# Expected Value, Reward Outcome, and Temporal Difference Error Representations in a Probabilistic Decision Task

Edmund T. Rolls, Ciara McCabe and Jerome Redoute

University of Oxford, Department of Experimental Psychology, South Parks Road, Oxford OX1 3UD, UK

In probabilistic decision tasks, an expected value (EV) of a choice is calculated, and after the choice has been made, this can be updated based on a temporal difference (TD) prediction error between the EV and the reward magnitude (RM) obtained. The EV is measured as the probability of obtaining a reward × RM. To understand the contribution of different brain areas to these decision-making processes, functional magnetic resonance imaging activations related to EV versus RM (or outcome) were measured in a probabilistic decision task. Activations in the medial orbitofrontal cortex were correlated with both RM and with EV and confirmed in a conjunction analysis to extend toward the pregenual cingulate cortex. From these representations, TD reward prediction errors could be produced. Activations in areas that receive from the orbitofrontal cortex including the ventral striatum, midbrain, and inferior frontal gyrus were correlated with the TD error. Activations in the anterior insula were correlated negatively with EV, occurring when low reward outcomes were expected, and also with the uncertainty of the reward, implicating this region in basic and crucial decision-making parameters, low expected outcomes, and uncertainty.

**Keywords:** expected utility, expected value, fMRI, insula, orbitofrontal cortex, reward magnitude, reward outcome, temporal difference error, ventral striatum

## Introduction

Reward and punishment magnitude are represented in the orbitofrontal cortex, as shown by investigations in which these are parametrically varied. One type of evidence has been obtained with food reward devaluation produced by sensory-specific satiety (Rolls et al. 1989; Kringelbach et al. 2003; Rolls 2004, 2005). Another has been with trial-by-trial variation of the monetary gain and loss, allowing correlation of activations of different parts of the orbitofrontal cortex with the magnitude of the monetary gain or loss (O'Doherty et al. 2001; Knutson et al. 2003).

However, the question arises of how a decision is influenced if the reward magnitude (RM) is high, but there is a small probability of obtaining the reward. Here we can adopt approaches used in reinforcement learning and neuroeconomics and use the terms RM for the magnitude of the reward obtained on a trial (that is the reward outcome, outcome value, or reward value) and the expected value (EV) as the probability of obtaining the reward multiplied by the RM (Sutton and Barto 1998; Dayan and Abbott 2001; Glimcher 2003, 2004; Rolls 2005). In an approach related to microeconomics, expected utility theory has provided a useful estimate of the desirability of actions (or at least of the action that is performed) and indicates that, except for very high and very low probabilities and for decisions with framing issues, expected utility, indicated by choices, approximately tracks EV with a concave (e.g., loga-rithmic) function that places less weight on high EVs (Bernoulli 1738; von Neumann and Morgenstern 1944; Kahneman and Tversky 1979, 1984; Tversky and Kahneman 1986; Gintis 2000). We show that the subjects' choices in our task, that is, the expected utility, are related to the EV (the probability multiplied by the RM).

We describe here a functional magnetic resonance imaging (fMRI) study of human decision making with a period on each trial after the decision has been taken in which the EV was known to the subjects (as a result of learning on the preceding trials). This was followed by a period in which the reward outcome was made known. One aim was to show whether the activation in some brain areas represents both EV (the prediction of what is likely to be obtained, given the choice just made) and RM (the amount that was obtained later in the trial). Such brain areas might compute EV by weighting RM by the probability of obtaining the reward. A second aim was to investigate how some brain areas may respond differently to EV (the predicted or anticipated outcome) and RM (the actual outcome) and thus contribute to decision making by providing separate representations of these key basic decision parameters. A third aim was to investigate whether in a decision-making task some brain areas have activations when RMs (outcomes) are better than expected (a positive temporal difference [TD] error or positive reward prediction error) or are worse than expected (a negative TD error or negative reward prediction error), as these signals may be key basic decision parameters (Sutton and Barto 1998; Doya 1999; Rolls 2005). The overall aim is to understand better the brain bases of decision making. We compare the results of this study with other functional neuro-imaging studies of reward expectation and reward outcome representations in which regions of interest have been identified in the orbitofrontal cortex, pregenual cingulate cortex, ventral striatum, and midbrain (Breiter et al. 2001; Knutson et al. 2003, 2005; O'Doherty et al. 2003, 2004; Daw et al. 2006; Kim et al. 2006) as described further in the Discussion.

Neuronal recording studies have shown that the relative desirability of action (or the expected utility of a choice) in oculomotor tasks is reflected in the firing of parietal and cingulate cortex neurons (Platt and Glimcher 1999; Glimcher 2003, 2004; Dorris and Glimcher 2004; Sugrue et al. 2004, 2005; McCoy and Platt 2005) and that dopamine neurons may respond to conditioned stimuli predicting reward with increasing probability and to reward uncertainty (Schultz 2006).

## Methods

### Task

In this probabilistic monetary reward decision task, the subjects could choose either on the right to obtain a large reward with a RM of 30

pence or on the left to obtain a smaller reward with a magnitude of 10 pence with a probability of 0.9. On the right, in different trial sets, the probability of the large reward was 0.9 (making the EV defined as probability × RM = 27 pence); the probability was 0.33 making the EV 10 pence; or the probability was 0.16 making the EV 5 pence (see Fig. 1*a*). (On the trials on which a reward was not obtained, 0 pence was the RM.) The participants learned in the sets of 30 trials with the different EVs on the right whether to press on the right or the left to maximize their winnings. They took typically less than 10 trials to adjust to the unsignalled change in EV every 30 trials (see example in Fig. 2*a* and Results), and neuroimaging analysis was performed for the last 20 trials of each set when the EV had been learned, and the choices made (i.e., the expected utility) stably reflected the EV. The 3 sets of 30 trials were run in a different randomized order for each of the 13 participants, and the transition between trial sets was not signalled to the subjects, who had been informed that the best choice to make would alter during the experiment, so that they should sample both sides and should make choices to maximize their winnings. The design of the task enabled probability of a reward to be separated from RM and EV, in that, for example, an EV of 9 pence on the left was linked with a probability of 0.9 and an EV of 10 pence on the right was linked with a probability of 0.33. The results in the figures showed that this aspect of the task design was useful, for whenever activations were related to EV, the percent blood oxygenation level-dependent (BOLD) signal change was very similar for the EV of 9 pence and of 10 pence, even though the probabilities of obtaining a reward were very different, 0.9 and 0.33. Further, in no case did the percent BOLD signal change fit a monotonic function of the probability of obtaining a reward, as can be seen by inspecting the percent BOLD change parts of the figures. Moreover, we used an event-related analysis design as shown below, so that the different probabil-

ities were mixed throughout the 90 trials (as the probability depended on whether a left or right choice was made, see Fig. 1*a*).

The timing of the task is shown in Figure 1*b*. On each trial, the subject was shown a yellow cross at time zero and selected using a button box either a right or a left response. At the end of the decision period, the choice was indicated by altering one of the left or right squares to white, and the subject was in a 6-s delay period for which the EV of the choice was known to the subject (as shown by the stable probability of right vs. left choices described in Results) for trials 11–30 of a given set based only on the rewards received given the choices made on preceding trials. (No explicit indication of the EVs of the choices was made known to the participants, who learned about the EVs based on the outcomes of their choices.) This EV period, time 4–10 s after the start of the trial, was used to analyze parts of the brain where activation reflected EV. In this period, the EV was known to the subject (as indicated by the subject's previous choices and analyzed below in Fig. 2*b*), but the actual reward that would be obtained was not yet known as it was probabilistic.

