Neuron
Review

Frontal Cortex and Reward-Guided Learning and Decision-Making

Matthew F.S. Rushworth,[1,2,*] MaryAnn P. Noonan,[1,2] Erie D. Boorman,[1,2,3] Mark E. Walton,[1,2] and Timothy E. Behrens[1,2]
[1]Department of Experimental Psychology, University of Oxford, South Parks Road, OX1 3UD, UK
[2]Centre for Functional Magnetic Resonance Imaging of the Brain, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU, UK
[3]Computation and Neural Systems, California Institute of Technology, 1200 E. California Boulevard, Mail Code 136-93, Pasadena, CA 91125, USA
*Correspondence: matthew.rushworth@psy.ox.ac.uk
DOI 10.1016/j.neuron.2011.05.014

Reward-guided decision-making and learning depends on distributed neural circuits with many components. Here we focus on recent evidence that suggests four frontal lobe regions make distinct contributions to reward-guided learning and decision-making: the lateral orbitofrontal cortex, the ventromedial prefrontal cortex and adjacent medial orbitofrontal cortex, anterior cingulate cortex, and the anterior lateral prefrontal cortex. We attempt to identify common themes in experiments with human participants and with animal models, which suggest roles that the areas play in learning about reward associations, selecting reward goals, choosing actions to obtain reward, and monitoring the potential value of switching to alternative courses of action.

Introduction

Recent years have witnessed a great surge of interest in understanding the neural mechanisms of reward-guided learning and decision-making. These developments partly reflect the realization that the structure of decision-making and learning systems reflect a major evolutionary pressure on animals including humans and other primates–the need to seek food and other rewards (Platt and Glimcher, 1999; Glimcher, 2002). Other important factors have been the exploitation of approaches derived from economic decision-making theory that have proven useful in guiding investigation of the ways in which reward value is represented in the brain (Plassmann et al., 2007; Glimcher et al., 2009) and other formal and computational descriptions of reward-guided learning and decision-making (Doya, 2008; Lee, 2008; Platt and Huettel, 2008; Rangel et al., 2008).

Rather than attempting another survey of these recent trends the aim of the current review is to focus more specifically on the role of frontal cortex in reward-guided behavior. It is proposed that current evidence suggests at least four frontal cortex regions can be identified with distinct roles in reward-guided behavior: ventromedial prefrontal cortex and adjacent medial orbitofrontal cortex (vmPFC/mOFC), lateral orbitofrontal cortex (lOFC), anterior cingulate cortex (ACC), and a lateral anterior prefrontal cortex (aPFC) region in, or at least adjacent to, the lateral part of the frontal pole (Figure 1). In reviewing the functions of these areas we draw on work conducted not only with human subjects but also with animal models for which more precise details of neuroanatomical connections and neurophysiological mechanisms are available. In addition to highlighting points of convergence between the studies of various researchers we also note outstanding debates and points of controversy.

Value and vmPFC

The human vmPFC/mOFC has perhaps been the most intensively studied frontal cortical area in investigations of reward-guided decision-making. Functional magnetic resonance imaging (fMRI) measures a blood oxygen level-dependent (BOLD) signal that reflects aspects of underlying neural activity. The correlation between the BOLD signal and behavior or between the BOLD signal and internal states of subjects that can be inferred from behavior is examined. A widely replicated finding is that the vmPFC/mOFC BOLD signal is correlated with the reward value of a choice.

There is agreement about some aspects of the location of the vmPFC/mOFC area that represents reward value but uncertainty about others. On the one hand, studies from many laboratories have identified reward-related signal changes at similar positions in the standard coordinate systems used for reporting fMRI results. The focus lies in the vicinity of the rostral sulcus, ventral and anterior to the rostrum of the corpus callosum on the medial surface of the frontal lobe. The activations extend onto the medial orbital gyrus and sulcus (Beckmann et al., 2009). On the other hand, exactly what name the area should be given has been less clear.

Beckmann et al. (2009) used diffusion-weighted magnetic resonance imaging (DW-MRI) to estimate the anatomical connectivity of medial frontal cortex and then used these estimates to parcellate the cortex into regions, each with different connectivity patterns. The distribution of reward-related BOLD signal changes in a meta-analysis of fMRI studies tracked the location of an area Beckmann et al. (2009) referred to as cluster 2 (Figure 2A). It seems clear that cluster 2 corresponds to a particular cytoarchitectonic area; several anatomists (Ongür et al., 2003; Mackey and Petrides, 2010), but not all (Vogt, 2009) agree that there is an area with a distinctive granular layer 4 with a similar position and orientation to cluster 2. Mackey and Petrides (2010) call the area 14 m. They refer to the adjacent area in the medial orbital gyrus where reward-related activity is also found as 14c (Figure 2B). Mackey and Petrides (2010) locate areas with similar granular and pyramidal cell layers that they

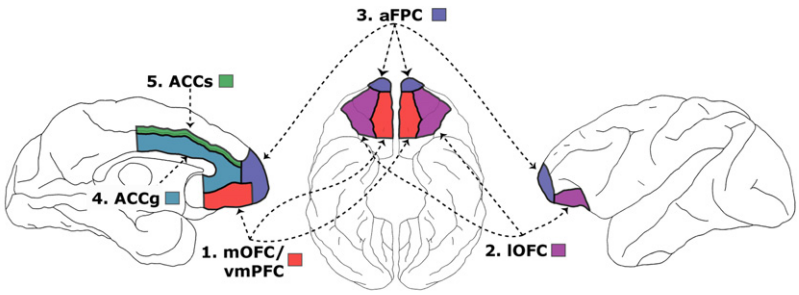

Figure 1. Frontal Brain Regions in the Macaque Involved in Reward-Guided Learning and Decision-Making

vmPFC/mOFC, ventromedial prefrontal cortex and adjacent medial orbitofrontal cortex; lOFC, lateral orbitofrontal cortex; ACCs, anterior cingulate cortex sulcus; ACCg, anterior cingulate cortex gyrus; aPFC, anterior prefrontal cortex.

also refer to as areas 14 m and 14r on the medial orbital gyrus of the macaque. In other words, human vmPFC/mOFC has important similarities with the tissue on the medial orbital gyrus in macaques.

Several ingenious approaches have been used to estimate the value that an object holds for a participant in an experiment in order to examine the correlation between subjective value and the vmPFC/mOFC signal. Plassmann et al. (2007) borrowed the Becker-DeGroot-Marschak method (Becker et al., 1964) used in experimental economics to determine the value of visually presented objects. Participants saw a series of images of food items on a computer monitor while in an MRI scanner and they were asked to indicate how much they were prepared to pay for each item. If the participant's bid exceeded the value of a subsequently generated random number then that participant forfeited the money and was obliged to take the item instead. Subjects made repeated bids over the course of many trials and the choice on one trial was selected at random at the end of the experiment and given to the participant to eat. The procedure provides an estimate of a particpant's "true" valuation of the items under consideration on every trial because subjects have no incentive to bid more or less than they are really "willing to pay" under these conditions. On each trial the vmPFC/mOFC BOLD signal increases with the value that the item has for the participant.

An alternative approach is to let subjects choose between different possible arbitrary stimuli over the course of many trials in an attempt to identify one that is associated with greater reward (typically visual tokens that indicate monetary rewards that will be paid at the end of an experiment). Reinforcement learning algorithms can then be used to estimate the value that is expected on the basis of past experience from choosing each stimulus (Sutton and Barto, 1998). Each time an item is chosen and it yields more reward than expected (in other words, when there is a "positive prediction error") the estimate for the item's value is adjusted upwards. Likewise, when the object is chosen and yields less reward than expected (a "negative prediction error") the item's reward value is revised downwards. The degree of adjustment in each case is a function of the participant's learning rate. Again, estimates of participants' valuations derived from reinforcement learning models are positively correlated with vmPFC/mOFC BOLD signal (Tanaka et al., 2004; Behrens et al., 2008; Gläscher et al., 2009; Wunderlich et al., 2010). A similar effect is seen when the options the subjects are

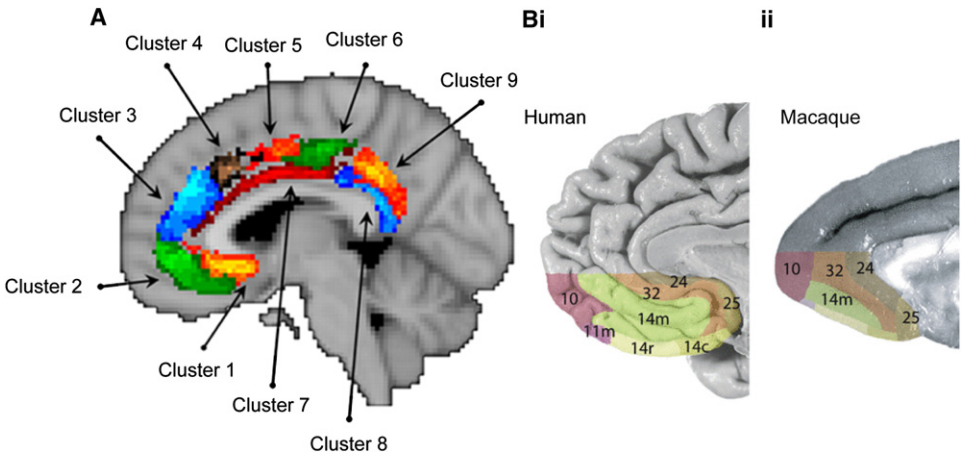

Figure 2. Anatomy of Medial Frontal Cortex

(A) Parcellation of human cingulate cortex and medial frontal cortex by using diffusion-weighted magnetic resonance imaging (DW-MRI) and probabilistic tractography. A meta-analysis of reward-related activations reported in fMRI studies found they were especially prominent in cluster 2 in vmPFC/mOFC. This figure is adapted and reprinted with permission of Beckmann et al., 2009) and the Society for Neuroscience.

(Bi) Cytoarchitectonic parcellation of human orbital and ventromedial brain surface.

(Bii) Cytoarchitectonic parcellation of macaque orbital and ventromedial brain surface. Cluster 2, where many reward expectation-related activations are found, corresponds in shape, size, and position to area 14 m. This figure is adapted and reprinted with permission of Mackey and Petrides (2010) and John Wiley and Sons.

