be recorded with fMRI in the ACC when people witness decisions made by others or make decisions that will influence others. Although the precise location of the activation depends on whether they are making a decision or observing one, the critical regions, as in the macaque, always appear to lie within the ACC gyrus[52]. It may seem counterintuitive to compare the valuation of social information to the valuation of information in reinforcement learning models, and indeed social information has been argued to impede the operation of feedback-based mechanisms of social learning[53]. Nevertheless, social expectations themselves are governed by reinforcement learning principles; for example, social expectations implicit in the intention to trust another individual seem to be activated at increasingly early time points during successive social exchanges, as would be predicted by reinforcement learning models[54]. The activations recorded in ACC and adjacent medial frontal cortex during social exchanges are enhanced when individuals believe that they are interacting with another person as opposed to an artificial agent such as a computer[55]. Such results raise the intriguing, but as yet untested, possibility that parallel ACC sulcal and gyral mechanisms encode the value of reward and social information for decision making.

Prefrontal cortex and reward expectation

Neurons with activity patterns that encode reward expectations have also been reported in the prefrontal cortex, particularly in orbital frontal cortex (OFC), in both macaques and rats[56,57]. Although there is little evidence for a prediction error signal in the OFC and comparatively little information about actions to select rewards[21,57–59], it is clear that OFC neurons encode details concerning the identities of rewards; in particular, their visual appearance, taste, smell and texture[60,61]. In some cases, OFC activity encodes expectations about the receipt of a reward of a particular magnitude[59] and identity[57,62]. Several lines of evidence suggest that the OFC is necessary for the representation of the expected value of specific reinforcement outcomes. First, when the value of an outcome is reduced by satiety, then the activity of OFC neurons decreases. Satiety can be specific to a certain food type, and satiety-related decrements in OFC neuron activity can also be specific to a certain food[63]. OFC lesions disrupt the ability to adapt choices in the light of a change in a reward's value, either through satiety or through an association with illness[64–66], and the effects are specific to choices guided by stimulus-reward associations rather than by response-reward associations[58,67].

In some neurons, firing rate also varies with the value of whichever outcome is expected, independently of the reward type[68]. When the food choices of individual macaques are used to derive an operational measure of their values, some OFC neurons' activities increase in proportion to the value that the individual animal assigns to the expected reward type and the quantity expected. OFC activity encodes the absolute value of the offer, at least in the short term, as the monkey proceeds from one task trial to the next; a given neuron's activity

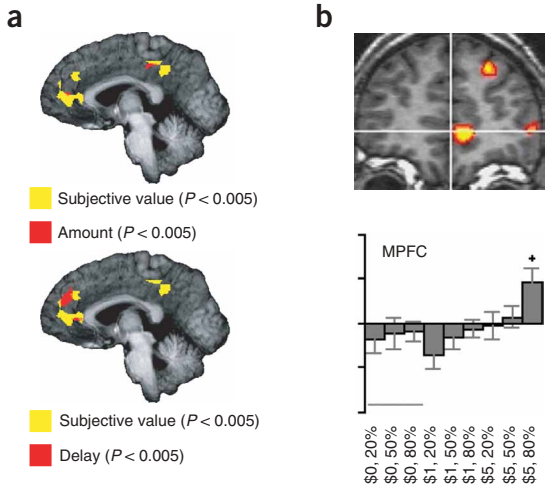

Figure 6 VMPFC activation and representation of subjective value. (**a**) Human subjects made decisions about whether to opt for delayed monetary rewards of various amounts at various delays or for a standard payment of US$20 that would be made immediately. The rates at which the values of delayed rewards were discounted by subjects were calculated from each individual's choice data. The resulting estimates of the subjective values of choice options were then regressed against the fMRI-recorded brain activity. Activity in the VMPFC and adjacent ACC and in the posterior cingulate cortex was better correlated with (top) the subjective values of the choice options (yellow) than with the objective amount of monetary reward (red) or with (bottom) the objective delay to the monetary reward (red). Reprinted with permission from Kable and Glimcher[72] (*Nature Neuroscience*). (**b**) Subjects were shown stimuli that predicted different reward magnitudes with different probabilities. VMPFC activity (top, crosshairs) increased linearly both with increasing reward probability and with increasing reward magnitude (bottom; error bars, s.e.m.). Activity therefore encoded the expected value of the stimulus. Reprinted with permission from Knutson *et al.*[70] (*Journal of Neuroscience*).