At time = 10–16 s after the start of the trial, the actual amount of reward that would be obtained (the reward outcome) was indicated by the display changing to green for 30 pence, blue for 10 pence, and red for 0 pence (see Fig. 1). In this period, the RM for that trial was known to the subject, and brain activations that reflected the RM were analyzed in this "RM" period (see Fig. 1*b*). At time *t* = 16 s, the reward was delivered with a message "you won 30 pence," "you won 10 pence," or "you did not win." (The design is to facilitate comparisons with neuronal recordings during decision making in macaques [cf., Sugrue et al. 2005], in which the color in the period labeled RM in Figure 1*b* indicates how much juice reward will be obtained on that particular trial, but the juice is delivered in the reward period labeled "Reward" in Figure 1*b*. This will allow the firing rate of a neuron to be measured
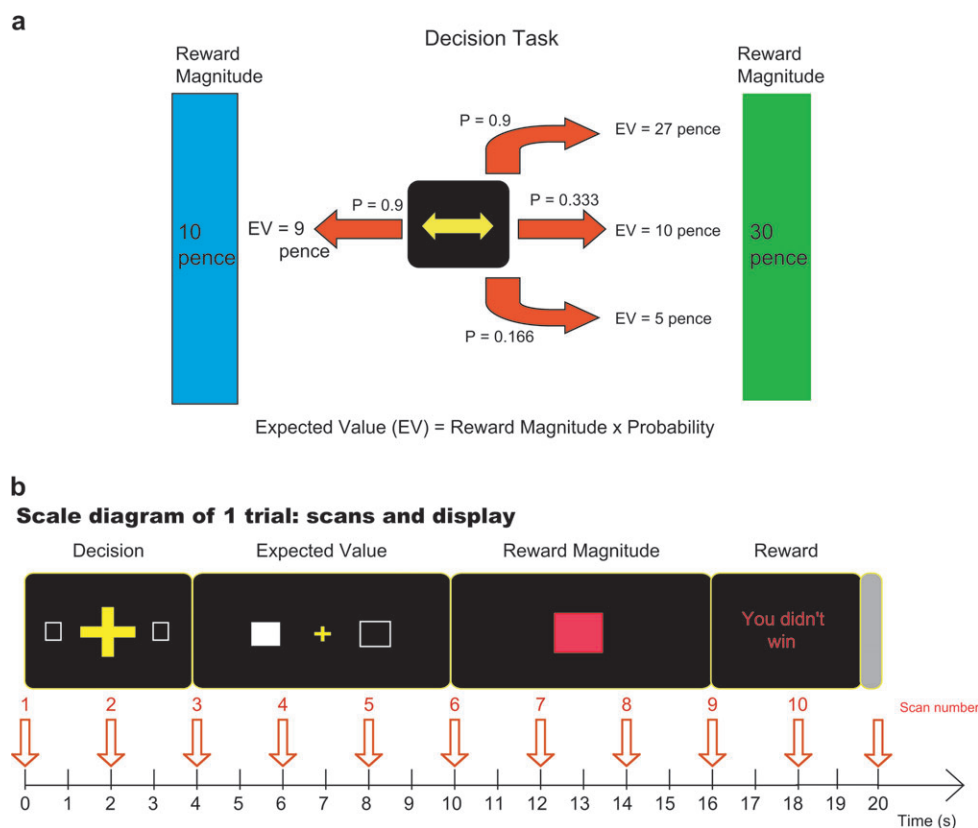


**Figure 1.** (*a*) Task design. In the probabilistic monetary reward decision task, subjects could choose either on the right to obtain a large reward with a magnitude of 30 pence or on the left to obtain a smaller reward with a magnitude of 10 pence with a probability of 0.9. On the right, in different sets each 30 trials long, the probability of the large reward was 0.9 (making the EV defined as probability × RM = 27 pence); the probability was 0.33 (EV = 10 pence); or the probability was 0.17 (EV = 5 pence). (On the trials on which a reward was not obtained, 0 pence was the RM). (*b*) Task timing (see Methods). The timing is shown in seconds with the corresponding scan numbers in each trial. Activations related to EV started at *t* = 4 s and to RM (or outcome) at *t* = 10 s (see Methods).

**a**

Single subject's responses to changes in the Expected Value
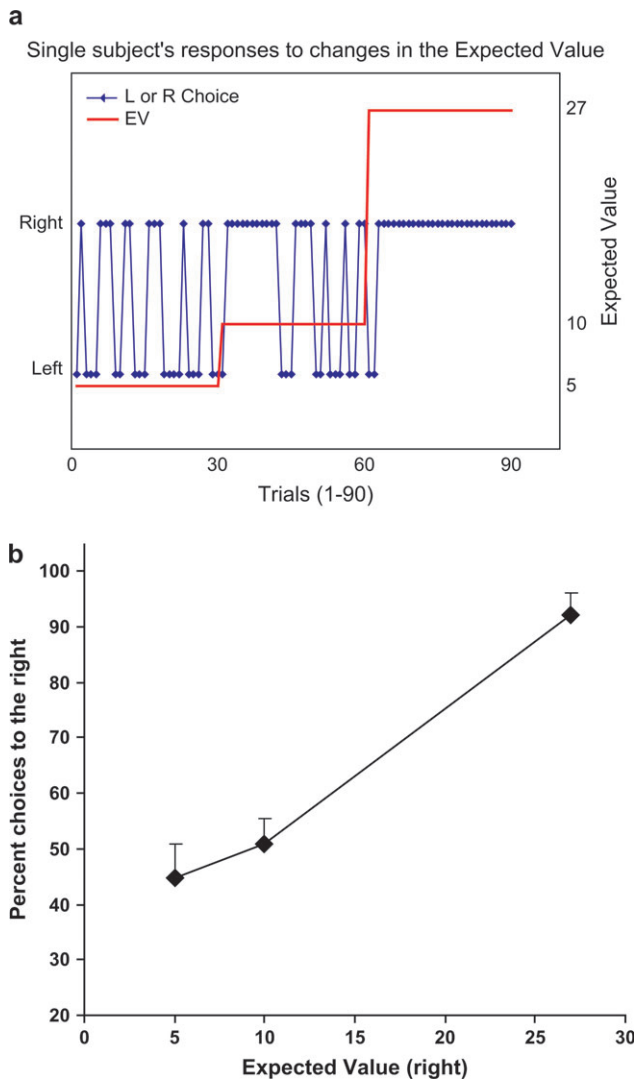


**b**

**Figure 2.** Task performance. (*a*) The choices of a subject in the decision task. The solid line shows the EV on the right in 3 different sets of 30 trials. Right and Left responses are shown. Each subject was run with a random order of the different EV trial sets. (*b*) The percentage of choices to the right as a function of the EV on the right (mean ± standard error of mean across all the subjects for the last 10 trials in a set). The EV for a left choice was 9 pence, in that 10 pence was obtained for a left choice with a probability of 0.9.

both when a particular EV applies, then later in the same trial when a particular RM has been signalled by the color, and then later in the same trial when a particular juice Reward has been delivered. The aim of the design is to show whether the same single neurons code for both the EV and the RM and, if so, whether the representation is on the same scale for EV and for RM. There are thus 2 times in each trial at which learning can potentially occur: at the transition between the EV and RM part of the trial at time = 10 s in Figure 1*b* and at the transition between the RM and Reward part of the trial at time = 16 s in Figure 1*b*. In this situation, there are TD errors that can vary at these 2 transition points at different stages of the learning. It is to facilitate comparison with results obtained by neurophysiology in macaques that we use the term TD error in this paper. In fact, in this human imaging study, the association between the color in the RM part of the trial and the Reward part of the trial had been learned in pretraining before the imaging started, so there was one transition point in the trial, which at time = 10 s when the subject knowing the EV was informed about the RM that would be obtained. This error, though called a TD error in this paper for the reasons just given, could equally well be called a reward prediction error in this human fMRI investigation.) The trials were spaced 20 s apart.