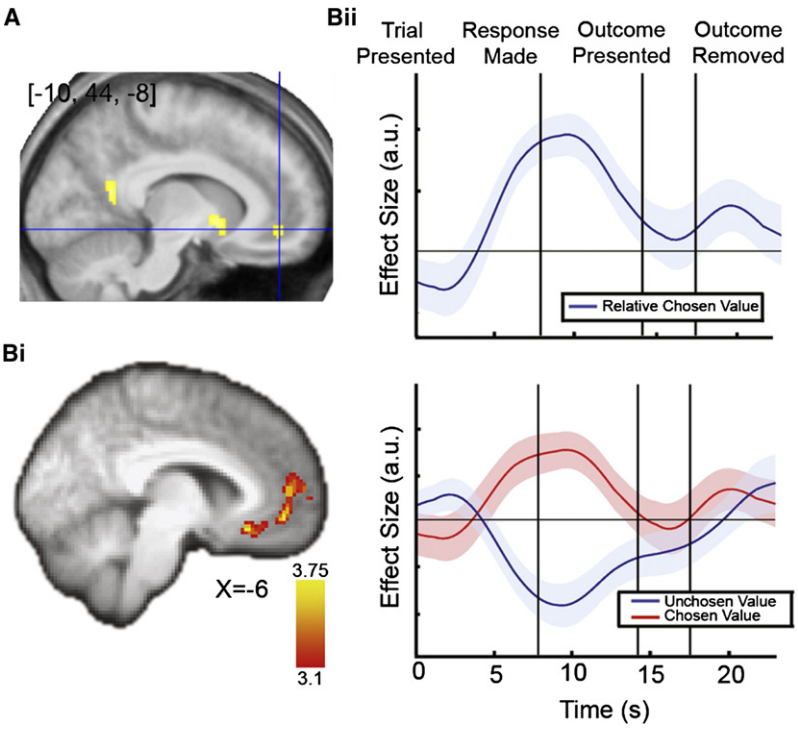

Figure 3. Value Encoding in vmPFC/mOFC

(A) Lebreton et al. (2009) argue that vmPFV/mOFC (crosshairs) are part of an automatic and generic system for valuation. Participants in their experiment provided pleasantness ratings for a range of different stimuli including faces, houses, and paintings and at the end of the experiment indicated preferences between pairs of stimuli. The activity in vmPFC/mOFC on exposure to the stimulus was a function of the interaction between preference and value even when subjects were making different types of judgments (age judgments) at the time that the stimuli were first presented and fMRI data were recorded. This figure is adapted and reprinted with permission of Lebreton et al. (2009).

(Bi) Sagittal slices through Z statistic maps relating to the relative chosen value (chosen – unchosen expected value) of two options during decision-making.

(Bii) Top panel: time course for the effect size of the relative chosen value in the vmPFC is shown throughout the duration of the trial. Bottom panel: the same time course is shown with the signal decomposed into chosen and unchosen action values. Thick lines show mean effect sizes while shadows indicate standard error of the mean effect sizes. There is a positive correlation with chosen value and a negative correlation with unchosen value during the decision-making phase. This figure is adapted and reprinted with permission of Boorman et al. (2009).

choosing between are actions rather than stimuli (Gläscher et al., 2009) and regardless of whether the reward in question is a token promising money or something else that the participant finds rewarding, such as an erotic image (Prévost et al., 2010; Sescousse et al., 2010). Likewise, a similar effect is seen even when participants not only estimate the value of the options on the basis of their own past experience of taking them but also on the basis of advice from another individual and their knowledge of that person's truthfulness (Behrens et al., 2008).

What Type of Value Signal Does the vmPFC/mOFC Encode?

Considerable emphasis has been placed on the possibility that vmPFC/mOFC and lOFC are relatively more concerned with the representation of positive outcomes, such as rewards, and negative outcomes, such as reward omission or punishment (O'Doherty et al., 2001; Kringelbach, 2005), but this dichotomy appears increasingly untenable. In the case of vmPFC/mOFC (see Comparing vmPFC/mOFC and lOFC and Comparing People and Other Primates) it is now clear that the signal reflects not only expectations of monetary gain but also expectations of monetary loss (Tom et al., 2007; Basten et al., 2010); vmPFC/mOFC BOLD signal decreases in proportion to the value of an anticipated loss (Tom et al., 2007) and with willingness to pay to avoid having to eat an unpleasant food (Plassmann et al., 2010). vmPFC/mOFC BOLD signal also decreases with other factors that diminish the value of rewards, for example, the presence of a delay before the reward is given (Kable and Glimcher, 2007; Prévost et al., 2010).

While there is now a broad consensus that vmPFC/mOFC signals reflect some aspect of both expected reward value prior to the making of a choice and the received reward value after a choice is made (Sescousse et al., 2010; Smith et al., 2010) current research has been directed at addressing several outstanding issues. First, it has been argued that the valuation signal in vmPFC/mOFC is an automatic one that reflects the value of an object even when no choice need be made. Lebreton et al. (2009) reported that vmPFC/mOFC activity reflects participants' preferences for stimuli even when they need not choose between them and instead are asked to make unrelated judgments about the stimuli. In the study by Lebreton et al. (2009) vmPFC/mOFC reflected participants' preferences for face stimuli even while the participants were making judgments about the faces' ages and explicit preference judgments were only made in later stages of the experiment (Figure 3A).

One difficulty for the claim that vmPFC/mOFC signals are automatically generated value signals is that they are not always found in every experiment. For example, the value-related vmPFC/mOFC BOLD signal observed in experiments on willingness to pay (Plassmann et al., 2007, 2010) disappears when subjects are given no option to choose but instead are instructed on which response they should make. Perhaps more of a concern is that when experimental participants watch other individuals playing a public goods game without taking part themselves, the vmPFC/mOFC BOLD signal may reflect aspects of the expected value of choices for the other playing individuals (Cooper et al., 2010). One possibility is that vmPFC/mOFC BOLD signal reflects the value of choices to the individuals to whom the scanned participant's attention is drawn; the vmPFC/mOFC

signal in public good games is larger when the scanned participant's attention is directed to the common good of the group by the experimental instructions. An alternative interpretation, however, might be that the vmPFC signal recorded by Cooper et al. (2010) does actually reflect something about how rewarding the situation is to the subject. The vmPFC/mOFC signal may reflect the fact that the subject is likely to have "other regarding preferences" (Fehr and Camerer, 2007; Behrens et al., 2009) and is therefore unlikely to solely consider his or her own best interests when judging whether an action is rewarding but instead to naturally perceive choices as rewarding when they benefit others.

In contrast to the view that vmPFC/mOFC activity automatically reflects the value of options there is also evidence that vmPFC/mOFC valuation signals reflect a comparison between the values of different options that might be chosen (Boorman et al., 2009; FitzGerald et al., 2009; Basten et al., 2010; Philiastides et al., 2010). In other words, even if value signals in vmPFC/mOFC appear to be automatically generated and present in the absence of choice they are closely tied to the guidance of decisions. For example, Boorman et al. (2009) showed that vmPFC/mOFC BOLD signal is positively correlated with the value of an option that a subject chooses and negatively correlated with the value of an option that a subject rejects (Figures 3Bi and 3Bii). In other studies, however, this value comparison signal is not seen (Wunderlich et al., 2010). Nevertheless, even in the absence of a clear value comparison signal the value of both options is represented at the beginning of the choice period but only the value of the chosen option is represented at later stages in a trial (Wunderlich et al., 2010).

A difficulty for an account of vmPFC/mOFC function emphasizing value comparison and decision-making is its activation in any experiment that does not require a decision from participants (Lebreton et al., 2009; Cooper et al., 2010). Particularly intriguing is the finding that vmPFC/mOFC value comparison signals reflect the value of the chosen option minus the value of the unchosen option (Boorman et al., 2009). In other words, the vmPFC/mOFC signal has a frame of reference that reflects the choice that the participant will ultimately make. One interpretation is that vmPFC/mOFC reflects not only the expected benefit of the course of action taken (in the positive correlation between the BOLD signal and the chosen option value) but also the opportunity costs associated with the unchosen action (in the negative correlation between the BOLD signal and the unchosen option's value). The precise nature of the vmPFC/mOFC signal remains to be elucidated but if it is a decision signal then it is important to note that it differs from a parietal cortical action selection signal (Gold and Shadlen, 2007). While the vmPFC/mOFC signal increases with the difference in value between possible choices the BOLD signal in the parietal cortex and some other motor association areas increases as the choice selection becomes more difficult, as indexed by reaction time. The parietal signal therefore often has characteristics that are the opposite of the vmPFC/mOFC signal; its size is negatively correlated with the difference in value between choices (Basten et al., 2010). Exactly how vmPFC/mOFC and the posterior parietal cortex make different contributions to decision-making remains to be determined.

One rather confusing feature of vmPFC activity in fMRI experiments is that it is often more active at rest than during task performance. The area is close to, or part of, the "default network," a set of brain areas with similar activity (Raichle and Snyder, 2007). "Activations" reported in vmPFC, therefore, actually correspond to different degrees of deactivation, in comparison to rest, in different task conditions. In general, activity in vmPFC decreases monotonically with the level of task engagement, which in turn is a function of a number of task features such as stimulus salience. Salience, however, is often correlated with value; high-value stimuli are often salient. Litt et al. (2011) have recently tried to determine whether the vmPFC BOLD signal is driven by saliency or value. They exploited the fact that salience of a stimulus also increases as it becomes more aversive, and therefore less valuable, as well as when it becomes more appetitive, and therefore more valuable. They examined BOLD activity related to both appetitive and aversive foods so that the impact of value and salience could be separated. vmPFC activity was correlated with value rather than stimulus saliency.

One way to test whether vmPFC/mOFC signals are causally important for guiding decision-making is to investigate what happens when vmPFC/mOFC lesions are made. If vmPFC/mOFC is essential for the value comparison process then a lesion should impair the value discrimination process. If the vmPFC/mOFC is critical for deciding and discriminating between potential choices on the basis of their relative values then the impairment should increase as a function of the proximity of the choices' values. By analogy, lesions of visual feature discrimination mechanisms disrupt discrimination as a function of the similarity in the visual stimuli (Buckley et al., 1997). Noonan et al. (2010) report just such an effect (Figure 4A). The lesions were made in the medial orbitofrontal cortex in macaques, in the region Mackey and Petrides (2010) argue corresponds to the human reward-related vmPFC/mOFC region (Figure 2B).

Comparing vmPFC/mOFC and lOFC and Comparing People and Other Primates

The vmPFC/mOFC reward signal in human fMRI studies is often discussed in the context of neural recordings and lesion studies of the orbital cortex of macaques and rats (Murray et al., 2007). The focus of studies conducted in animals, however, is often on the more accessible lOFC rather than the vmPFC/mOFC itself. Although there is evidence that neurons on the orbital surface of the frontal lobe encode the value of offered and chosen rewards (Tremblay and Schultz, 1999; Padoa-Schioppa and Assad, 2006; Kennerley et al., 2009; Morrison and Salzman, 2009) the majority of recordings are made in the tissue that lies lateral to the medial orbital sulcus in the lOFC. Comparatively little is known of the activity of single neurons in vmPFC. The lOFC has distinct anatomical connections to vmPFC/mOFC that suggest it has access to different types of information and is able to exert different types of influences on the rest of the brain; in other words, its functions are likely to be distinct (Ray and Price, 1993; Carmichael and Price, 1994, 1995a, 1995b, 1996; Ongür et al., 1998; Ferry et al., 2000; Kondo et al., 2003, 2005; Saleem et al., 2008).

One influential idea is that lOFC and vmPFC/mOFC are relatively more concerned with negative and positive outcomes,

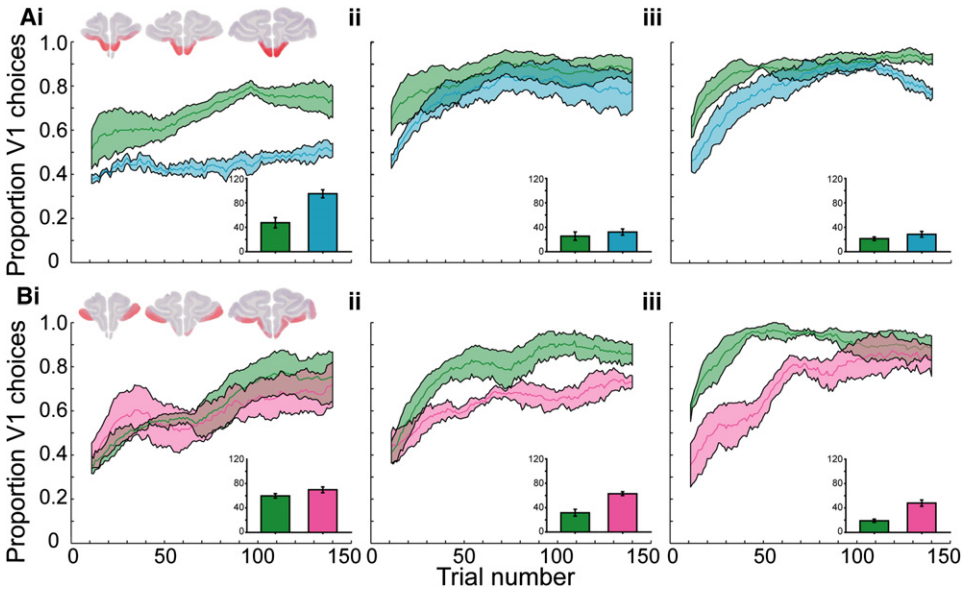

Figure 4. Double Dissociation between Effects of mOFC and lOFC Lesions

(A and B) Effect of choice option value proximity for macaques with mOFC lesions (A) and lOFC lesions (B).