in response to a particular reward remains invariant regardless of whether it is paired with a less or more preferred option[62]. Such an activity pattern may underlie the transitivity of value judgments. Over the longer term, however, OFC neuron activity recorded in relation to a particular outcome expectation changes depending on the other rewards that are available[57], such that OFC neurons encoding a particular reward will show more activity if this reward is the preferred one among the options available. Such OFC activity might underlie context-dependent preferences. Activity in parietal areas also varies with the relative reward expectations associated with different possible actions, but it is not clear whether regions outside the OFC contain representations that are specific to both stimulus and outcome and that contain at least some information about absolute value over some time scales[41,69].

Further support for the idea of value representations in the orbitofrontal cortex has been adduced from human fMRI studies. A consistent finding is that activity in a ventromedial prefrontal cortex (VMPFC) is correlated with the expected reward in a great many situations[17,70–75]. In one recent study[74], hungry humans were asked how much they would be prepared to pay for different food items; VMPFC activity varied with the subjects' valuation of each item as judged by their willingness to pay. In another study[72] (**Fig. 6**), a choice procedure was used to determine the value assigned by different individuals to monetary rewards of different sizes and expected delays. VMPFC activity was correlated with individual valuations.

Integrating rewards and costs in decision making

It is tempting to relate the findings of VMPFC activation reported in human subjects with studies of reward-related single-neuron activity in the OFC of rats and macaques. Such an argument would tend to conclude that the OFC represents value in an abstract and context-independent manner that could provide a 'common currency' for decision making. It is not clear, however, that it is correct to make this argument. The precise location of the VMPFC activation varies from one study to the next and even between subjects within a study[72], but it is usually located on the medial rather than the orbital surface (**Fig. 6**). The neurons in the macaque, on the other hand, that have been reported to encode value lie in a distinct region in the central part of the orbital surface between the medial and lateral orbital sulci[57,62]. Anatomically, the VMPFC region in humans and other primates contains three distinct divisions—medial orbitofrontal cortex, ventral cingulate cortex and posterior parts of the frontopolar cortex[76,77]—and one could argue for identifying VMPFC activity with any of these three main divisions. Even if VMPFC activations are within the OFC, then they are likely to fall within its most medial part, which, in the macaque, is strongly interconnected with the ACC and medial frontal cortex but comparatively weakly interconnected with the remainder of the OFC[78]. Very little is known of the functions or physiology of macaque VMPFC.

Whereas many single-neuron studies have concentrated on the encoding of rewards in a controlled learning environment, in real-life situations, decisions and actions involve costs[79]. For example, a course of action may lead to a benefit such as a large reward but only after a cost, such as a long delay, is encountered. Both the reward and the cost in conjunction determine a choice's value. If the OFC does maintain a context-independent representation of value as a common currency for decision making, then OFC neurons that increase their firing rates in relation to increases in the expected magnitude of a reward should also decrease their activity in relation to the expected cost of the action. Although this may occur in macaques[80], such encoding is not readily observable in the OFC of rats[81] (**Fig. 7**). The discrepancy between results in rat and monkey may reflect either a genuine species difference or a difference in the extent of training (substantial and minimal in the monkey and rat studies, respectively). Nevertheless, the evidence from the rat suggests that OFC represents the costs of a choice option independently from its associated reward magnitude[81]. The pattern of results suggest that the distributed activity of OFC neurons does not just encode the integrated current value of a choice outcome in a unitary manner[57,62,68], but also maintains a rich representation of many aspects of an expected reward, including its intrinsic features, such as its identity[57,62,68], taste, smell and texture[60,61]. At least in rodents, the OFC also encodes other variable reward features, such as associated delay[81] and spatial position[82], that are also often important for decision making.