The design of the task meant that sometimes the participants were expecting a low probability of a high reward of 30 pence and unexpectedly obtained a high reward value of 30 pence. On these trials, the TD prediction error from the EV part of the trial when the decision was being made and in the delay period after it was made to the RM or outcome part of the trial when subjects were informed whether they would obtain the large reward was positive. On other trials when the EV was high but probabilistically no reward was obtained, the TD prediction error from the EV part of the trial to the RM part of the trial was negative. TD errors and TD learning are described by Sutton and Barto (1998), Doya (1999), Dayan and Abbott (2001), and Rolls (2005). We calculated the TD reward prediction error at time $t = 10$ s in each trial as just described (as this is the time at which the EV for that trial changes to the RM value representing the actual outcome) and used this as a regressor in a correlation analysis with the BOLD signal. The TD error calculated was therefore the difference on any trial between the EV, which was constant for the last 20 trials of each trial set as indicated by stable choice behavior reflecting the learning that had occurred in the first 10 trials of each trial set, and the RM actually indicated at $t = 10$ s (see Fig. 1). If, for example, the EV was 27 pence and the RM was 0 pence, then the TD error was –27 pence. If the EV was 5 pence and the RM was 30 pence, then the TD error was +25 pence.

We were also able to measure another fundamental decision parameter in this task, reward uncertainty, which is the variance of the magnitude of the reward obtained for a particular choice (Tobler et al. 2007). (The variance = $[\Sigma_i (RM_i - EV)^2]/n$, where RM is the RM, $n$ is the number of elements [outcomes associated with each stimulus], and $i$ is index = 1–$n$.) The reward variance values that corresponded to each choice were as follows: EV = 27, variance = 8.1; EV = 10, variance = 200; EV = 9, variance = 41; and EV = 5, variance = 125, and it can be seen that there is no monotonic correspondence between EV and variance in this study.

### Participants

Thirteen healthy right-handed participants (age range 23–35 years, 8 males and 5 females) participated in the study. Written informed consent from all participants and ethical approval (Oxfordshire Research Ethics Committee) were obtained before the experiment.

### fMRI Data Acquisition

Images were acquired with a 3.0-T VARIAN/SIEMENS whole-body scanner at the Centre for Functional Magnetic Resonance Imaging at Oxford (FMRIB), where 19 $T_2$*-weighted echo planar imaging (EPI) coronal slices with in-plane resolution of 3 × 3 mm and between plane spacing of 6 mm were acquired every 2 s (time repetition = 2). We used the techniques that we have developed over a number of years (e.g., O'Doherty et al. 2001; de Araujo et al. 2003) and as described in detail by Wilson et al. (2002) to carefully select the imaging parameters in order to minimize susceptibility and distortion artifact in the orbitofrontal cortex. The relevant factors include imaging in the coronal plane, minimizing voxel size in the plane of the imaging, as high a gradient switching frequency as possible (960 Hz), a short echo time of 25 ms, and local shimming for the inferior frontal area. The matrix size was 64 × 64, and the field of view was 192 × 192 mm. Continuous coverage was obtained from +62 (A/P) to –52 (A/P). Acquisition was carried out during the task performance yielding 900 volumes in total. A whole-brain $T_2$*-weighted EPI volume of the above dimensions and an anatomical $T_1$ volume with coronal plane slice thickness 1.5 mm and in-plane resolution of 1.5 × 1.5 mm were also acquired.

### fMRI Data Analysis

The imaging data were analyzed using SPM2 (Wellcome Institute of Cognitive Neurology). Preprocessing of the data used SPM2 realignment, reslicing with sinc interpolation, normalization to the Montreal Neurological Institute (MNI) coordinate system (Collins et al. 1994), and spatial smoothing with a 8-mm full width at half maximum isotropic Gaussian kernel. The time series at each voxel was low-pass filtered with a hemodynamic response kernel. Time series nonsphericity at each voxel was estimated and corrected (Friston et al. 2002), and a high-pass filter with a cutoff period of 128 s was applied. In the single event design, a general linear model was then applied to the time course of activation where stimulus onsets were modeled as single impulse response

functions and then convolved with the canonical hemodynamic response function (Friston et al. 1994; duration 0). Linear contrasts were defined to test specific effects. Time derivatives were included in the basis functions set. Following smoothness estimation (Kiebel et al. 1999), linear contrasts of parameter estimates were defined to test the specific effects of each condition with each individual data set. Voxel values for each contrast resulted in a statistical parametric map of the corresponding $t$-statistic, which was then transformed into the unit normal distribution (SPM $Z$). The statistical parametric maps from each individual data set were then entered into second-level, random effects analyses accounting for both scan-to-scan and subject-to-subject variability. More precisely, the sets of individual statistical maps corresponding to a specific effect of interest were entered as covariates in multiple regression models (analysis of variance [ANOVA] without a constant) as implemented in SPM2, and the corresponding group effects were assessed by applying linear contrasts (again following smoothness estimation) to the (second level) parameter estimates generating a $t$-statistics map for each group effect of interest. The correlation analyses of the fMRI BOLD signal with given parameters of interest (e.g., the EV and the RM) were performed at the second level through applying 1-sample $t$-tests to the first-level $t$-maps resulting from performing linear parametric modulation as implemented in SPM2. The correlation and contrast analyses for RM were formed on the basis of the RM on that trial (30, 10, or 0 pence) signalled at $t = 10$ s in the trial and using an onset time of $t = 10$ s. The correlation and contrast analyses for EV were formed on the basis of the EV on that trial (taking into account the subject's choice of right vs. left), which was available by $t = 4$ s in the trial and using an onset time of $t = 4$ s, which was scan 3. For example, the RMs were performed at the second level through applying 1-sample $t$-tests to the first-level $t$-maps resulting from performing linear parametric modulation as implemented in SPM2. Unless otherwise stated, reported $P$ values for each cluster based on this group analysis are corrected (fc) for the number of comparisons (resels) in the entire volume ("whole-brain" multiple comparisons; Worsley et al. 1996). Peaks are reported for which $P < 0.05$. For brain regions where there was a prior hypothesis as described in the Introduction and under Methods, namely, in the parts of the orbitofrontal cortex, pregenual cingulate cortex, ventral striatum, and midbrain shown to be of interest because of activations found in prior studies in reward paradigms or because neurophysiological, lesion, or connection evidence links them to reward processing, including activation that depends on monetary RM (Breiter et al. 2001; O'Doherty et al. 2001, 2003; Fiorillo et al. 2003; Knutson et al. 2003, 2005; Kringelbach and Rolls 2004; Rolls 2005, 2006), we used small volume corrections (SVCs). These activations correspond to voxels significant when corrected for the number of comparisons made within each region (SVC applied with a sphere of 8 mm chosen to be greater than or equal to the spatial smoothing kernel; Worsley et al. 1996). Peaks with $P < 0.05$ false discovery rate corrected across the small volume were considered significant. The mean percent change in the BOLD signal for different EVs and RMs within regions of interest identified from the correlation analyses were extracted using the SPM Marsbar toolbox documented at http://marsbar.sourceforge.net/.