(A) Proportion of choices which were the best value option when the difference in value between the best and second best option was small (i, 0.2), medium (ii, 0.4), and large (iii, 0.6). Shown are control prelesion (green) and post-mOFC (blue) lesion performance (A).

(B) Postoperative lOFC (pink) and matched unoperated control (green) performance. Thick line lines indicate mean proportion of choices of best value option while shadows indicate standard error of the mean. Once again, the different panels illustrate results when the difference in value between the best and second best option was small (i, 0.2), medium (ii, 0.4), and large (iii, 0.6).

Insets show mean number of trials to reach 70% V1 choices (error bars indicate standard error of the mean). mOFC lesions caused impairments when the best and second best value differences were small (Ai) but lOFC lesions impaired performance when the value differences between the best and second best option were larger (Bii and Biii).

(Panel insets) mOFC (top) and lOFC (bottom) lesion locations are represented on an unoperated control brain, with red indicating lesion overlap (mOFC: one to four animals; lOFC: one to three animals). This figure is adapted and reprinted with permission of Noonan et al. (2010) and the National Academy of Sciences, ©2010.

respectively (O'Doherty et al., 2001). There have certainly been frequent replications of the finding that vmPFC/mOFC activity is higher when reward outcomes are received for choices while lOFC activity is higher after punishment or on error trials when potential rewards are not given (Kringelbach and Rolls, 2004). As we have already seen, however, one problem for the reward versus error view of vmPFC versus lOFC is that vmPFC/mOFC appears sufficient to signal both aversive and rewarding value expectations (Tom et al., 2007; Plassmann et al., 2010).

Even more problematic for a view that emphasizes the separation of appetitive and aversive outcomes in OFC is evidence that information about both converges on the same OFC neurons. Morrison and Salzman (2009) reported no anatomical separation within the orbitofrontal area bounded by the medial and lateral orbitofrontal sulci, in neurons that responded to aversive and appetitive outcomes, such as air puffs and juice rewards, respectively. They even found neurons that responded to both types of outcome and that responded to conditioned stimuli predictive of either type of outcome.

The impact of lOFC and vmPFC/mOFC lesions on macaques' sensitivities to error and reward outcomes has also been tested (Noonan et al., 2010). Control macaques normally tended to switch choices more often after errors than after rewards. Both lesions led to higher switch rates after both types of trials—those after reward and those after errors. In other words, there was no evidence that lOFC and vmPFC/mOFC lesions caused relatively greater alterations in error or reward sensitivity.

lOFC lesions do, however, produce the opposite pattern of impairment to vmPFC/mOFC lesions on the value-guided decision task. Again, the impairment is a function of the difference in value of the options (Figure 4B) but while the vmPFC/mOFC lesion-induced impairment increases with the proximity of option values, lOFC lesion-induced impairments do the opposite; impairments increase as value differences between choice options increase and decisions become easier (Noonan et al., 2010) (Figure 4B). vmPFC/mOFC lesions impair performance to a greater degree as the values of the best and second best option are closer and harder to distinguish (Figure 4A) while lOFC lesions cause greater impairments when the decisions are easy and the choice values are very distinct (Figure 4B). While the ability of control animals to identify the best value choice increases with the difference in value between the best and second best value options there is no improvement after lOFC lesions. Such a radically different impairment pattern suggests that lOFC has relatively little role in comparing reward values.

Credit Assignment and the Assignment of Specific Rewards to Specific Stimuli in lOFC

Rather than comparing the values of options lOFC is more concerned with learning about the values of options. The lOFC is

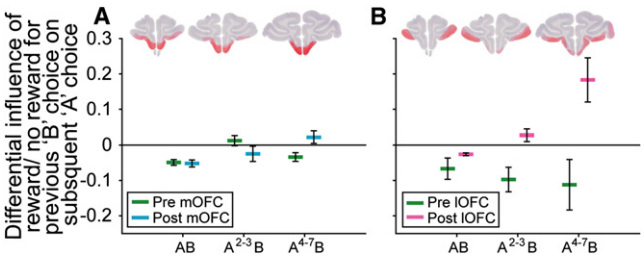

Figure 5. Credit Assignment after Orbitofrontal Lesions

(A and B) Reward-credit assignment during value learning in macaques with mOFC lesions (A) and lOFC lesions (B). Influence of rewards in the current trial on valuation of stimuli chosen in recent trials. Shown is the difference in likelihood of choosing option A on trial n after previously selecting option B on trial n-1 as a function of whether or not a reward was received for this choice. Data are plotted based on the length of choice history on A: (Left) one previous choice of A; (center) two to three previous choices of A; (right) four to seven previous choices of A. After lOFC lesions (B), but not mOFC lesions (A), the credit for the outcome (reward or no reward) received for choosing B on the current trial (n) is partly assigned to stimuli chosen in earlier trials. Error bars indicate standard error of mean performance levels.

(Panel insets) mOFC (left) and lOFC (right) lesion locations are represented on an unoperated control brain, with red indicating lesion overlap (mOFC: one to four animals; lOFC: one to three animals). This figure is adapted and reprinted with permission of Noonan et al. (2010) and the National Academy of Sciences, ©2010, and Walton et al. (2010).

especially important for credit assignment—the process by which visual stimuli are associated with reward values during associative learning (Walton et al., 2010). Normally, monkeys learn to attribute value to a stimulus as a function of the precise history of reward received in association with the choice of that particular stimulus. Animals with lOFC lesions instead value a stimulus as a recency-weighted function of the history of all rewards received approximately at the time of its choice even when the rewards were actually caused by choices of alternative stimuli on preceding and subsequent trials.

Two analyses reveal impairments of credit assignment after lOFC lesions. The first examines the degree to which the recent history of choices made by an animal influences how stimulus-outcome associations are updated when the monkey has just switched to choose a different stimulus. Note that this process of updating the value representation of a new stimulus after a long history of choosing an alternative stimulus mirrors the type of situation found during reversal learning. If credit is assigned correctly, animals should be more likely to repeat the choice of the new stimulus (e.g., stimulus B) on the next trial if its selection was rewarded than if it did not result in reward. Similarly, in this situation, they should be less likely to return to choosing the option they previously were selecting (e.g., stimulus A). Importantly, this effect should be independent of recent choice history (Figure 5). However, this was not the pattern of choices seen in the lOFC-lesioned animals (Figure 5B). Instead, these animals assigned credit for a new outcome based on the integrated recent history of choices, meaning that the outcome for choosing stimulus B is partly assigned to stimulus A. Moreover, the longer the recent history of choices of this other stimulus A, the stronger the influence of an outcome after a new B choice is on the value representation of stimulus A. Indeed, after four to seven consecutive choices of stimulus A, a reward for a new choice of stimulus B makes the reselection of option A on the next trial more likely than if no reward is received for the stimulus B choice. No such effect was seen after vmPFC/mOFC lesions (Noonan et al., 2010) (Figure 5A).

In addition to credit assignment in the lOFC-lesioned animals being affected by their recent choice history, it was also influenced by recent reinforcement history. An option (e.g., stimulus B) was more likely to be reselected if a recent choice of another option (e.g., stimulus A) had been rewarded than if it had not been because the reward for the preceding option A was erroneously assigned to the subsequently chosen option B (Walton et al., 2010). The effect was clearest when the reward for the prior choice of A had been delivered on the previous trial. No evidence of the same impairment was seen after vmPFC/mOFC lesions (Noonan et al., 2010).

The lOFC lesion impairment in credit assignment can explain the otherwise counterintuitive finding that lOFC lesions lead to a failure to improve on "easy" decisions when the reward values of the possible choices are very disparate. While normal animals exploring the stimuli in such easy situations credit each stimulus with its own distinct value, by contrast, the credit assignment impairment leads to animals with lOFC lesions crediting all the stimuli they explore with approximately their mean value. The human lOFC BOLD signal on error trials can also be reinterpreted in the light of the credit assignment hypothesis. It is on just such error trials that subjects are updating the value that should be assigned to an option. The hypothesis, however, also predicts that a similar lOFC signal should be seen when subjects receive positive reinforcement for a choice for the first time because these trials are also ones on which revaluation of an option occurs. Such "first correct" trials are rarely analyzed separately in fMRI experiments; instead they are often lumped together with other trials on which rewards are received. If, however, a subject has considerable experience of consistently receiving reward for a choice then there will be little updating of valuation when yet another reward is received for making the same choice. The hypothesis that lOFC assigns credit for reward to specific stimuli is also consistent with the relatively greater interconnectedness of lOFC versus vmPFC with the temporal and perirhinal cortical areas that represent visual and some other stimuli (Carmichael and Price, 1995b; Kondo et al., 2003, 2005; Saleem et al., 2008).

A similar dichotomy in the effects of vmPFC/mOFC and lOFC lesions in macaques has been reported by Rudebeck and Murray (2011). Rather than testing assignment of credit for rewards to particular stimuli they tested knowledge of how two different types of reward, peanuts and raisins, had been assigned to a large number of stimuli. Macaques learned associations between many arbitrary visual stimuli and either one reward or the other. One of the rewards was then devalued by letting the animals feed on it to satiety; after they are sated on a reward animals prefer the alternative reward. Knowledge of the reward-type-to-stimulus assignment was then tested by giving animals choices between pairs of stimuli, each associated with the two different rewards. Animals with lOFC, but not vmPFC/mOFC lesions, were impaired; they made fewer choices of stimuli to which the unsated reward had been assigned. In another task Rudebeck and Murray (2011) tested the ability to make fine-grained value discriminations by letting the macaques make choices between different pairs drawn from a set of five

stimuli, each associated with different food items. Control animals exhibit consistent, fine-grained differences in the valuations they make of the different stimuli in the consistency and transitivity of their preferences. For example, if a control animal preferred A to B and B to C then it would also be likely to prefer A to C. Such consistent, fine-grained differences in valuations were absent after vmPFC/mOFC lesions.

Recordings of the activity of single neurons in lOFC are also consistent with a role in credit assignment. One way for a credit assignment mechanism to work would be for it to reactivate a representation of the choice that had just been made at the time that the reward was received (and on error trials at the time that the absence of the reward was registered). Tsujimoto et al. (2009) have reported that lOFC neurons do indeed act in this way. Unlike dorsolateral prefrontal neurons, orbitofrontal neurons encode relatively little information about which response is made at the time of the response or during the interval between response and reward. At the time of reward delivery, however, lOFC encodes the choice that led to the delivery of reward. In addition to reactivating choice representations orbitofrontal neurons are also able to maintain representations of particular reward types over a delay even when distracting reward outcomes are presented in the intervening period (Lara et al., 2009). Although neurons in a number of brain areas encode the recent history of rewards and the recent history of choices (Seo and Lee, 2007, 2008) only a few areas, such as lOFC, may encode the conjoint history of which rewards were received for making particular choices.