OFC lesions alter decisions about rewards that are expected after a delay[83]. Normally rats can learn to choose a course of action that ultimately leads to a larger reward after a delay, but after OFC lesions rats sometimes prefer actions that lead to an immediate small reward[84,85]. The deficit suggests that the OFC is essential for certain aspects of value representation, but it does not indicate a fundamental impairment in all aspects of value representation because animals with lesions in the OFC can make appropriate decisions when the cost to be taken into account concerns the effort that must be expended before the reward is acquired rather than a delay. ACC lesions, by contrast, cause the opposite pattern: impairment in

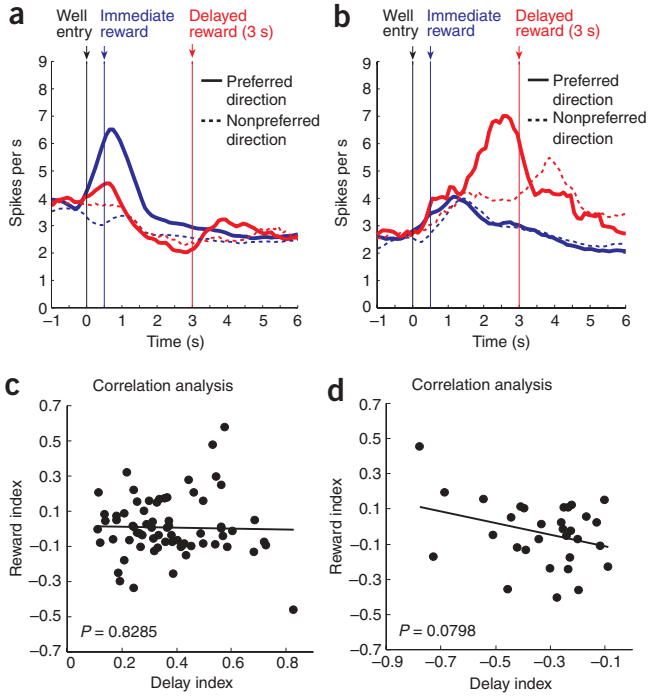

Figure 7 OFC single-neuron activity does not represent reward value in a simple unitary code. (**a**) Population histogram representing firing rate as a function of time during the trial for neurons that fired more strongly during a reward epoch with short delays (blue) than in an epoch with longer delays (red) of 3 s ($n = 65$). Activity is aligned on entry to the food well in order to retrieve food pellets. Solid lines, preferred direction of food reward; dashed lines, nonpreferred direction of food reward. The neurons were more active on immediate reward trials than during expectation of delayed rewards. (**b**) In contrast, other neurons were more active during expectation of delayed rewards than immediate rewards ($n = 27$). Regardless of the neurons' preference for delay, there was no systematic tendency for the neurons to have related preferences for reward size. (**c**,**d**) Relation of firing dependent on delay length to firing dependent on reward size for those neurons that fired more strongly after long delays. Neurons with a preference for immediate reward delivery were no more likely to have a preference for larger rewards (**c**), whereas neurons with a preference for later reward delivery were no more likely to have a preference for small rewards (**d**). The delay index and reward index are computed on the basis of firing during the reward epoch. Delay index $= (S - L)/(S + L)$, where S and L represent firing rates on short and long delay trials, respectively. Reward index $= (B - S)/(B + S)$, where B and S represent firing rates on big and small reward trials, respectively. Reprinted with permission from Roesch *et al.*[81] (*Neuron*).

effort-based decision making but normal performance on the delay-based decision making task[85–87].

Although the OFC maintains a rich and detailed representation of the various features of a potential reward, such as its anticipated delay, its relative value compared to other possible outcomes, and its associations with objects and stimuli in the environment, it is possible that the ACC represents the integrated value of a course of action to reflect both the action's intrinsic costs and benefits[58]. There are certainly neurons in the ACC that reflect not just expected reward but also progress through a course of actions toward the reward[88], as successive actions predict the same future reward but less future incurred cost. Such a potential distinction between OFC and ACC is reminiscent of the distinction[89] between a goal-based representation of rewards and the representation of an action's cached value. ACC-dependent reinforcement learning mechanisms may operate on the integrated action value rather than operating independently on costs and benefits and then combining the two. The precise homologies between the rat ACC and the primate medial frontal cortex are unclear, but several fMRI studies find value representations in the human ACC as well as in adjacent parts of the medial frontal cortex and VMPFC[70,72,75].

Risk and decision costs

The values of rewards that are distant in time should be diminished because there is risk as to whether or not the reward will ever be delivered or whether the animal will be present if the reward eventually becomes available. This risk represents a cost that should be taken into account before the delayed reward is pursued and the animal forgoes the opportunity to pursue alternative rewards.