## Results

All the 13 subjects learned to alter their choices when the probabilities changed at the start of each of the 3 sets of 30 trials each and had adjusted within 10 trials (see example in Fig. 2$a$). Moreover, Figure 2$b$ shows that the subjects' choices (indicating the desirability of the actions and thus reflecting expected utility) reflected the different EVs, as shown by the percentage of choices to the right as a function of the EV on the right ($F_{2,24} = 21.6$, $P < 10^{-5}$). This graph shows that (with an EV of 9 pence on the left in all cases, i.e., of obtaining 10 pence with a probability of 0.9), when the EV on the right was 27 pence (30 pence with probability 0.9), the percentage of choices on the right was 92%; when the EV on the right was 10 pence (30 pence with probability 0.33), the percentage of choices on the right was 51%; and when the EV on the right was 5 pence (30

pence with probability 0.17), the percentage of choices on the right was 44%. Indeed, the line in Figure 2$b$ intersects the 50% right versus left choice point at approximately 10 pence, showing that the subjects chose the right and the left equally often when the EV on the right was 10 pence, which is close to the EV for a left choice (9 pence). This shows very clearly that in this experiment, the EV was known to the subjects in that it affected the desirability of the choice (or expected utility) of right versus left choices and that it is therefore appropriate to measure how brain activations are related to the EV. Indeed, at the 50% right–left choice point, the subjects' choice was guided by EV, in that their choices matched the EV, even though on the left this was achieved by obtaining 10 pence with probability 0.9 and on the right by obtaining 30 pence with probability 0.33. In microeconomic terms, with these parameters, the left/right choice had equal expected utility as shown by the choices made when the EVs of a left versus right choice were equal. (The probability of a choice reflects relative EVs in a probabilistic decision task, and this has been investigated in the matching law literature [Sugrue et al. 2004, 2005]).

The mean number of trials across subjects for learning to occur and the choices to settle after the EV on the right changed was $5.7 \pm 3$ (standard deviation) (taking as a criterion, the first trial at which 8 of the 10 choices had the same proportion to the right as those that occurred later in the trial set). Analyses for the EV period thus were performed for trials 11–30 within each set of 30 trials, when the probability of choices of right versus left had settled to a stable value and the EVs in a set were influencing the choices made in that set. Every subject had a consistent probability of choices to the right in the last 20 trials, so that in these 20 trials the subject was demonstrating by the choices being made that the estimate of EV of each choice was stable for these last 20 trials. Moreover, the probability of the choices of every subject was influenced by the EV (reflected also in the small standard errors in Fig. 2$b$). We note that the correlation analyses were based on the last 20 trials of each set so that 60 trials for both EV and RM were available and that many trials were available (given the subjects' choices and the probabilistic nature of the task) for all RMs as well as all EVs. EV analyses started at the start of the EV period shown in Figure 1$b$, that is, at $t = 4$ s from the start of the trial. In that the EV for trials 11–30 had settled to a value that produced stable choices, the imaging period starting at 4 s on these trials was in a period in which the EV (but not the RM) was known to the subjects. We note that the response that indicated the decision was made typically at between $t = 2$ and $t = 4$ s from the start of the trial and that further statistical analyses for the decision period starting at $t = 2$ s produced similar patterns of brain activation to those for the EV period starting at $t = 4$ s. The EV activations described below thus reflect activity when the EV of a choice was known to the subjects (as they had just made the decision of left vs. right and had received 10 or more trials with that probability) and in the immediately preceding decision period while they were making the choice. RM analyses started at the start of the RM period shown in Figure 1$b$, that is, at $t = 10$ s from the start of the trial, when the outcome of the choice was made known to the subject by the color of the display. (Similar patterns of brain activations were found if the RM activations were calculated later into the RM period, including the time when words were used to convey the outcome at the time labeled Reward in Fig. 1$b$.) The results are described for each of the main brain regions in which

significant correlations ($P < 0.05$ fully corrected or $P < 0.05$ SVC in regions of interest based on prior studies as described in Methods) with EV, RM, or both were found.

### Medial Orbitofrontal Cortex

Activations in the medial orbitofrontal cortex [–6, 42, –22] had a positive correlation with RM ($P = 0.003$ SVC) as shown in Figure 3a. A positive correlation with EV was also found in the orbitofrontal cortex (with peak voxel at [19, 51, –3] [$P = 0.015$ SVC]), and it was of interest that at, for example, the [–6, 42, –22] medial orbitofrontal cortex site shown in Figure 3a, the activations for different EVs did approximately parallel the RM (see Fig. 3b).

To investigate whether there was a region of the orbitofrontal cortex where activations reflected both RM and EV, we performed a conjunction analysis at the second level for the correlations of the BOLD signal with the RM and the EV. (Separate regressors from different SPM analyses were used, and sphericity correction was applied). The result, shown in Figure 3c, was that a large region of the medial orbitofrontal cortex extending up toward the pregenual cingulate cortex was activated by this conjunction (peak at [2, 38, –14] [$P < 0.03$ SVC]). (The BOLD signal from the same region of the medial orbitofrontal cortex is shown in Fig. 3b). The implication of this result is that the medial orbitofrontal cortex and adjoining pregenual cingulate cortex represent both the actual amount of reward received at reward outcome time (RM) and the amount of reward expected probabilistically for a particular choice measured earlier in a trial (the EV). No other brain area had statistically significant effects in this conjunction analysis. We note that as described under Methods, Task, the probability of obtaining any reward greater than 0 pence did not fit the percent BOLD change graphs, and therefore, the activations described as related to EV were related to the EV of the reward that would be obtained and not just to the probability that any reward would be obtained. (The dissociation was made possible by the inclusion of the 9 pence and 10 pence EVs in the design, which had very different reward probabilities. In no case did any activation in this paper reflect probability of reward as distinct from EV [cf., Abler et al. 2006; Yacubian et al. 2006; Tobler et al. 2007]).

### Ventral Striatum

The group random effects analysis showed a correlation of the BOLD signal with the RM in the ventral striatum at MNI coordinates [12, 16, –12] ($P < 0.038$ fully corrected) as shown in Figure 4a. This correlation is shown clearly in Figure 4b, which shows the percent change in the BOLD signal for the 3 different RMs, 30 pence, 10 pence, or 0 pence for the region of interest defined by the area in which the main analysis showed a significant correlation. The group analysis showed no correlation in this brain region with the EV on each trial (measured in the EV period starting at time $t = 4$ s in each trial), and this is also evident in the percent BOLD signal in this region of interest, which had no significant differences for the different EVs, as confirmed by the error bars in Figure 4b. However, some evidence of a positive correlation with EV was found in a more posterior and dorsal part of the ventral striatum at [–2, 0, –2] ($P < 0.038$ SVC). However, it was in a separate part of the ventral striatum to that with a correlation with RM, it was less statistically significant, and it was not reflected in the percent BOLD analysis at [12, 16, –12], as shown in Figure 4b. Thus, the