In summary, it is argued that lOFC is relatively more specialized for assigning credit for both rewards and errors to specific stimulus choices. When different types of reward outcome are available then lOFC represents the assignment of a particular reward type to a particular stimulus. By contrast, it is argued that vmPFC/mOFC value representations are not so much of the specific identify of a reward outcome but of its value and that it is these value representations that determine the goals and choices that primates pursue. The few neuron recording studies that have compared the areas support this interpretation. Rolls (2008) reports that neurons encoding dimensions of reward outcomes, such as taste and texture, are more prevalent in lOFC than vmPFC/mOFC in the macaque. Bouret and Richmond (2010) report that lOFC neurons are more active than vmPFC/mOFC neurons when macaques see visual stimuli that predict rewards. By contrast vmPFC/mOFC neurons have greater access to information about the macaque's current motivational state; the activity of vmPFC/mOFC neurons, but not lOFC neurons, was modulated by satiety (Figure 6).

A very influential observation has been the report of neurons encoding the values of potential choices ("offer-value"-correlated activity) and the values of choices that are actually taken ("chosen-value"-correlated activity) in the lateral bank of the medial orbital sulcus and the adjacent posterior orbitofrontal cortex (Padoa-Schioppa and Assad, 2006, 2008; Padoa-Schioppa, 2009), a region at the transition between vmPFC/mOFC and lOFC divisions (Ongür and Price, 2000). It is tempting to relate the activity of such neurons to human vmPFC/mOFC BOLD signals that reflect the values of available choices and of taken choices (Boorman et al., 2009; FitzGerald et al., 2009; Philiastides et al., 2010; Wunderlich et al., 2010) but it is not clear whether the frequency of such neural patterns changes between vmPFC/mOFC and lOFC.

Counterfactual Choices and aPFC

In many experiments it is assumed that during decision-making people first weigh and compare the values of all of the different options that are available in order to make a choice and second, that these values are learned from the experience of previously choosing these options. Neither of these assumptions may be true. Instead the choice made and the best alternative may each have a special status. Moreover, learning about the value of choices can sometimes occur even without taking the choice if the right feedback is provided. Recent studies of aPFC provide the key evidence for both of these propositions.

The aPFC carries a very distinct signal to the vmPFC. While vmPFC/mOFC encodes the value of the choice that is being made the aPFC encodes information about the value of alternative options that are not chosen (Boorman et al., 2009). While there is a positive relationship between vmPFC/mOFC BOLD signal and the value of chosen options and a negative relationship between vmPFC/mOFC BOLD signal and unchosen options, the relationships are reversed in the aPFC; there is a positive correlation between aPFC BOLD signal and the reward probability associated with an unchosen option but a negative correlation between aPFC BOLD signal and the reward probability associated with the chosen option (Boorman et al., 2009) (compare Figure 7A and Figure 3Bii). One interpretation is that while vmPFC/mOFC encodes the value of the choice that a participant is taking now the aPFC encodes the value of the unchosen option, or what might be referred to as a "counterfactual" choice. It therefore represents the value that switching to an alternative choice might have on a future occasion. Concordant with this notion are findings that aPFC activity reflects the probability of switching on the next trial (Figure 7B) and individual differences in aPFC signal strength are correlated with individual differences in trial-by-trial switching rates (Figure 7B). Koechlin and colleagues have argued that aPFC maintains a representation of a pending state of behavior in which a person might engage in the near future even while a different course of action is actually being followed (Koechlin et al., 1999; Koechlin and Hyafil, 2007).

Not only might aPFC have a role in decision-making but it may also be part of a circuit for learning about the values of counterfactual options. Boorman et al. (2011) gave their participants feedback about counterfactual choices; they indicated to the participants how successful the choice they did not take would have been had it been taken. A counterfactual prediction error signal was seen in the aPFC that paralleled the prediction error signal for chosen options in the ventral striatum. The aPFC signal did not reflect the actual outcome of the choice taken. The exclusive nature of the aPFC prediction error signal therefore makes it distinct from the "fictive" prediction error signal that has been reported in the striatum which reflects the best possible outcome that could have been attained minus the experienced outcome actually received (Lohrenz et al., 2007; Chiu et al., 2008). While the striatal fictive prediction error can only influence subjects by leading them to rechoose the same option, but more of it next time, learning according to the aPFC counterfactual

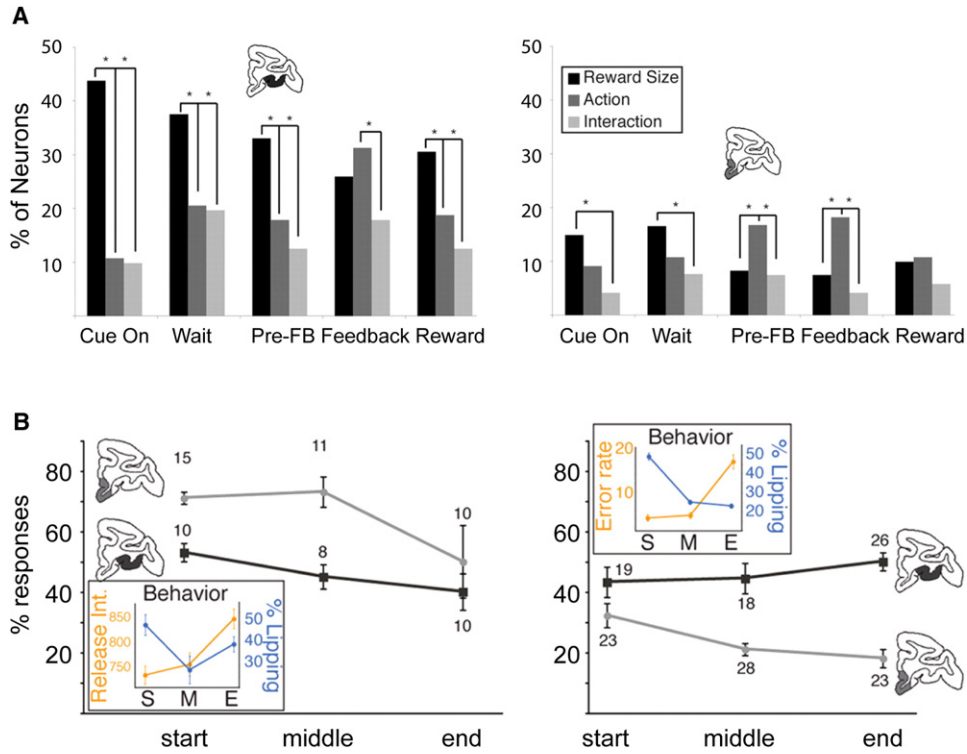

Figure 6. Distinct Patterns of Reward-Related Activity in mOFC and lOFC

(A) Visual stimuli instructed macaque monkeys about whether a manual response was required (active trials) or was not required (passive trials) in order for a reward to be delivered and the size of the reward. lOFC neurons respond to the visual cues associated with reward. Shown is the proportion of responding neurons across epochs of a trial during trials: proportion of neurons responding to reward size (black), action (passive versus active; dark gray), and their interaction (light gray) in each of the five epochs in lOFC (left) and vmPFC/mOFC (right). More lOFC (left) than vmPFC/mOFC (right) neurons responded overall. In lOFC (left), neurons predominantly encode reward size except between the feedback (FB) and reward (Reward) delivery, where the encoding of action peaked. The proportion of neurons encoding action after the feedback was greater than in all the other epochs. In vmPFC more neurons encode action than reward size both before and after the feedback.

(B) The responses of vmPFC/mOFC neurons, but not lOFC neurons, decline with satiety during the course of a session. Response modulates with progression in a session and satiety. Mean percentages (and SEM) of responses for neurons recorded at the very beginning (start), halfway through (middle), and at the end (end) of a recording session for visually cued trials (left) and a distinct set of self-initiated trials (right). Numbers indicate number of neurons at each point. Insets show corresponding behavioral performances (mean and SEM). Monkeys showed a decrease in lipping (blue) and bar-release responses (orange) as they progressed through a session. In self-initiated trials, the proportion of selective neurons decreased in both areas. In cued trials, there was a significant decrease in the proportion of selective neurons in vmPFC, but not in lOFC. S, Start; M, middle; E, end. This figure is adapted and reprinted with permission of Bouret and Richmond (2010) and the Society for Neuroscience.

prediction error can lead subjects to the selection of a completely new course of action.

The aPFC is not the only area in which there is evidence for the encoding of counterfactual prediction errors. In addition there is evidence that they are encoded in a dorsal part of the ACC and in the posterior cingulate cortex (Boorman et al., 2011). The activity of single neurons in the ACC has also been reported to reflect not just rewards that are received but also counterfactual rewards that might have been received for making a different choice (Hayden et al., 2009). The ACC neurons' response rates were predictive of whether animals would switch to a better choice on the next trial.

The aPFC may not encode only the value of an alternative choice but also information about the diversity of alternative options available. Yoshida and Ishii (2006) trained their subjects to navigate through a virtual maze and then took fMRI scans while the subjects tried to work out where they had subsequently been placed within the maze. The BOLD signal in an aPFC region just anterior to that highlighted by Boorman et al. (2009) varied with the subjects' uncertainty about their position in the maze and therefore with the range of alternative options that the subjects might reasonably choose as their next response. In other words, aPFC activity increases when one considers many alternative options as opposed to just a few.

While aPFC may encode the number of possible alternative choices it is important to note that it does not code all options in the same way. The representation of one alternative, the one with the highest value, seems to have a special status. Boorman et al. (2011) tested subjects on a decision-making task with three choices. They found a positive correlation between aPFC BOLD signal and the value of the better alternative choice. There was, however, a negative relationship between aPFC BOLD signal and the value of both the chosen option and the worse alternative. It seems, therefore, that aPFC activity reflects the benefits of switching to the better alternative and the opportunity cost of switching away from both the current choice and the other

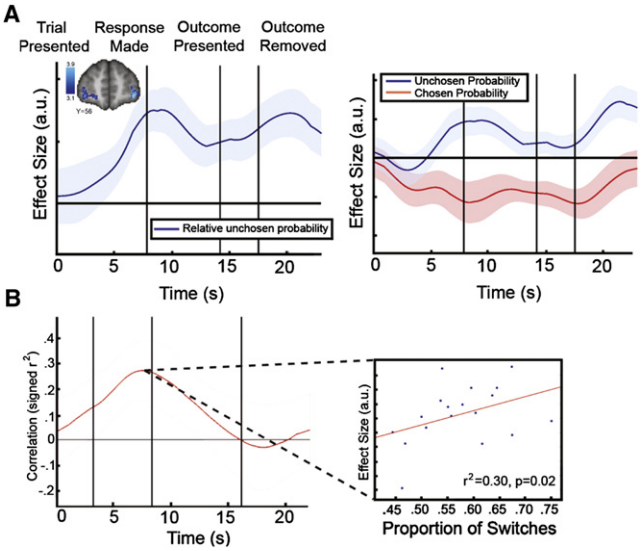

Figure 7. Counterfactual Choice and Switching in the aPFC

(A, inset) Coronal slices through Z statistic maps in aPFC relating to the relative unchosen probability (unchosen option probability − chosen option probability) during decision-making. Color bar indicates Z score.