The most direct way to investigate neural coding of risk is by comparing responses to stimuli with different reward probabilities (**Fig. 1**). Whereas activities in dopaminergic[10] and striatal[90] regions are correlated with prediction risk that is temporally dissociated from their phasic response to probability, similar responses in the lateral OFC predict individual risk attitudes in decision making[91].

A strong evolutionary correlate of risk is the delay to the time at which a reward is expected, a further term sometimes used in reinforcement learning equations (for example, equation (1)) for updating expected value. As the size of the delay discounting term is changed, rewards that are further away in time can come to have an increasingly small impact on decision making, so that only the most immediate rewards are treated as important. Several fMRI studies have sought correlates of the discount term, reporting activity in lateral prefrontal cortex (LPFC) and OFC[17,71,92].

The previous section discussed the role of the OFC in representing many aspects of potential reward expectations; related signals are also seen in the LPFC. In the macaque, the LPFC, like the OFC, encodes expectations of particular types and amounts of reward[59,93,94], but LPFC neurons also encode information about the monkey's responses and states of the environment[95–97]. LPFC neurons are therefore able to represent the sequence of steps and state transitions that lead from the present to the desired goal[98,99]. In addition, the ensemble of LPFC neurons represents uncertainty about the agent's position in the unfolding plan. Human subjects' conditional uncertainty about their progress through a maze, estimated by a bayesian incremental belief-updating approach, is correlated with activation in the LPFC[44]. In the macaque, LPFC neurons encode particular routes through mazes, but the emergence of distinctive encoding parallels the animal's certainty, as indexed by reaction time measures, concerning the best route to take to obtain an eventual reward[100]. Because LPFC activity in delayed-reward tasks does not scale with expected reward in a simple way[17], it may not be important for representing the distant rewards themselves. Instead, LPFC activity may represent the environmental states and responses, and uncertainty about those states and responses, during progression toward a distant reward.

Summary and conclusions

Learning and decision making have been explained by formalisms such as reinforcement learning, and activity in several brain areas has been interpreted as reflecting components of reinforcement learning mechanisms. Despite a justifiable focus on the striatum and the dopaminergic system, a full account of the neural mechanisms of learning and decision making also requires an understanding of prefrontal and cingulate cortex. This evidence derives from neuroimaging studies that reflect indirect but pooled neuronal activity, from

single-neuron recordings that may only focus on a subset of neurons but confirm that the spiking output of these neurons represents the signals, and from lesion studies that demonstrate the causal importance of such activity for behavior.

Such evidence suggests that ACC activity preceding a decision encodes the integrated value of an action, whether in terms of immediate gains and costs or in terms of information to aid future decision making. On the observation of an outcome, ACC activity encodes the degree to which the resulting information should influence future decisions. A functional division may exist in the ACC between the sulcus, implicated in reward-based learning and decision making, and the gyrus, implicated in social learning and decision making. Although in some circumstances OFC activity encodes integrated value signals, OFC activity also contains information about a wealth of other features that define reinforcement goals.

ACKNOWLEDGMENTS

Funded by the UK Medical Research Council.

Published online at http://www.nature.com/natureneuroscience
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions