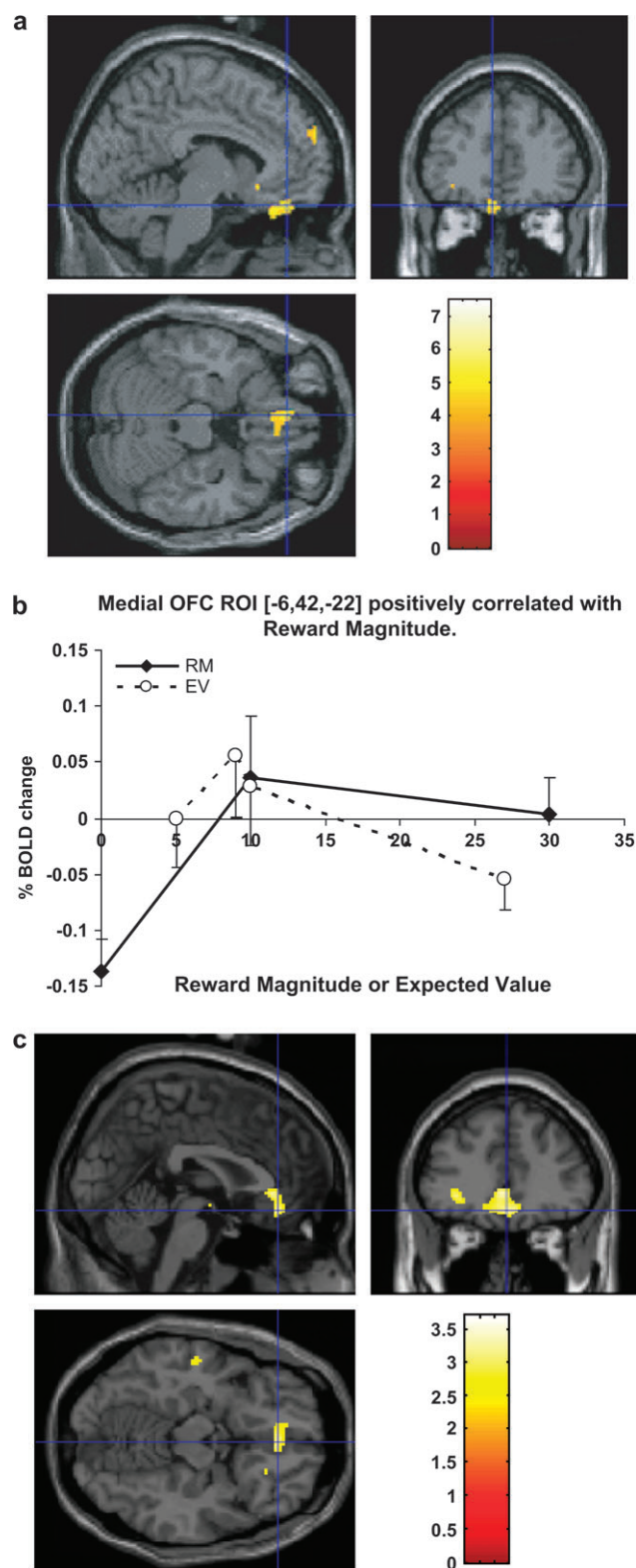Medial OFC ROI [–6,42,–22] positively correlated with Reward Magnitude.



**Figure 3.** (a) Medial orbitofrontal cortex. A positive correlation of the BOLD signal with the RM was found in the medial orbitofrontal cortex at MNI coordinates [−6, 42, −22] ($P < 0.003$ SVC). (b) The percent change in the BOLD signal for the 3 RMs (30 pence, 10 pence, or 0 pence) for the region of interest defined by the correlation analysis. The means and standard errors are shown. The percent change in the BOLD signal for the 4 EVs (EVs of 27 pence, 10 pence, 9 pence, and 5 pence) for the same region of interest are also shown. (c) Medial orbitofrontal cortex. Conjunction analysis of correlations with EV and correlations with RM at MNI coordinates [2, 38, −14] ($P < 0.03$ SVC).
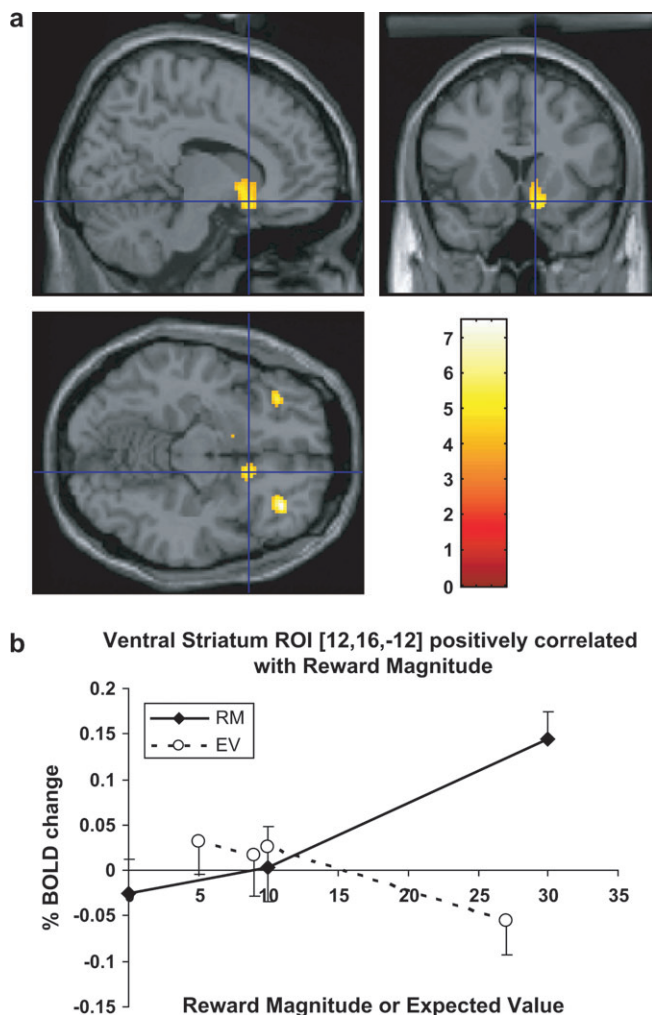
**Figure 4.** Nucleus accumbens/ventral striatum. (*a*) The group random effects analysis showed a positive correlation between the BOLD signal and the RM in the ventral striatum at MNI coordinates [12, 16, −12] (*P* < 0.038 fully corrected). (*b*) The percent change in the BOLD signal for the 3 RMs (RMs of 30 pence, 10 pence, or 0 pence) for the region of interest defined by the correlation analysis. The means and standard errors are shown. The percent change in the BOLD signal for the 4 EVs (EVs of 27 pence, 10 pence, 9 pence, and 5 pence) for the same region of interest are also shown.



**Figure 5.** Insula. A negative correlation between the BOLD signal and the EV was found in the insular cortex at MNI coordinates [−38, 24, 16] (*P* < 0.001 fully corrected) and [42, 14, 0] (*P* = 0.01 SVC). The percent change in the BOLD signal for the 4 EVs (EVs of 27 pence, 10 pence, 9 pence, and 5 pence) for the region of interest defined by the correlation analysis are shown. The percent change in the BOLD signal for the 3 RMs (RMs of 30 pence, 10 pence, or 0 pence) for the same regions of interest are also shown. The means and standard errors averaged over the left and right are shown.

most significant result for the ventral striatum was a positive correlation with RM.

### Insula

A very different type of result was found in the insula (bilaterally), where activations were negatively correlated with EV ([−38, 24, 16], *P* < 0.001 fully corrected), as shown in Figure 5*a*. The basis for this is shown in Figure 5*b*, which also confirms that the activations were not related to the RM in this region. (The difference between the correlations with EV and RM was confirmed by ANOVA). This region of activation continued anteriorly to the junction between the insula and the orbitofrontal cortex at [−46, 22, −14] (*P* = 0.015 SVC), in what is probably cytoarchitecturally the agranular insular cortex (de Araujo et al. 2003; Kringelbach and Rolls 2004) and posteriorly to beyond *y* = 14. This activation was also present in the decision period (e.g., at scan 2, *P* < 0.001 fully corrected).