(A, left panel) Time course for the effect size of the relative unchosen probability in the aPFC is shown throughout the duration of the trial. Right panel: the same time course is shown with the signal decomposed into log unchosen and log chosen action probabilities. There is a positive correlation with log unchosen probability and a negative correlation with log chosen probability. Thick lines: mean effect sizes. Shadows: standard error of the mean (± SEM).

(B) aPFC activity during the ITI predicts within and between-subject variability in behavior. (Left) time series of between-subject correlation between the effect size in aPFC and the proportion of trials in which subjects switched to the better option. (Right) a scatter plot of the effect size (i.e., regression coefficient) for the relative unchosen probability in the aPFC during the ITI for each subject was plotted against the proportion of trials in which a subject switched as a fraction of the trials in which it would have been advantageous to switch. This figure is adapted and reprinted with permission of Boorman et al. (2009).

worse alternative. Such a pattern of activity with just one alternative choice held in a pending state suggests that not all potential alternative choices are considered equally when we change our minds and pursue a different course of action. Behavioral evidence also suggests that we are not able to represent all alternative choices equally and that we switch more effectively to one alternative as opposed to another at any given time (Boorman et al., 2011; Charron and Koechlin, 2010).

The same aPFC region has also been identified in a study of exploratory decision-making (Daw et al., 2006). Although it is obviously advantageous for organisms to choose the most valuable option they can identify it is essential that they also explore alternative options; an organism that fails to explore alternatives will fail to identify choices that might be even higher in value, especially when the environment is changing. There is, therefore, a balance to be struck between exploiting choices of known value and exploring alternatives. Daw et al. (2006) found that vmPFC/mOFC activity was highest when exploitative choices of high-value options were being made but aPFC was more active when lower value, presumably exploratory, choices were made. The results reported by Boorman et al. (2009, 2011) suggest that the high-aPFC signal during exploration reported by Daw et al. (2006) may reflect either a high probability that subjects will switch to yet another alternative or the high value of the options that the participants are already foregoing by exploring.

Comparing aPFC in Humans and Monkeys

The precise anatomical identity of the human aPFC region and its correspondence to regions in other primate species is currently being elucidated. The aPFC region lies either in area 10 in the frontal pole or in a region that Rajkowska and Goldman-Rakic (1995) suggested was a transition zone between area 10 and the dorsolateral prefrontal area 46. The frontal pole is especially large in humans (Semendeferi et al., 2001) and its increase in size is due to its lateral expansion in hominoids into the approximate region in which Boorman et al. (2009, 2011) and Daw et al. (2006) reported fMRI results. Mars et al. (2011) used a combination of diffusion-weighted MRI tractography and examination of the patterns of correlation in the fMRI signals in aPFC and in other brain regions to estimate and compare aPFC's connections in humans and macaques. In the human brain there was evidence of connections linking aPFC to a central region of the inferior parietal lobule (IPL) because the BOLD signals in the two regions were correlated. No similar evidence could be found to link IPL, or indeed any parietal region, and aPFC in macaques. Petrides and Pandya (2007) have also reported no connections between frontal polar area 10 and parietal cortex in the macaque.

One way in which neuroanatomical differences are known to arise during speciation is that parts of areas, perhaps already specialized modules, become spatially separate in some species. The invasion of new connections into an area may also lead to species differences in brain structure and function (Krubitzer, 1995, 2007). It is perhaps not surprising then that in the macaque a similar central IPL region is interconnected to more rostral parts of prefrontal cortex, albeit in area 46 rather than in area 10, than is the case for any other parietal region (Rozzi et al., 2006). In humans, however, the tissue in the aPFC in the transition region between dorsolateral prefrontal cortex and the frontal pole may have coalesced into a distinctive region.

Interactions between the aPFC and the central region of the IPL seem to be especially important at the moment that human participants actually switch from taking one choice to another (Boorman et al., 2009). The signals in the two areas become more highly correlated on switching than in trials in which the same choice is just repeated. It is as if aPFC were able to represent the relative advantage that would accrue from switching choices but it is only through interactions with IPL that the switch is accomplished. Very similar aPFC and central IPL regions are coactive during exploratory choices (Daw et al., 2006).

Despite its prominence in human neuroimaging studies, until recently no recordings had been made of single neuron activity in aPFC area 10 in the monkey. Tsujimoto et al. (2010) have now reported that neurons in this area are active when macaques receive reward feedback after making choices. Moreover, the activity that was recorded at the time of reward delivery also reflected the decision that led to the reward. The activity, however, disappeared when the monkey was not free to make its own decisions but instead was selecting options on the basis of instructional cues. In summary, the pattern of activity is

consistent with area 10 having a role in reward-guided decision-making and possibly in reward-guided learning in macaques as well as in humans. So far, however, there is not clear evidence that area 10 in the macaque is especially concerned with the representation of counterfactual options. It is possible that the recordings were made in too medial a location; as already explained, the human aPFC region implicated in counterfactual choice representation is situated laterally in a transition zone between the frontal pole and the dorsolateral prefrontal cortex. It is also possible that it will be difficult to identify an exact homolog of the human lateral aPFC area in the macaque.

Action-Reward Associations and the ACC

An important aspect of ACC function, supported by research conducted with several techniques including single neuron recording and recording of event-related potentials such as the error-related negativity (ERN) and fMRI, concerns its responsiveness to errors and the initiation of subsequent changes in behavior (Shima and Tanji, 1998; Nieuwenhuis et al., 2004; Jocham and Ullsperger, 2009). It is now clear, however, from both fMRI (Walton et al., 2004) and single neuron recording (Matsumoto et al., 2007; Sallet et al., 2007; Quilodran et al., 2008; Luk and Wallis, 2009) that the most investigated region in the ACC sulcus responds not only to errors but also to rewards in both humans and macaques (cluster 4 in Figure 2A). The critical area is immediately anterior to, and may extend into, the rostral cingulate motor area in areas 24c′ and 24c. In monkeys, it may extend into medial parts of areas 9 and 6 in the dorsal bank of the cingulate sulcus. While some studies suggest that more ACC neurons are responsive to error feedback than to positive, rewarding feedback, the ratio of error-responding to reward-responding cells may be approximately 5:4 (Quilodran et al., 2008) and some neurons respond to both types of feedback. In order to see positive feedback-related activity in ACC in both human fMRI and macaque single neuron recording studies it is, however, critical that the positive feedback is "informative"; that is, it is present when the monkey or person is uncertain which is the correct response to make and is exploring the different possible alternatives (Walton et al., 2004; Quilodran et al., 2008).

Lesion studies also demonstrate that ACC is essential for determining the way in which monkeys respond to rewards (Kennerley et al., 2006). A pressing concern, therefore, is to determine how the ACC's contribution to reward-guided behavior differs from that of the lOFC and vmPFC/mOFC (Rushworth et al., 2007). The emerging picture is that there is some overlap in the function of the ACC and these other areas, perhaps not surprisingly given their anatomical interconnection (Van Hoesen et al., 1993), but that there are also ways in which they differ.

The anatomical connections of ACC provide one important insight into how its function might differ from lOFC. The rostral cingulate motor area is connected to primary motor cortex, several premotor areas, and even to the ventral horn of the spinal cord (Van Hoesen et al., 1993; Morecraft and Tanji, 2009). Such connections mean that it is better placed to influence action selection and to be influenced by action selection, than lOFC. By contrast, ACC has far fewer connections with inferior temporal and perirhinal areas concerned with object recognition than does lOFC (Kondo et al., 2005; Saleem et al., 2008; Yukie and Shibata, 2009). Consistent with these differences in connections, lesion studies in the macaque have shown that ACC and lOFC are relatively more specialized for learning action-reward and stimulus-reward associations (Rudebeck et al., 2008). Ostlund and Balleine (2007) have reported a possibly similar relative specialization for learning action-reward and stimulus-reward associations in a medial frontal cortex area, the prelimbic cortex, and in the rat's OFC. Neurophysiological studies have also shown that ACC neurons have response properties that would allow them to associate actions with rewards. Hayden and Platt (2010) report that ACC neurons that are reward sensitive are also tuned for the direction of saccades at the time that the saccades are made and reward is received even if they are not tuned in this way at earlier times during motor planning. Kennerley et al. (2009) reported a greater number of response-selective neurons in ACC than in OFC when both areas were investigated in the same paradigm in the same individual monkeys.

Exactly how vmPFC/mOFC and ACC interact during reward-guided decision-making remains unclear. The two regions are anatomically interconnected (Van Hoesen et al., 1993; Morecraft and Tanji, 2009). Moreover, vmPFC/mOFC activity reflects the expected value of a choice whether the choice is made between stimuli or actions (Gläscher et al., 2009; Wunderlich et al., 2010). One possibility is that while vmPFC/mOFC determines the reward goal that is to be pursued the ACC is particularly concerned with the association between reward and action and the determination of the action that is to be made to obtain the goal. In many experiments the process of choosing a reward goal is confounded with the choice of an action to achieve the goal but these two aspects of selection can be separated.

Another possibility is that ACC is encoding a parameter related to the rate at which reward is being received per response. It is known that a number of animals including primates are sensitive to the rate at which rewards are being harvested and that they move and start foraging elsewhere once the rate of reward falls below the average for the animal's broader environment (Charnov, 1976; Agetsuma, 1999). The failure of animals with ACC lesions to identify the response with the better reward yield (Kennerley et al., 2006) could be interpreted as the consequence of an impairment in a mechanism for encoding the reward rate associated with a response or an impairment in the use of such information to decide whether or not to try switching to making an alternative response. A related idea, discussed in more detail below, is the possibility that ACC encodes certain types of costs, as well as the benefits, that are associated with a choice.

The exact nature of the response that will be made at the end of the decision does appear to be important for ACC neurons. Kennerley et al. (2009) trained each of their two macaques to respond in different modalities with either eye movements or arm movements. Response selectivity was more apparent in ACC in the second case. The difference in selectivity might reflect the precise placement of the recording electrodes with respect to regions of cortex specialized for representing one type of response or the other (Wang et al., 2004; Amiez and Petrides, 2009). Alternatively, however, it may reflect the fact that the nature of the response itself may impact on the value of a

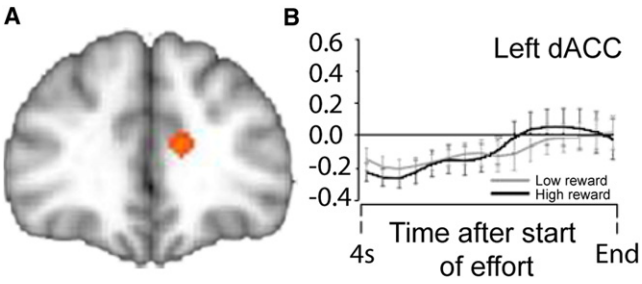

Figure 8. ACC Activity Changes with Both Reward and Effort Expectation

(A) Increased BOLD signal identified by the net value contrast in left ACCd. (B) Activity during the effort investment period on high-effort trials in the left ACCd. Black and gray continuous lines indicate high-effort trials with high- and low-reward expectation, respectively. The vertical lines indicate SEM. The effort period varied in length from subject to subject as did the degree of effort required on each trial, but here all data have been normalized to the mean length of the effort period. The baseline is an implicit baseline representing the unexplained variance in each subject's time series. The interval before the start of the effort period was jittered so this activity was not confounded with cue-related activity. This figure is adapted and reprinted with permission of Croxson et al. (2009) and the Society for Neuroscience.

course of action; if a course of action is difficult to execute or effortful then the costs of pursuing that course of action may need to be weighed against the potential benefits before a choice is made.