1. Rescorla, R. & Wagner, A. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. in *Classical Conditioning* (eds. Black, A.H. & Prokasy, W.F.) 64–99 (Appleton-Century-Crofts, New York, 1972).
2. Sutton, R. & Barto, A.G. *Reinforcement Learning* (MIT Press, Cambridge, Massachusetts, 1998).
3. Dayan, P., Kakade, S. & Montague, P.R. Learning and selective attention. *Nat. Neurosci.* **3** (suppl.): 1218–1223 (2000).
4. Courville, A.C., Daw, N.D. & Touretzky, D.S. Bayesian theories of conditioning in a changing world. *Trends Cogn. Sci.* **10**, 294–300 (2006).
5. Behrens, T.E., Woolrich, M.W., Walton, M.E. & Rushworth, M.F. Learning the value of information in an uncertain world. *Nat. Neurosci.* **10**, 1214–1221 (2007).
6. Schultz, W. Behavioral theories and the neurophysiology of reward. *Annu. Rev. Psychol.* **57**, 87–115 (2006).
7. Roesch, M.R., Calu, D.J. & Schoenbaum, G. Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat. Neurosci.* **10**, 1615–1624 (2007).
8. Tobler, P.N., Fiorillo, C.D. & Schultz, W. Adaptive coding of reward value by dopamine neurons. *Science* **307**, 1642–1645 (2005).
9. Satoh, T., Nakai, S., Sato, T. & Kimura, M. Correlated coding of motivation and outcome of decision by dopamine neurons. *J. Neurosci.* **23**, 9913–9923 (2003).
10. Fiorillo, C.D., Tobler, P.N. & Schultz, W. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* **299**, 1898–1902 (2003).
11. Samejima, K., Ueda, Y., Doya, K. & Kimura, M. Representation of action-specific reward values in the striatum. *Science* **310**, 1337–1340 (2005).
12. O'Doherty, J.P., Deichmann, R., Critchley, H.D. & Dolan, R.J. Neural responses during anticipation of a primary taste reward. *Neuron* **33**, 815–826 (2002).
13. Haruno, M. *et al.* A neural correlate of reward-based behavioral learning in caudate nucleus: a functional magnetic resonance imaging study of a stochastic decision task. *J. Neurosci.* **24**, 1660–1665 (2004).
14. Haruno, M. & Kawato, M. Different neural correlates of reward expectation and reward expectation error in the putamen and caudate nucleus during stimulus-action-reward association learning. *J. Neurophysiol.* **95**, 948–959 (2006).
15. O'Doherty, J. *et al.* Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* **304**, 452–454 (2004).
16. Haruno, M. & Kawato, M. Heterarchical reinforcement-learning model for integration of multiple cortico-striatal loops: fMRI examination in stimulus-action-reward association learning. *Neural Netw.* **19**, 1242–1254 (2006).
17. Tanaka, S.C. *et al.* Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nat. Neurosci.* **7**, 887–893 (2004).
18. Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B. & Dolan, R.J. Cortical substrates for exploratory decisions in humans. *Nature* **441**, 876–879 (2006).
19. Cohen, M.X. & Ranganath, C. Reinforcement learning signals predict future decisions. *J. Neurosci.* **27**, 371–378 (2007).
20. Williams, S.M. & Goldman-Rakic, P.S. Widespread origin of the primate mesofrontal dopamine system. *Cereb. Cortex* **8**, 321–345 (1998).
21. Matsumoto, K., Suzuki, W. & Tanaka, K. Neuronal correlates of goal-based motor selection in the prefrontal cortex. *Science* **301**, 229–232 (2003).
22. Amiez, C., Joseph, J.P. & Procyk, E. Reward encoding in the monkey anterior cingulate cortex. *Cereb. Cortex* **16**, 1040–1055 (2006).
23. Seo, H. & Lee, D. Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J. Neurosci.* **27**, 8366–8377 (2007).
24. Amiez, C., Joseph, J.P. & Procyk, E. Anterior cingulate error-related activity is modulated by predicted reward. *Eur. J. Neurosci.* **21**, 3447–3452 (2005).
25. Matsumoto, M., Matsumoto, K., Abe, H. & Tanaka, K. Medial prefrontal cell activity signaling prediction errors of action values. *Nat. Neurosci.* **10**, 647–656 (2007).
26. Mars, R.B. *et al.* Neural dynamics of error processing in medial frontal cortex. *Neuroimage* **28**, 1007–1013 (2005).
27. Ullsperger, M., Nittono, H. & von Cramon, D.Y. When goals are missed: dealing with self-generated and externally induced failure. *Neuroimage* **35**, 1356–1364 (2007).
28. Klein, T.A. *et al.* Genetically determined differences in learning from errors. *Science* **318**, 1642–1645 (2007).
29. Walton, M.E., Devlin, J.T. & Rushworth, M.F.S. Interactions between decision making and performance monitoring within prefrontal cortex. *Nat. Neurosci.* **7**, 1259–1265 (2004).
30. Debener, S. *et al.* Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *J. Neurosci.* **25**, 11730–11737 (2005).