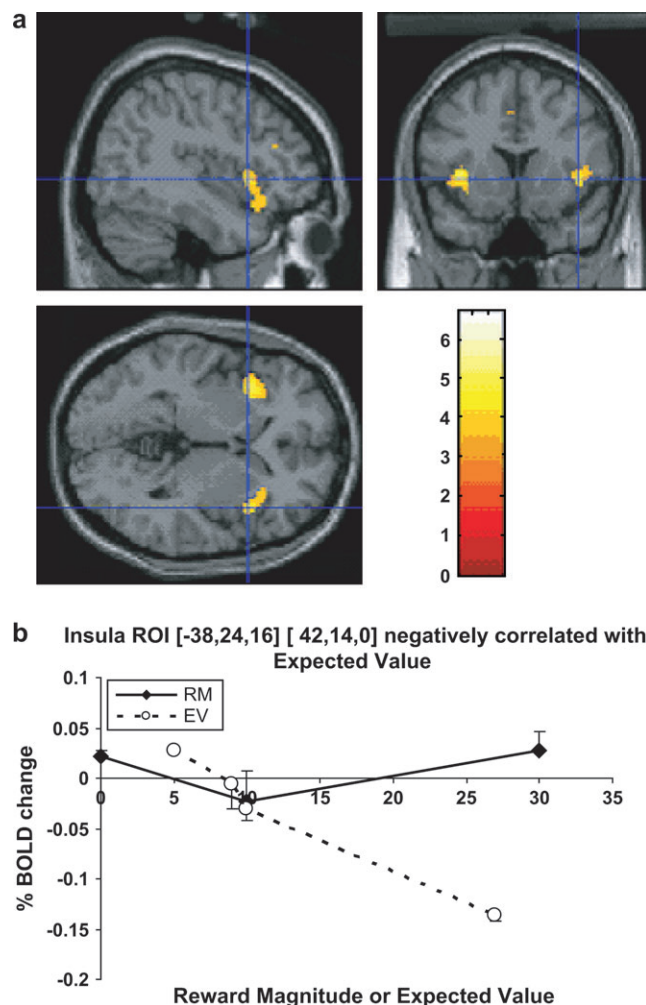To investigate the basis for the activations in the anterior insula further, we performed a correlation analysis with the variance of the reward obtained for each choice made in the task. The only region with a significant correlation was the anterior insula [46, 16, 6] (*Z* = 3.05 SVC 0.02), with some evidence for a correlation in the dorsal cingulate cortex [0, 16, 44] (*Z* = 2.26 SVC 0.058). Interestingly, the insula activation was restricted to the anterior part of the insula and did not extend behind *y* = +7 to the mid and posterior insula. High variance corresponds to high uncertainty about the magnitude of the reward that will be obtained on a particular trial, and the high uncertainty may be unpleasant. Together, these findings suggest that the anterior insula may reflect in this task low EV and much uncertainty about the reward that will be obtained by a particular choice (see Discussion).

### TD Error-Related Activations

Correlations in the following 3 brain regions were related to the TD error. Given the design of this investigation, these correlations with TD error can be understood in terms of the change from a given EV on each trial to a given RM at *t* = 10 s, and how EV and RM are represented in a given brain region, as follows.

## Nucleus Accumbens

A correlation analysis with the TD reward prediction error at time $t = 10$ s in the trial (calculated as described in Methods) showed a positive correlation in the nucleus accumbens at MNI coordinates [8, 8, –8] ($P < 0.048$ fully corrected) as shown in Figure 6a. The percent BOLD signal in a region of interest defined by this activation bilaterally showed that in this cluster, the BOLD signal was strongly correlated with the RM (calculated at $t = 10$ s), and there is no relation to the EV as shown in Figure 6b. The implication of the results shown in Figure 6 is that the TD error correlation arose in the nucleus accumbens because at the time that the EV period ended and the subject was informed about how much reward had been obtained on that trial (the reward outcome, signalled at time $t = 10$ s as shown in Fig. 1b), the BOLD signal generally changed to a higher value for large RM, and to a lower value for low or 0 RM, from a value that was not a function of the EV on that trial. This is confirmed by Figure 6c, which shows the percent change in the BOLD signal as a function of the TD error in the nucleus accumbens, with the TD error calculated at time = 10 s in the trial (see Fig. 1).

## Inferior Frontal Gyrus

A positive correlation of brain activations with the TD (reward prediction) error was found in the left hemisphere at [–44, 4, 22] ($P = 0.006$ fully corrected) close to the border between area 44 (Broca's area) and area 9, as shown in Figure 7a. The basis for this TD error is suggested by what is shown in Figure 7b, in that it is clear that almost independently of the EV (i.e., apart from EV9), there will be an increase in activation at the transition between the EV and RM part of each trial at time $t = 10$ s if the RM is 10 pence or 30 pence and a decrease in percent BOLD signal if the RM is 0 pence on a trial. The TD error correlation in this brain region is thus mainly related to the fact that the activations are related to RM and much less to EV.

## Midbrain

In a part of the midbrain at [14, –20, –16], there was a positive correlation with the TD error ($P = 0.032$ SVC) (see Fig. 8a). The results shown in Figure 8b suggest that this was related to higher values of the percent BOLD signal for the higher RMs and lower values of the percent BOLD signal for higher expected reward values. In another part of the midbrain [–10, –10, –4], which may be closer to the dopamine neurons, a negative correlation with the TD error was found ($P = 0.02$ SVC), and this could arise because, for example, this region shows activation when high EV is followed by a low RM. The only other location where a negative correlation with TD error was found was in the anterior cingulate cortex [8, 36, 42] ($P < 0.03$ SVC) in a dorsal part close to the region that is activated on reversal trials (when expected rewards are not being obtained) in a reversal task (Kringelbach and Rolls 2003). This part of the midbrain [–10, –10, –4], and the anterior cingulate cortex, may thus be responding in relation to frustrative nonreward or to errors that arise when less is obtained than is expected, so that behavioral choice should change, as in a reversal task (Kringelbach and Rolls 2003).