Cost-Benefit Decision-Making

A second and related dimension of difference between ACC and OFC concerns the way in which the areas encode the costs, in addition to the benefits, of a choice. Rangel and Hare (2010) argue that there is an important difference between costs that are tied to the outcome itself and costs that are tied to the action that is used to obtain the outcome. The first type of cost might include an aversive outcome that occurs at the same time as an appetitive outcome or the delay that elapses before the reward arrives. The second type of cost might include the effort that has to be expended in order to perform the action that is needed to obtain a reward. lOFC and vmPFC/mOFC are more concerned with the first type of cost and the ACC is more concerned with the second type of cost.

In the rat OFC lesions lead to impulsive decision-making and an impaired ability to wait for a longer time in order to receive a larger reward (Rudebeck et al., 2006). By contrast ACC lesions lead to apathetic patterns of decision-making such that a rat is no longer prepared to invest effort in taking a course of action in order to obtain a larger reward (Walton et al., 2002, 2003; Rudebeck et al., 2006).

The effort versus delay cost distinction has also proved useful for understanding differences between primate vmPFC/mOFC and ACC. Prévost et al. (2010) have reported that human cost-benefit decision-making behavior is influenced in a similar way by both delay and effort costs but that the two types of decisions are associated with different neural structures, vmPFC/mOFC and ACC, respectively. The BOLD response in the vmPFC/mOFC is positively correlated with the temporally discounted subjective reward expectation (Kable and Glimcher, 2007; Prévost et al., 2010). Prevost et al. (2010) argue that vmPFC/mOFC does not encode the effort to be expended in reaching the reward. Croxson et al. (2009) have also reported the existence of a more lateral posterior OFC region that is sensitive to expectations about reward magnitude but which does not carry information about the effort to be exerted before a reward is received.

Exactly how ACC encodes effort remains uncertain. Although both Croxson et al. (2009) and Prévost et al. (2010) report that ACC activity reflects both anticipated effort and anticipated reward there are differences between the patterns of modulation seen in the two studies. The differences may reflect the degree to which cueing of effort expectations and actual effort exertion are separated in time. When the cue that indicates the reward and effort expectations is separated in time from the period when the response is made and effort is actually exerted then different BOLD signals at the two times can be identified (Croxson et al., 2009). At the time that an instruction cue is presented the ACC signal reflects the interaction of both reward and effort expectations; the ACC is most active in anticipation of high rewards to be obtained with the least effort. As the participant begins to engage in the "effort period" and makes a series of movements, the ACC signal increases as the reward approaches (Figure 8).

Comparisons have also been made of single neuron activity in the ACC and OFC when monkeys are presented with cues instructing reward and effort expectations (Kennerley et al., 2009) and as they move through a sequence of responses toward rewards (Shidara and Richmond, 2002; Simmons and Richmond, 2008). In the experiment conducted by Kennerley et al. (2009) animals chose between two cues with learned associations with expected reward payoff size, probability of reward delivery, and effort (expected number of lever presses). In both ACC and lOFC, neurons were equally sensitive to each facet of value. Single ACC neurons, however, were significantly more likely to encode all three aspects of value. In other words, the activity of single neurons in the ACC integrates information about the effort costs and the reward benefits of actions and does not distinguish what aspect of a choice makes it valuable (Figure 9).

The responses of ACC and OFC neurons have also been recorded in a task that required monkeys to respond with a series of bar presses to a series of cues with a gradually changing appearance that indicated approach toward a reward (Shidara and Richmond, 2002; Simmons and Richmond, 2008). One-third of the task-sensitive neurons in ACC tracked the monkey's position in the response schedule and had firing rates that changed monotonically as the monkeys approached the reward. The activity disappeared when information about the animal's position in the sequence of responses, and therefore information about how many more responses had to be made before reward could be expected, was absent. By contrast, only 3% of OFC neurons exhibited activity that was monotonically dependent on position in the response sequence. OFC neurons were not simply unresponsive to the task but also had activity that reflected other features of the task, for example, whether or not a reward had just been delivered on the last trial or whether or not reward was expected on the current trial. The OFC sensitivity to recent reward delivery was independent of whether or not there was information available to inform the animal about the schedule of responses to be made. In summary, while OFC

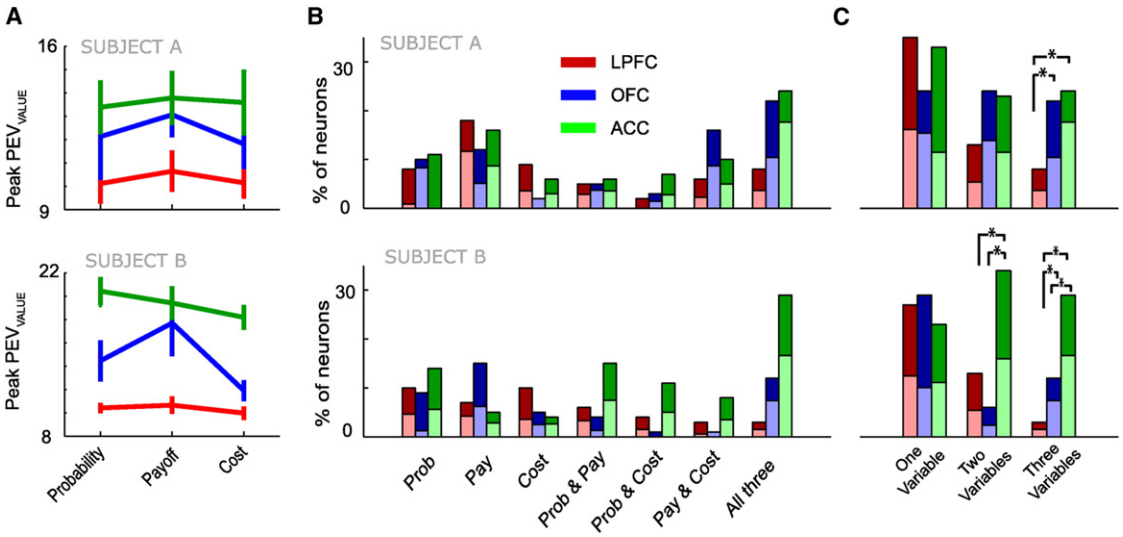

Figure 9. Distinct Patterns of Value Encoding in AC and Orbitofrontal Cortex

(A) Macaque monkeys made choices between stimuli on the basis of their reward probability, reward payoff size, and effort, by a number of joystick movements that had to be made before a reward was given. Shown is the mean (± standard error) of the peak selectivity (PEV_{VALUE}, percentage of the total variance in the neuronal activity that the chosen option's value explained) reached during the choice epoch for every neuron that coded some aspect of value ($p < 0.001$) for three consecutive data time bins.

(B) Percentage of all neurons that encode the value of the chosen option depending on which decision variable (reward probability, reward payoff size, or effort) was being manipulated. The light shading indicates the proportion of neurons that increased their firing rate as value increased, whereas the dark shading indicates those that increased their firing rate as value decreased.

(C) Percentage of neurons that encode value across one, two, or three decision variables. Asterisks indicate the proportions that are significantly different from one another (χ^2 test, $p < 0.05$). This figure is adapted and reprinted with permission of Kennerley et al. (2009) and MIT Press, ©2008 Massachusetts Institute of Technology.

activity might facilitate the relative comparison of rewards ACC activity reflects the integration of both future reward and effort expectations.

In a task that involves the execution of a plan involving a sequence of actions directed toward a distant reward it is clear that integrated value neurons in the ACC will have particular importance. In addition to value selectivity, these cells either themselves possess response selectivity or are adjacent to and interconnected with neurons that have response selectivity. The fact that OFC neurons represent different facets of choice values independently, such as effort, reward probability, and reward payoff size (Kennerley et al., 2009), as well as the taste and texture of rewards (Rolls, 2008), may mean that the OFC is especially important when the context means that only certain features of the reward are relevant for the decision in hand. The OFC's representation of reward value may also be important when a particular feature of the reward is changing in value, for example, when a reward's value changes as a result of satiety (Murray and Izquierdo, 2007; Murray and Wise, 2010).

Finer Grained Anatomical Divisions with Frontal Cortical Systems for Reward-Guided Behavior

The aim of this review has been to sketch emerging evidence concerning the role of frontal cortical areas in reward-guided learning and decision-making. Dividing frontal cortex into four regions facilitates comparisons between studies conducted with monkeys and humans and makes it possible to review distinct hypotheses about frontal functions. It is not, of course, meant to imply that there are not interactions between the areas; nor is it meant to suggest that they do not operate in tandem in many real-world situations. Equally it is important to acknowledge that further and more fine-grained functional subdivisions are emerging within these four major regions.

For example, within ACC there is evidence that two regions can be distinguished that encode how informative feedback is for learning about the value of one's own choices and about other individuals (Behrens et al., 2007, 2008). The value of a choice should be updated when the choice is made and the reward received is better or worse than anticipated. An organism might revise its estimate of the choice's value by a small or a considerable degree each time it witnesses a prediction error. The optimal degree of value updating, however, ought to be a function of the speed with which the reward environment is changing. If the reward environment is volatile and changing rapidly then it makes sense to update valuations substantially as each prediction error is observed. By contrast, in a more stable environment, dramatic revaluation with each prediction error is less optimal and it is preferable to base estimates of an action's value on a longer history of reward events. The impact that volatility has on action valuation is associated with activity changes in the ACC sulcus region implicated in reward-guided action selection (Behrens et al., 2007, 2008; Jocham et al., 2009) (cluster 4, Figure 2A). By contrast, the impact that volatility has on evaluation of other people is associated with changes in the adjacent ACC gyrus (Behrens et al., 2008) (cluster 7, Figure 2A). Pharmacological manipulations that alter the importance ascribed to other individuals activate a similar ACC gyral region (Baumgartner et al., 2008). Information about the value of one's own actions and information about the

value of information from other individuals may be brought together in adjacent ACC regions because both types of information are often important guides to what choices we should make next. There is also evidence that other parts of the ACC are concerned with the control of autonomic activity in the body; different regions within the ACC may be concerned with different aspects of autonomic control or autonomic activity in different body regions (Critchley, 2005). Although a discussion of autonomic control is beyond the scope of the current review it is important to note that autonomic changes may be instigated during reward-guided decision-making and autonomic feedback may contribute to the appraisal of a choice.

The vmPFC/mOFC region includes a variety of distinct if interconnected anatomical areas and it is likely that they make distinct contributions to valuation. Localizing BOLD signal changes in this region is difficult because of the proximity of the sinuses but nevertheless there is already emerging evidence of regional differences in function. Grabenhorst et al., 2008 asked their subjects either to rate the pleasantness of temperature stimuli or to make a decision about whether the stimulus should be repeated. They argued that the responses in the most posterior vmPFC, which was probably in perigenual cingulate cortex, reflected value on a continuous scale while activity in more anterior vmPFC (cluster 2, Figure 2A and area 14 m, Figure 2B) was more closely related to the making of a binary choice about whether or not to repeat the stimulus. Yet another vmPFC/mOFC region, area 25 in the subcallosal region (cluster 1, Figure 2A), may track the value that is ascribed to oneself; activity in this region is altered in depression (Murray et al., 2010) and correlates with mood changes induced by inflammation after infection (Harrison et al., 2009). In other words, major challenges to a person's evaluation of themselves and their own value and their sense of well-being are associated with changes in area 25. Information about the value currently assigned to oneself and about the value of one's prospects and decisions may be brought together in adjacent vmPFC regions in order to provide the best estimate of the organism's value in the future.