31. Holroyd, C.B. & Coles, M.G. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.* **109**, 679–709 (2002).
32. Oliveira, F.T., McDonald, J.J. & Goodman, D. Performance monitoring in the anterior cingulate is not all error related: expectancy deviation and the representation of action-outcome associations. *J. Cogn. Neurosci.* **19**, 1994–2004 (2007).
33. Bayer, H.M. & Glimcher, P.W. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47**, 129–141 (2005).
34. Morris, G., Nevet, A., Arkadir, D., Vaadia, E. & Bergman, H. Midbrain dopamine neurons encode decisions for future action. *Nat. Neurosci.* **9**, 1057–1063 (2006).
35. Tobler, P.N., Dickinson, A. & Schultz, W. Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. *J. Neurosci.* **23**, 10402–10410 (2003).
36. Ljunberg, T., Apicella, P. & Schultz, W. Responses of monkey dopamine neurons during learning of behavioral reactions. *J. Neurophysiol.* **67**, 145–163 (1992).
37. Dum, R.P. & Strick, P.L. Spinal cord terminations of the medial wall motor areas in macaque monkeys. *J. Neurosci.* **16**, 6513–6525 (1996).
38. Van Hoesen, G.W., Morecraft, R.J. & Vogt, B.A. Connections of the monkey cingulate cortex. in *Neurobiology of Cingulate Cortex and Limbic Thalamus* (eds. Vogt, B.A. & Gabriel, M.) 249–284 (Birkhauser, Boston, 1993).
39. Akkal, D., Bioulac, B., Audin, J. & Burbaud, P. Comparison of neuronal activity in the rostral supplementary and cingulate motor areas during a task with cognitive and motor demands. *Eur. J. Neurosci.* **15**, 887–904 (2002).
40. Hoshi, E., Sawamura, H. & Tanji, J. Neurons in the rostral cingulate motor area monitor multiple phases of visuomotor behavior with modest parametric selectivity. *J. Neurophysiol.* **94**, 640–656 (2005).
41. Sugrue, L.P., Corrado, G.S. & Newsome, W.T. Matching behavior and the representation of value in the parietal cortex. *Science* **304**, 1782–1787 (2004).
42. Kennerley, S.W., Walton, M.E., Behrens, T.E., Buckley, M.J. & Rushworth, M.F. Optimal decision making and the anterior cingulate cortex. *Nat. Neurosci.* **9**, 940–947 (2006).
43. Corrado, G.S., Sugrue, L.P., Seung, H.S. & Newsome, W.T. Linear-nonlinear-Poisson models of primate choice dynamics. *J. Exp. Anal. Behav.* **84**, 581–617 (2005).
44. Yoshida, W. & Ishii, S. Resolution of uncertainty in prefrontal cortex. *Neuron* **50**, 781–789 (2006).
45. Deiber, M.-P. *et al.* Cortical areas and the selection of movement: a study with positron emission tomography. *Exp. Brain Res.* **84**, 393–402 (1991).
46. Forstmann, B.U., Brass, M., Koch, I. & von Cramon, D.Y. Voluntary selection of task sets revealed by functional magnetic resonance imaging. *J. Cogn. Neurosci.* **18**, 388–398 (2006).
47. Procyk, E., Tanaka, Y.L. & Joseph, J.P. Anterior cingulate activity during routine and non-routine sequential behaviors in macaques. *Nat. Neurosci.* **3**, 502–508 (2000).
48. 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).
49. Shepherd, S.V., Deaner, R.O. & Platt, M.L. Social status gates social attention in monkeys. *Curr. Biol.* **16**, R119–R120 (2006).
50. Rudebeck, P.H., Buckley, M.J., Walton, M.E. & Rushworth, M.F. A role for the macaque anterior cingulate gyrus in social valuation. *Science* **313**, 1310–1312 (2006).
51. Rudebeck, P.H. *et al.* Distinct contributions of frontal areas to emotion and social behavior in the rat. *Eur. J. Neurosci.* **26**, 2315–2326 (2007).
52. Tomlin, D. *et al.* Agent-specific responses in the cingulate cortex during economic exchanges. *Science* **312**, 1047–1050 (2006).
53. Delgado, M.R., Frank, R.H. & Phelps, E.A. Perceptions of moral character modulate the neural systems of reward during the trust game. *Nat. Neurosci.* **8**, 1611–1618 (2005).
54. King-Casas, B. *et al.* Getting to know you: reputation and trust in a two-person economic exchange. *Science* **308**, 78–83 (2005).
55. Amodio, D.M. & Frith, C.D. Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* **7**, 268–277 (2006).
56. Schoenbaum, G., Chiba, A.A. & Gallagher, M. Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nat. Neurosci.* **1**, 155–159 (1998).
57. Tremblay, L. & Schultz, W. Relative reward preference in primate orbitofrontal cortex. *Nature* **398**, 704–708 (1999).
58. Rushworth, M.F., Behrens, T.E., Rudebeck, P.H. & Walton, M.E. Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behavior. *Trends Cogn. Sci.* **11**, 168–176 (2007).
59. Wallis, J.D. & Miller, E.K. Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur. J. Neurosci.* **18**, 2069–2081 (2003).