## Discussion

In this investigation, the RM on a trial was correlated with the activation in parts of the orbitofrontal cortex, and these cortical areas may be the origins of some of the signals found in other
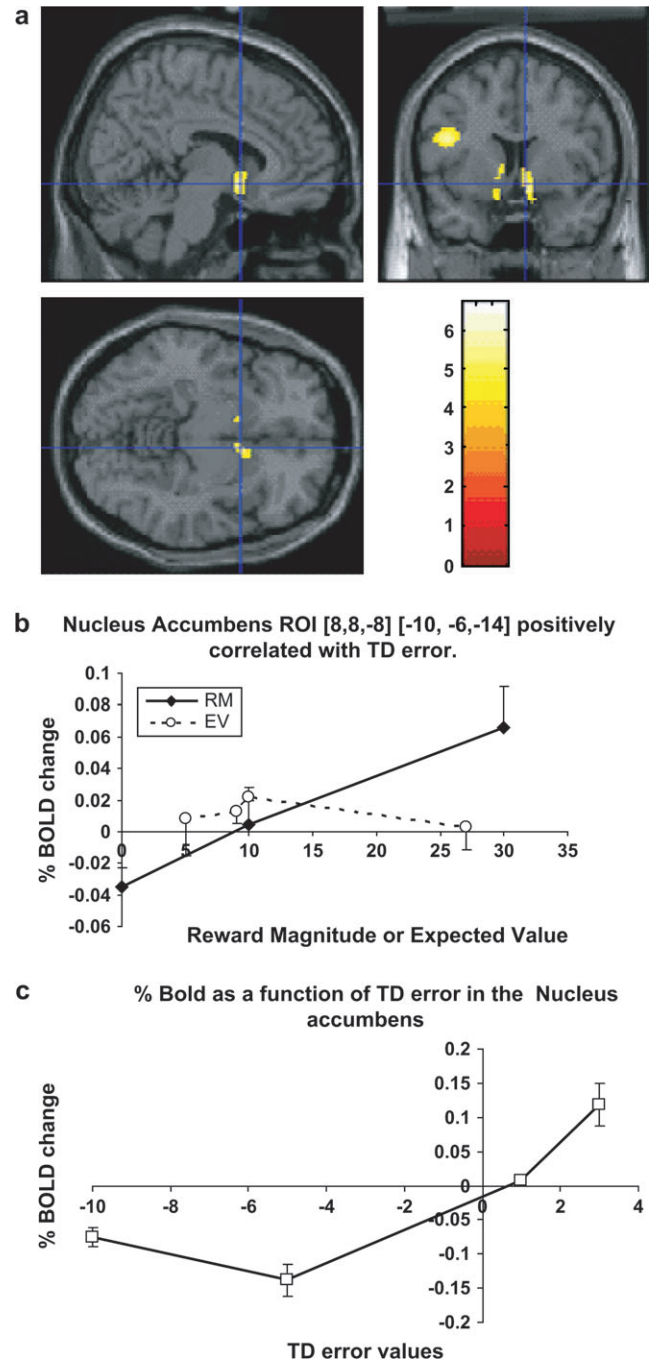


**Figure 6.** Nucleus accumbens/ventral striatum. (a) A positive correlation between the BOLD signal and the TD error was found in the ventral striatum at MNI coordinates [8, 8, −8] ($P < 0.048$ fully corrected) and [−10, 6, −14] ($P < 0.01$ SVC). (b) The percent change in the BOLD signal for the 3 RMs (RMs of 30 pence, 10 pence, or 0 pence) for the regions of interest defined by the correlation analysis. The means and standard errors over the left and right are shown. The percent change in the BOLD signal for the 4 EVs (EVs of 27 pence, 10 pence, 9 pence, and 5 pence) for the same region of interest are also shown. (c) The percent change in the BOLD signal as a function of the TD error in the nucleus accumbens (means and standard errors of the means are shown).

brain areas such as the ventral striatum. Activation of the orbitofrontal cortex by monetary reward value (or "outcomes") has been found in previous studies (O'Doherty et al. 2001; Knutson et al. 2003; Rogers et al. 2004). In the medial orbitofrontal cortex, the activation in the EV period paralleled
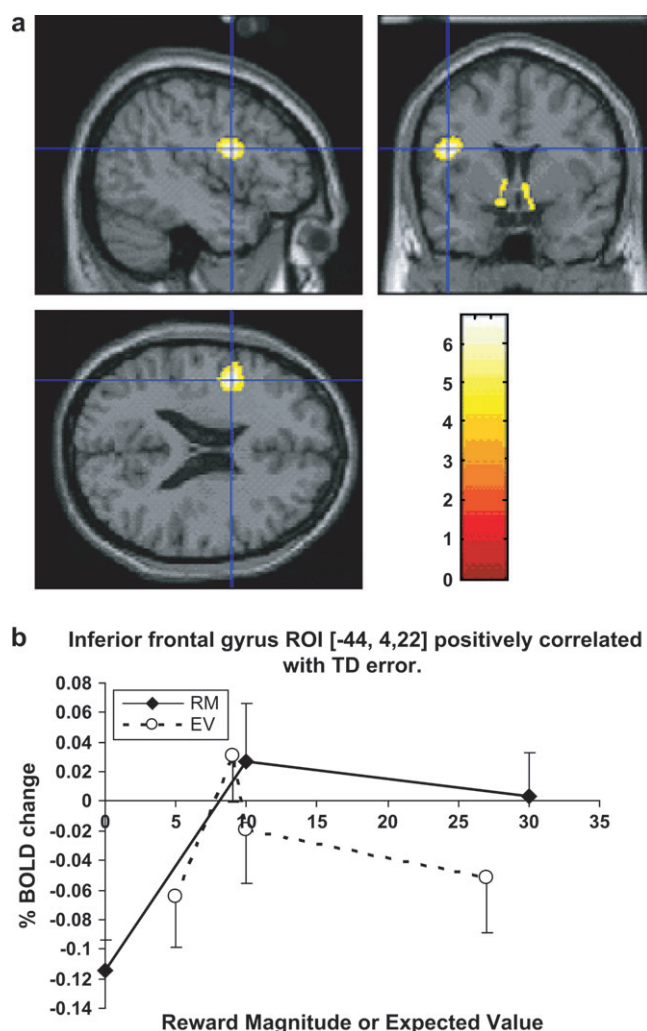
**Figure 7.** (a) A positive correlation between the BOLD signal and the TD error was found in area 44 at MNI coordinates [−44, 4, 22] (P = 0.006 fully corrected). (b) The percent change in the BOLD signal for the 3 RMs (30 pence, 10 pence, or 0 pence) for the region of interest defined by the correlation analysis. The means and standard errors are shown. The percent change in the BOLD signal for the 4 EVs (27 pence, 10 pence, 9 pence, and 5 pence) for the same region of interest are also shown.

**Figure 8.** (a) A positive correlation between the BOLD signal and the TD error was found in the midbrain at MNI coordinates [14, −20, −16] (P = 0.032 SVC). (b) The percent change in the BOLD signal for the 3 RMs (30 pence, 10 pence, or 0 pence) for the region of interest defined by the correlation analysis. The means and standard errors are shown. The percent change in the BOLD signal for the 4 EVs (EVs of 27 pence, 10 pence, 9 pence, and 5 pence) for the same region of interest are also shown.