Conclusions

Although investigations of reward-guided decision-making in the primate have often focused on human vmPFC/mOFC and on macaque lOFC it is becoming increasingly clear that there are important differences in the functions of these areas and other areas such as ACC and aPFC. Relatively little is known of activity at neuronal level in some of these areas, including vmPFC/mOFC and aPFC. Future progress is likely to depend not only on more refined descriptions of behavior and more detailed descriptions of neurophysiology, but also on an increasing knowledge of the interactions of the various frontal lobe areas with one another and with other brain regions (Schoenbaum et al., 2009).

ACKNOWLEDGMENTS

This research was supported by MRC and Wellcome Trust.

REFERENCES

Agetsuma, N. (1999). Simulation of patch use by monkeys using operant conditioning. J. Ethol. *16*, 49–55.

Amiez, C., and Petrides, M. (2009). Anatomical organization of the eye fields in the human and non-human primate frontal cortex. Prog. Neurobiol. *89*, 220–230.

Basten, U., Biele, G., Heekeren, H.R., and Fiebach, C.J. (2010). How the brain integrates costs and benefits during decision making. Proc. Natl. Acad. Sci. USA *107*, 21767–21772.

Baumgartner, T., Heinrichs, M., Vonlanthen, A., Fischbacher, U., and Fehr, E. (2008). Oxytocin shapes the neural circuitry of trust and trust adaptation in humans. Neuron *58*, 639–650.

Becker, G.M., DeGroot, M.H., and Marschak, J. (1964). Measuring utility by a single-response sequential method. Behav. Sci. *9*, 226–232.

Beckmann, M., Johansen-Berg, H., and Rushworth, M.F. (2009). Connectivity-based parcellation of human cingulate cortex and its relation to functional specialization. J. Neurosci. *29*, 1175–1190.

Behrens, T.E., Woolrich, M.W., Walton, M.E., and Rushworth, M.F. (2007). Learning the value of information in an uncertain world. Nat. Neurosci. *10*, 1214–1221.

Behrens, T.E., Hunt, L.T., Woolrich, M.W., and Rushworth, M.F. (2008). Associative learning of social value. Nature *456*, 245–249.

Behrens, T.E., Hunt, L.T., and Rushworth, M.F. (2009). The computation of social behavior. Science *324*, 1160–1164.

Boorman, E.D., Behrens, T.E., Woolrich, M.W., and Rushworth, M.F. (2009). How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. Neuron *62*, 733–743.

Boorman, E.D., Behrens, T.E., and Rushworth, M.F.S. (2011). Counterfactual choice and learning in a neural network centered on human lateral frontopolar cortex. PLoS Biology, in press.

Bouret, S., and Richmond, B.J. (2010). Ventromedial and orbital prefrontal neurons differentially encode internally and externally driven motivational values in monkeys. J. Neurosci. *30*, 8591–8601.

Buckley, M.J., Gaffan, D., and Murray, E.A. (1997). Functional double dissociation between two inferior temporal cortical areas: perirhinal cortex versus middle temporal gyrus. J. Neurophysiol. *77*, 587–598.

Carmichael, S.T., and Price, J.L. (1994). Architectonic subdivision of the orbital and medial prefrontal cortex in the macaque monkey. J. Comp. Neurol. *346*, 366–402.

Carmichael, S.T., and Price, J.L. (1995a). Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. J. Comp. Neurol. *363*, 615–641.

Carmichael, S.T., and Price, J.L. (1995b). Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. J. Comp. Neurol. *363*, 642–664.

Carmichael, S.T., and Price, J.L. (1996). Connectional networks within the orbital and medial prefrontal cortex of macaque monkeys. J. Comp. Neurol. *371*, 179–207.

Charnov, E.L. (1976). Optimal foraging, the marginal value theorem. Theor. Popul. Biol. *9*, 129–136.

Charron, S., and Koechlin, E. (2010). Divided representation of concurrent goals in the human frontal lobes. Science *328*, 360–363.

Chiu, P.H., Lohrenz, T.M., and Montague, P.R. (2008). Smokers' brains compute, but ignore, a fictive error signal in a sequential investment task. Nat. Neurosci. *11*, 514–520.

Cooper, J.C., Kreps, T.A., Wiebe, T., Pirkl, T., and Knutson, B. (2010). When giving is good: ventromedial prefrontal cortex activation for others' intentions. Neuron *67*, 511–521.

Critchley, H.D. (2005). Neural mechanisms of autonomic, affective, and cognitive integration. J. Comp. Neurol. *493*, 154–166.

Croxson, P.L., Walton, M.E., O'Reilly, J.X., Behrens, T.E., and Rushworth, M.F. (2009). Effort-based cost-benefit valuation and the human brain. J. Neurosci. *29*, 4531–4541.

Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B., and Dolan, R.J. (2006). Cortical substrates for exploratory decisions in humans. Nature *441*, 876–879.

Doya, K. (2008). Modulators of decision making. Nat. Neurosci. *11*, 410–416.

Fehr, E., and Camerer, C.F. (2007). Social neuroeconomics: the neural circuitry of social preferences. Trends Cogn. Sci. (Regul. Ed.) *11*, 419–427.

Ferry, A.T., Ongür, D., An, X., and Price, J.L. (2000). Prefrontal cortical projections to the striatum in macaque monkeys: evidence for an organization related to prefrontal networks. J. Comp. Neurol. *425*, 447–470.

FitzGerald, T.H., Seymour, B., and Dolan, R.J. (2009). The role of human orbitofrontal cortex in value comparison for incommensurable objects. J. Neurosci. *29*, 8388–8395.

Gläscher, J., Hampton, A.N., and O'Doherty, J.P. (2009). Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. Cereb. Cortex *19*, 483–495.

Glimcher, P. (2002). Decisions, decisions, decisions: choosing a biological science of choice. Neuron *36*, 323–332.

Glimcher, P., Camerer, C., and Fehr, E.R.A.P. (2009). Neuroeconomics: Decision Making and the Brain (New York: Academic Press).

Gold, J.I., and Shadlen, M.N. (2007). The neural basis of decision making. Annu. Rev. Neurosci. *30*, 535–574.

Grabenhorst, F., Rolls, E.T., and Parris, B.A. (2008). From affective value to decision-making in the prefrontal cortex. Eur. J. Neurosci. *28*, 1930–1939.

Harrison, N.A., Brydon, L., Walker, C., Gray, M.A., Steptoe, A., and Critchley, H.D. (2009). Inflammation causes mood changes through alterations in subgenual cingulate activity and mesolimbic connectivity. Biol. Psychiatry *66*, 407–414.

Hayden, B.Y., and Platt, M.L. (2010). Neurons in anterior cingulate cortex multiplex information about reward and action. J. Neurosci. *30*, 3339–3346.

Hayden, B.Y., Pearson, J.M., and Platt, M.L. (2009). Fictive reward signals in the anterior cingulate cortex. Science *324*, 948–950.

Jocham, G., and Ullsperger, M. (2009). Neuropharmacology of performance monitoring. Neurosci. Biobehav. Rev. *33*, 48–60.

Jocham, G., Neumann, J., Klein, T.A., Danielmeier, C., and Ullsperger, M. (2009). Adaptive coding of action values in the human rostral cingulate zone. J. Neurosci. *29*, 7489–7496.

Kable, J.W., and Glimcher, P.W. (2007). The neural correlates of subjective value during intertemporal choice. Nat. Neurosci. *10*, 1625–1633.

Kennerley, S.W., Walton, M.E., Behrens, T.E., Buckley, M.J., and Rushworth, M.F. (2006). Optimal decision making and the anterior cingulate cortex. Nat. Neurosci. *9*, 940–947.

Kennerley, S.W., Dahmubed, A.F., Lara, A.H., and Wallis, J.D. (2009). Neurons in the frontal lobe encode the value of multiple decision variables. J. Cogn. Neurosci. *21*, 1162–1178.

Koechlin, E., and Hyafil, A. (2007). Anterior prefrontal function and the limits of human decision-making. Science *318*, 594–598.

Koechlin, E., Basso, G., Pietrini, P., Panzer, S., and Grafman, J. (1999). The role of the anterior prefrontal cortex in human cognition. Nature *399*, 148–151.

Kondo, H., Saleem, K.S., and Price, J.L. (2003). Differential connections of the temporal pole with the orbital and medial prefrontal networks in macaque monkeys. J. Comp. Neurol. *465*, 499–523.

Kondo, H., Saleem, K.S., and Price, J.L. (2005). Differential connections of the perirhinal and parahippocampal cortex with the orbital and medial prefrontal networks in macaque monkeys. J. Comp. Neurol. *493*, 479–509.

Kringelbach, M.L. (2005). The human orbitofrontal cortex: linking reward to hedonic experience. Nat. Rev. Neurosci. *6*, 691–702.

Kringelbach, M.L., and Rolls, E.T. (2004). The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. Prog. Neurobiol. *72*, 341–372.

Krubitzer, L. (1995). The organization of neocortex in mammals: are species differences really so different? Trends Neurosci. *18*, 408–417.

Krubitzer, L. (2007). The magnificent compromise: cortical field evolution in mammals. Neuron *56*, 201–208.

Lara, A.H., Kennerley, S.W., and Wallis, J.D. (2009). Encoding of gustatory working memory by orbitofrontal neurons. J. Neurosci. *29*, 765–774.

Lebreton, M., Jorge, S., Michel, V., Thirion, B., and Pessiglione, M. (2009). An automatic valuation system in the human brain: evidence from functional neuroimaging. Neuron *64*, 431–439.

Lee, D. (2008). Game theory and neural basis of social decision making. Nat. Neurosci. *11*, 404–409.

Litt, A., Plassmann, H., Shiv, B., and Rangel, A. (2011). Dissociating valuation and saliency signals during decision-making. Cereb. Cortex *21*, 95–102.

Lohrenz, T., McCabe, K., Camerer, C.F., and Montague, P.R. (2007). Neural signature of fictive learning signals in a sequential investment task. Proc. Natl. Acad. Sci. USA *104*, 9493–9498.

Luk, C.H., and Wallis, J.D. (2009). Dynamic encoding of responses and outcomes by neurons in medial prefrontal cortex. J. Neurosci. *29*, 7526–7539.

Mackey, S., and Petrides, M. (2010). Quantitative demonstration of comparable architectonic areas within the ventromedial and lateral orbital frontal cortex in the human and the macaque monkey brains. Eur. J. Neurosci. *32*, 1940–1950.

Mars, R.B., Jbabdi, S., Sallet, J., O'Reilly, J.X., Croxson, P.L., Olivier, E., Noonan, M.P., Bergmann, C., Mitchell, A.S., Baxter, M.G., et al. (2011). Diffusion-weighted imaging tractography-based parcellation of the human parietal cortex and comparison with human and macaque resting-state functional connectivity. J. Neurosci. *31*, 4087–4100.

Matsumoto, M., Matsumoto, K., Abe, H., and Tanaka, K. (2007). Medial prefrontal cell activity signaling prediction errors of action values. Nat. Neurosci. *10*, 647–656.

Morecraft, R., and Tanji, J. (2009). Cingulofrontal interactions and the cingulate motor areas. In Cingulate neurobiology and disease, B. Vogt, ed. (New York: Oxford University Press).