60. Rolls, E.T. & Baylis, L.L. Gustatory, olfactory, and visual convergence within the primate orbitofrontal cortex. *J. Neurosci.* **14**, 5437–5452 (1994).
61. Rolls, E.T., Critchley, H.D., Browning, A.S., Hernadi, I. & Lenard, L. Responses to the sensory properties of fat of neurons in the primate orbitofrontal cortex. *J. Neurosci.* **19**, 1532–1540 (1999).
62. Padoa-Schioppa, C. & Assad, J.A. The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nat. Neurosci.* **11**, 95–102 (2008).
63. Rolls, E.T., Sienkiewicz, Z.J. & Yaxley, S. Hunger modulates the responses to gustatory stimuli of single neurons in the caudolateral orbitofrontal cortex of the macaque monkey. *Eur. J. Neurosci.* **1**, 53–60 (1989).
64. Gallagher, M., McMahan, R.W. & Schoenbaum, G. Orbitofrontal cortex and representation of incentive value in associative learning. *J. Neurosci.* **19**, 6610–6614 (1999).
65. Baxter, M.G., Parker, A., Lindner, C.C.C., Izquierdo, A.D. & Murray, E.A. Control of response selection by reinforcer value requires interaction of amygdala and orbital frontal cortex. *J. Neurosci.* **20**, 4311–4319 (2000).
66. Izquierdo, A., Suda, R.K. & Murray, E.A. Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. *J. Neurosci.* **24**, 7540–7548 (2004).
67. Ostlund, S. & Balleine, B.W. The contribution of orbitofrontal cortex to action selection. *Ann. N Y Acad. Sci.* **1121**, 174–192 (2007).
68. Padoa-Schioppa, C. & Assad, J.A. Neurons in the orbitofrontal cortex encode economic value. *Nature* **441**, 223–226 (2006).
69. Platt, M.L. & Glimcher, P.W. Neural correlates of decision variables in parietal cortex. *Nature* **400**, 233–238 (1999).
70. Knutson, B., Taylor, J., Kaufman, M., Peterson, R. & Glover, G. Distributed neural representation of expected value. *J. Neurosci.* **25**, 4806–4812 (2005).
71. McClure, S.M., Laibson, D.I., Loewenstein, G. & Cohen, J.D. Separate neural systems value immediate and delayed monetary rewards. *Science* **306**, 503–507 (2004).
72. Kable, J.W. & Glimcher, P.W. The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* **10**, 1625–1633 (2007).
73. Hampton, A.N., Bossaerts, P. & O'Doherty, J.P. The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *J. Neurosci.* **26**, 8360–8367 (2006).
74. Plassmann, H., O'Doherty, J. & Rangel, A. Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J. Neurosci.* **27**, 9984–9988 (2007).
75. Knutson, B., Rick, S., Wimmer, G.E., Prelec, D. & Loewenstein, G. Neural predictors of purchases. *Neuron* **53**, 147–156 (2007).
76. Vogt, B.A., Vogt, L., Farber, N.B. & Bush, G. Architecture and neurocytology of monkey cingulate gyrus. *J. Comp. Neurol.* **485**, 218–239 (2005).
77. Ongur, D., Ferry, A.T. & Price, J.L. Architectonic subdivision of the human orbital and medial prefrontal cortex. *J. Comp. Neurol.* **460**, 425–449 (2003).
78. Carmichael, S.T. & Price, J.L. Connectional networks within the orbital and medial prefrontal cortex of macaque monkeys. *J. Comp. Neurol.* **371**, 179–207 (1996).
79. Walton, M.E., Kennerley, S.W., Bannerman, D.M., Phillips, P.E. & Rushworth, M.F. Weighing up the benefits of work: behavioral and neural analyses of effort-related decision making. *Neural Netw.* **19**, 1302–1314 (2006).
80. Roesch, M.R. & Olson, C.R. Neuronal activity in primate orbitofrontal cortex reflects the value of time. *J. Neurophysiol.* **94**, 2457–2471 (2005).
81. Roesch, M.R., Taylor, A.R. & Schoenbaum, G. Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. *Neuron* **51**, 509–520 (2006).
82. Feierstein, C.E., Quirk, M.C., Uchida, N., Sosulski, D.L. & Mainen, Z.F. Representation of spatial goals in rat orbitofrontal cortex. *Neuron* **51**, 495–507 (2006).
83. Kheramin, S. *et al.* Effects of quinolinic acid-induced lesions of the orbital prefrontal cortex on inter-temporal choice: a quantitative analysis. *Psychopharmacology (Berl.)* **165**, 9–17 (2002).
84. Mobini, S. *et al.* Effects of lesions of the orbitofrontal cortex on sensitivity to delayed and probabilistic reinforcement. *Psychopharmacology (Berl.)* **160**, 290–298 (2002).
85. Rudebeck, P.H., Walton, M.E., Smyth, A.N., Bannerman, D.M. & Rushworth, M.F. Separate neural pathways process different decision costs. *Nat. Neurosci.* **9**, 1161–1168 (2006).
86. Walton, M.E., Bannerman, D.M. & Rushworth, M.F.S. The role of rat medial frontal cortex in effort-based decision making. *J. Neurosci.* **22**, 10996–11003 (2002).
87. Walton, M.E., Bannerman, D.M., Alterescu, K. & Rushworth, M.F.S. Functional specialization within medial frontal cortex of the anterior cingulate for evaluating effort-related decisions. *J. Neurosci.* **23**, 6475–6479 (2003).
88. Shidara, M. & Richmond, B.J. Anterior cingulate: single neuronal signals related to degree of reward expectancy. *Science* **296**, 1709–1711 (2002).
89. Daw, N.D., Niv, Y. & Dayan, P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* **8**, 1704–1711 (2005).
90. Preuschoff, K., Bossaerts, P. & Quartz, S.R. Neural differentiation of expected reward and risk in human subcortical structures. *Neuron* **51**, 381–390 (2006).
91. Tobler, P.N., O'Doherty, J.P., Dolan, R.J. & Schultz, W. Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. *J. Neurophysiol.* **97**, 1621–1632 (2007).
92. McClure, S.M., Ericson, K.M., Laibson, D.I., Loewenstein, G. & Cohen, J.D. Time discounting for primary rewards. *J. Neurosci.* **27**, 5796–5804 (2007).
93. Watanabe, M. Reward expectancy in primate prefrontal neurons. *Nature* **382**, 629–632 (1996).
94. Kobayashi, S. *et al.* Influences of rewarding and aversive outcomes on activity in macaque lateral prefrontal cortex. *Neuron* **51**, 861–870 (2006).
95. Watanabe, M. & Sakagami, M. Integration of cognitive and motivational context information in the primate prefrontal cortex. *Cereb. Cortex* **17** (suppl. 1), i101–i109 (2007).
96. Wise, S.P. & Murray, E.A. Arbitrary associations between antecedents and actions. *Trends Neurosci.* **23**, 271–276 (2000).
97. Genovesio, A., Brasted, P.J. & Wise, S.P. Representation of future and previous spatial goals by separate neural populations in prefrontal cortex. *J. Neurosci.* **26**, 7305–7316 (2006).
98. Mushiake, H., Saito, N., Sakamoto, K., Itoyama, Y. & Tanji, J. Activity in the lateral prefrontal cortex reflects multiple steps of future events in action plans. *Neuron* **50**, 631–641 (2006).
99. Saito, N., Mushiake, H., Sakamoto, K., Itoyama, Y. & Tanji, J. Representation of immediate and final behavioral goals in the monkey prefrontal cortex during an instructed delay period. *Cereb. Cortex* **15**, 1535–1546 (2005).
100. Averbeck, B.B., Sohn, J.W. & Lee, D. Activity in prefrontal cortex during dynamic selection of action sequences. *Nat. Neurosci.* **9**, 276–282 (2006).