Morrison, S.E., and Salzman, C.D. (2009). The convergence of information about rewarding and aversive stimuli in single neurons. J. Neurosci. *29*, 11471–11483.

Murray, E.A., and Izquierdo, A. (2007). Orbitofrontal cortex and amygdala contributions to affect and action in primates. Ann. N Y Acad. Sci. *1121*, 273–296.

Murray, E.A., and Wise, S.P. (2010). Interactions between orbital prefrontal cortex and amygdala: advanced cognition, learned responses and instinctive behaviors. Curr. Opin. Neurobiol. *20*, 212–220.

Murray, E.A., O'Doherty, J.P., and Schoenbaum, G. (2007). What we know and do not know about the functions of the orbitofrontal cortex after 20 years of cross-species studies. J. Neurosci. *27*, 8166–8169.

Murray, E.A., Wise, S.P., and Drevets, W.C. (2010). Localization of dysfunction in major depressive disorder: prefrontal cortex and amygdala. Biol. Psychiatry, in press. Published online November 24, 2010. 10.1016/j.biopsych.2010.09.041.

Nieuwenhuis, S., Holroyd, C.B., Mol, N., and Coles, M.G. (2004). Reinforcement-related brain potentials from medial frontal cortex: origins and functional significance. Neurosci. Biobehav. Rev. *28*, 441–448.

Noonan, M.P., Walton, M.E., Behrens, T.E., Sallet, J., Buckley, M.J., and Rushworth, M.F. (2010). Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. Proc. Natl. Acad. Sci. USA *107*, 20547–20552.

O'Doherty, J., Kringelbach, M.L., Rolls, E.T., Hornak, J., and Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. Nat. Neurosci. *4*, 95–102.

Ongür, D., and Price, J.L. (2000). The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. Cereb. Cortex *10*, 206–219.

Ongür, D., An, X., and Price, J.L. (1998). Prefrontal cortical projections to the hypothalamus in macaque monkeys. J. Comp. Neurol. *401*, 480–505.

Ongür, D., Ferry, A.T., and Price, J.L. (2003). Architectonic subdivision of the human orbital and medial prefrontal cortex. J. Comp. Neurol. *460*, 425–449.

Ostlund, S.B., and Balleine, B.W. (2007). Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental conditioning. J. Neurosci. *27*, 4819–4825.

Padoa-Schioppa, C. (2009). Range-adapting representation of economic value in the orbitofrontal cortex. J. Neurosci. *29*, 14004–14014.

Padoa-Schioppa, C., and Assad, J.A. (2006). Neurons in the orbitofrontal cortex encode economic value. Nature *441*, 223–226.

Padoa-Schioppa, C., and Assad, J.A. (2008). The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. Nat. Neurosci. *11*, 95–102.

Petrides, M., and Pandya, D.N. (2007). Efferent association pathways from the rostral prefrontal cortex in the macaque monkey. J. Neurosci. *27*, 11573–11586.

Philiastides, M.G., Biele, G., and Heekeren, H.R. (2010). A mechanistic account of value computation in the human brain. Proc. Natl. Acad. Sci. USA *107*, 9430–9435.

Plassmann, H., O'Doherty, J., and Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. J. Neurosci. *27*, 9984–9988.

Plassmann, H., O'Doherty, J.P., and Rangel, A. (2010). Appetitive and aversive goal values are encoded in the medial orbitofrontal cortex at the time of decision making. J. Neurosci. *30*, 10799–10808.

Platt, M.L., and Glimcher, P.W. (1999). Neural correlates of decision variables in parietal cortex. Nature *400*, 233–238.

Platt, M.L., and Huettel, S.A. (2008). Risky business: the neuroeconomics of decision making under uncertainty. Nat. Neurosci. *11*, 398–403.

Prévost, C., Pessiglione, M., Météreau, E., Cléry-Melin, M.L., and Dreher, J.C. (2010). Separate valuation subsystems for delay and effort decision costs. J. Neurosci. *30*, 14080–14090.

Quilodran, R., Rothé, M., and Procyk, E. (2008). Behavioral shifts and action valuation in the anterior cingulate cortex. Neuron *57*, 314–325.

Raichle, M.E., and Snyder, A.Z. (2007). A default mode of brain function: a brief history of an evolving idea. Neuroimage *37*, 1083–1090, discussion 1097–1099.

Rajkowska, G., and Goldman-Rakic, P.S. (1995). Cytoarchitectonic definition of prefrontal areas in the normal human cortex: II. Variability in locations of areas 9 and 46 and relationship to the Talairach Coordinate System. Cereb. Cortex *5*, 323–337.

Rangel, A., and Hare, T. (2010). Neural computations associated with goal-directed choice. Curr. Opin. Neurobiol. *20*, 262–270.

Rangel, A., Camerer, C., and Montague, P.R. (2008). A framework for studying the neurobiology of value-based decision making. Nat. Rev. Neurosci. *9*, 545–556.

Ray, J.P., and Price, J.L. (1993). The organization of projections from the mediodorsal nucleus of the thalamus to orbital and medial prefrontal cortex in macaque monkeys. J. Comp. Neurol. *337*, 1–31.

Rolls, E.T. (2008). Functions of the orbitofrontal and pregenual cingulate cortex in taste, olfaction, appetite and emotion. Acta Physiol. Hung. *95*, 131–164.

Rozzi, S., Calzavara, R., Belmalih, A., Borra, E., Gregoriou, G.G., Matelli, M., and Luppino, G. (2006). Cortical connections of the inferior parietal cortical convexity of the macaque monkey. Cereb. Cortex *16*, 1389–1417.

Rudebeck, P.H., and Murray, E.A. (2011). Dissociable effects of subtotal lesions within the macaque orbital prefrontal cortex on reward-guided behavior. J. Neurosci., in press.

Rudebeck, P.H., Walton, M.E., Smyth, A.N., Bannerman, D.M., and Rushworth, M.F. (2006). Separate neural pathways process different decision costs. Nat. Neurosci. *9*, 1161–1168.

Rudebeck, P.H., Behrens, T.E., Kennerley, S.W., Baxter, M.G., Buckley, M.J., Walton, M.E., and Rushworth, M.F. (2008). Frontal cortex subregions play distinct roles in choices between actions and stimuli. J. Neurosci. *28*, 13775–13785.

Rushworth, M.F., Behrens, T.E., Rudebeck, P.H., and Walton, M.E. (2007). Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour. Trends Cogn. Sci. (Regul. Ed.) *11*, 168–176.

Saleem, K.S., Kondo, H., and Price, J.L. (2008). Complementary circuits connecting the orbital and medial prefrontal networks with the temporal, insular, and opercular cortex in the macaque monkey. J. Comp. Neurol. *506*, 659–693.

Sallet, J., Quilodran, R., Rothé, M., Vezoli, J., Joseph, J.P., and Procyk, E. (2007). Expectations, gains, and losses in the anterior cingulate cortex. Cogn. Affect. Behav. Neurosci. *7*, 327–336.

Schoenbaum, G., Roesch, M.R., Stalnaker, T.A., and Takahashi, Y.K. (2009). A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. Nat. Rev. Neurosci. *10*, 885–892.

Semendeferi, K., Armstrong, E., Schleicher, A., Zilles, K., and Van Hoesen, G.W. (2001). Prefrontal cortex in humans and apes: a comparative study of area 10. Am. J. Phys. Anthropol. *114*, 224–241.

Seo, H., and Lee, D. (2007). Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. J. Neurosci. *27*, 8366–8377.

Seo, H., and Lee, D. (2008). Cortical mechanisms for reinforcement learning in competitive games. Philos. Trans. R. Soc. Lond. B Biol. Sci. *363*, 3845–3857.

Sescousse, G., Redouté, J., and Dreher, J.C. (2010). The architecture of reward value coding in the human orbitofrontal cortex. J. Neurosci. *30*, 13095–13104.

Shidara, M., and Richmond, B.J. (2002). Anterior cingulate: single neuronal signals related to degree of reward expectancy. Science *296*, 1709–1711.

Shima, K., and Tanji, J. (1998). Role for cingulate motor area cells in voluntary movement selection based on reward. Science *282*, 1335–1338.

Simmons, J.M., and Richmond, B.J. (2008). Dynamic changes in representations of preceding and upcoming reward in monkey orbitofrontal cortex. Cereb. Cortex *18*, 93–103.

Smith, D.V., Hayden, B.Y., Truong, T.K., Song, A.W., Platt, M.L., and Huettel, S.A. (2010). Distinct value signals in anterior and posterior ventromedial prefrontal cortex. J. Neurosci. *30*, 2490–2495.

Sutton, R., and Barto, A.G. (1998). Reinforcement Learning (Cambridge, Massachusetts: MIT Press).

Tanaka, S.C., Doya, K., Okada, G., Ueda, K., Okamoto, Y., and Yamawaki, S. (2004). Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. Nat. Neurosci. *7*, 887–893.

Tom, S.M., Fox, C.R., Trepel, C., and Poldrack, R.A. (2007). The neural basis of loss aversion in decision-making under risk. Science *315*, 515–518.

Tremblay, L., and Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. Nature *398*, 704–708.

Tsujimoto, S., Genovesio, A., and Wise, S.P. (2009). Monkey orbitofrontal cortex encodes response choices near feedback time. J. Neurosci. *29*, 2569–2574.

Tsujimoto, S., Genovesio, A., and Wise, S.P. (2010). Evaluating self-generated decisions in frontal pole cortex of monkeys. Nat. Neurosci. *13*, 120–126.

Van Hoesen, G.W., Morecraft, R.J., and Vogt, B.A. (1993). Connections of the monkey cingulate cortex. In Neurobiology of Cingulate Cortex and Limbic Thalamus, B.A. Vogt and M. Gabriel, eds. (Boston: Birkhauser).

Vogt, B. (2009). Architecture, neurocytology, and comparative organization of monkey and human cingulate cortices. In Cingulate Neurobiology and Disease, B. Vogt, ed. (New York: Oxford University Press).

Walton, M.E., Bannerman, D.M., and Rushworth, M.F.S. (2002). The role of rat medial frontal cortex in effort-based decision making. J. Neurosci. *22*, 10996–11003.

Walton, M.E., Bannerman, D.M., Alterescu, K., and Rushworth, M.F.S. (2003). Functional specialization within medial frontal cortex of the anterior cingulate for evaluating effort-related decisions. J. Neurosci. *23*, 6475–6479.

Walton, M.E., Devlin, J.T., and Rushworth, M.F.S. (2004). Interactions between decision making and performance monitoring within prefrontal cortex. Nat. Neurosci. *7*, 1259–1265.

Walton, M.E., Behrens, T.E., Buckley, M.J., Rudebeck, P.H., and Rushworth, M.F. (2010). Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. Neuron *65*, 927–939.

Wang, Y., Matsuzaka, Y., Shima, K., and Tanji, J. (2004). Cingulate cortical cells projecting to monkey frontal eye field and primary motor cortex. Neuroreport *15*, 1559–1563.

Wunderlich, K., Rangel, A., and O'Doherty, J.P. (2010). Economic choices can be made using only stimulus values. Proc. Natl. Acad. Sci. USA *107*, 15005–15010.

Yoshida, W., and Ishii, S. (2006). Resolution of uncertainty in prefrontal cortex. Neuron *50*, 781–789.

Yukie, M., and Shibata, H. (2009). Temporocingulate interactions in the monkey. In Cingulate Neurobiology and Disease, B. Vogt, ed. (New York: Oxford University Press).