the activation in the RM period for the best estimate at each stage of the same amount of reward being available on that trial (Fig. 3), and consistent with this there was a positive correlation with EV in the orbitofrontal cortex at [19, 51, −3]. Further, the conjunction analysis in Figure 3c showed that the same extensive area of the medial orbitofrontal cortex extending up toward the pregenual cingulate cortex had activations correlated with both RM and EV in this probabilistic decision task. The activation of the same part of the medial orbitofrontal cortex by both RM and by expected value or utility has not to our knowledge been shown previously. An implication is that these medial orbitofrontal cortex activations tracked the best estimate of reward value throughout each trial. This is the EV for that trial (which was known to the participant based on previous learning) throughout the period from t = 4 (or earlier) until t = 10 s, and after that the actual RM (or outcome) for that trial signalled at time t = 10 s. Given the fact that activations in the medial orbitofrontal cortex do reflect many different types of reward, for example, taste (de Araujo et al. 2003), food flavor (de Araujo et al. 2003; Kringelbach et al. 2003; McCabe and Rolls
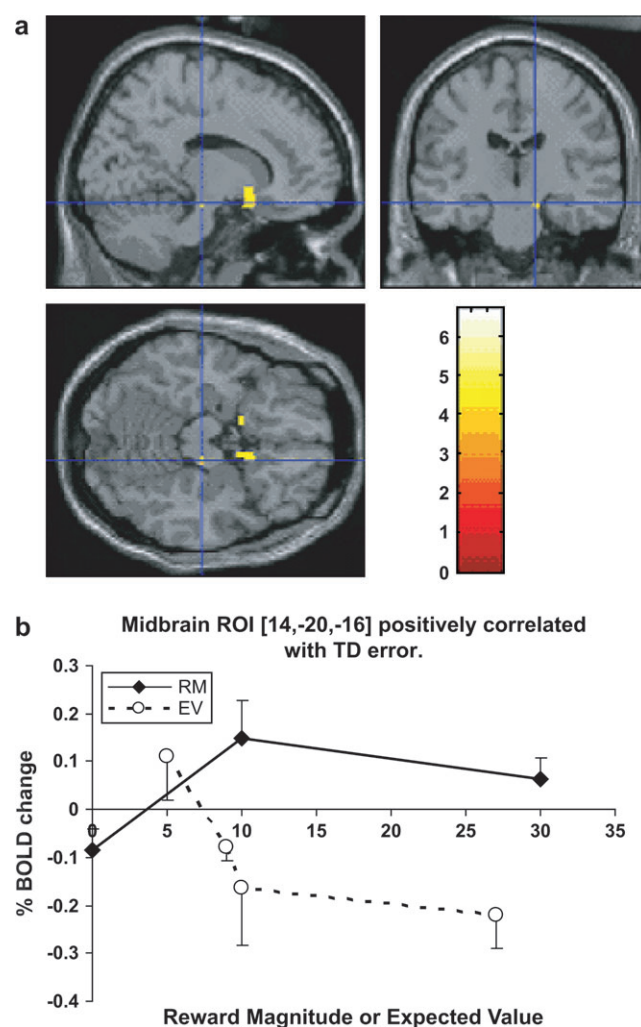
2007; cf., Rolls and Baylis 1994), olfactory (O'Doherty et al. 2000; Rolls et al. 2003; de Araujo et al. 2005), visual (O'Doherty et al. 2002, 2003), and monetary (O'Doherty et al. 2001), the finding that its activation may also be related to EV with probabilistic reward before the reward outcome is known for that trial, makes the medial orbitofrontal cortex a candidate brain region where the neurons that represent RM have their firing modulated by the probability that the RM will be obtained to represent EV. Moreover, in the orbitofrontal cortex, the activations to EV and RM may be on the same scale, in that in the medial orbitofrontal cortex the activations to different EVs and RMs are relatively close to following a similar function (see Fig. 3b).

Activations in one part of the ventral striatum had positive correlations with the RM (Fig. 4), and in this part, there was no correlation with EV. In a separate part of the ventral striatum [−2, 0, −2], a correlation with EV but not with RM was found. Consistently, the conjunction analysis for RM and EV did not show significant effects in the ventral striatum. The implication

is that although both signals reach the ventral striatum, they may be represented by different neurons.

The present study found activations of the medial orbitofrontal cortex related to both RM and to EV during a decision task in which the EVs are being learned based on the probability of the rewards obtained on recent trials. In a study by Knutson et al. (2005) in which the probability and reward value were signalled (i.e., it was not a decision task), it was found in a reward anticipation period (before the reward outcome on a given trial was known) that the main brain region where EV (defined in their study by valence × probability × RM) was correlated with activations was in a rather dorsal part of the anterior cingulate cortex. In another study without any decisions being made in which there was again a "prospect" or "expectancy" period in which one set of several sets of possible reward values was known, activations in the amygdala and orbital gyrus tracked the EV of the prospects, and activations in the nucleus accumbens, amygdala, and hypothalamus were related to the actual reward value on a trial (outcome) (Breiter et al. 2001). However, as noted, no decisions were required in those studies. Part of the interest of the present investigation is that decisions are actually being made, and in this situation, both the EV and the RM activate the same part of the medial orbitofrontal cortex, as shown by the conjunction analysis in Figure 3c. (In a gambling task study, activations in the medial orbitofrontal cortex were related to RM and in the ventromedial prefrontal cortex to EV, but a conjunction analysis was not performed to test whether these regions overlapped [Daw et al. 2006]. Consistent with the finding reported in the present investigation, in another study, overlap was reported but was not confirmed by a conjunction analysis of the type performed in the present investigation [Kim et al. 2006].) Moreover, the choices actually made in our task reflected the EV obtainable for a particular choice (Fig. 2b), showing that the EV (based on the recent history of rewards) was being processed and was influencing the decision making in the task. The fact that the EV influenced the choices made in the task (Fig. 2b) is consistent with the hypothesis we propose that the activations in the medial orbitofrontal cortex and pregenual cingulate cortex influence expected utility, that is, the choices made. Consistent with this, lesions of the orbitofrontal cortex do impair decision making (Rolls et al. 1994; Bechara et al. 1999; Fellows and Farah 2003; Hornak et al. 2004), and it would be of interest to investigate this more with the type of task used here in which expected utility can be measured by the choices made.

In the anterior insula [−38, 24, 16], we found that the activations were negatively correlated with EV in the decision task (Fig. 5) and that these activations started during the decision period. Thus, the new finding is that the anterior insula is activated by relatively low but still positive expected outcome or gain. This complements an earlier finding that the anterior insula is activated by loss prediction (Paulus et al. 2003). When we investigated the basis for the activations in the anterior insula further, we found that the activations were correlated with the variance of the reward obtained for each choice made in the task. High variance corresponds to high uncertainty about the magnitude of the reward that will be obtained on a particular trial, and the high uncertainty may be unpleasant. It is also known that part of the insula is activated by face expressions of disgust (Phillips et al. 1998, 2004), at, for example, $Y = 6$, which is at the posterior border of the activations we found. Thus, more activation of this part of the insula appears to be produced by unpleasant states, such as when making decisions and working for low compared with high EVs or when seeing disgust face expressions. Autonomic activity is also known to correlate with activations in parts of the insula (e.g., at $Y = 24$ mm anterior) (Nagai et al. 2004), and this could be part of the same pattern of responses to some unpleasant stimuli.

The design of this study also enabled analysis of how TD reward prediction error-related signals in neuroimaging investigations might be related to other representations (such as EV and RM) that might be present to different extents in different brain areas. (The TD error was based on the difference between the EV and the RM [or outcome] for that trial at time $t = 10$ s [see Fig. 1b] and was calculated over trials 10–30 as the desirability of the action—or expected utility—of choosing on the right was a function of the EV, as shown in Fig. 2b, and was stable in these trials, as shown at the start of the Results). In the ventral striatum, the activations were positively correlated with the reward actually obtained on that trial (the RM or outcome) but not in at least some parts with the EV (see Fig. 6). Thus, the positive TD error correlation arose in the ventral striatum/nucleus accumbens because at the time that the EV period ended and the participant was informed about how much reward had been obtained on that trial, the BOLD signal changed to a higher value for large rewards, and to a lower value for low or no reward, from a value that was not a function of the EV in the earlier part of the trial (as shown in Fig. 6). These fMRI findings exemplify the fact that activation of the ventral striatum does reflect the changing expectations of reward during the trials of a task, and indeed this was shown by recordings in the macaque ventral striatum (Williams et al. 1993), which show that some neurons altered their firing rate within 170 ms of the monkey being shown a visual stimulus that indicated whether reward or saline was available on that trial (corresponding to the start of the RM period at $t = 10$ s in the present study). Given that ventral striatal neurons alter their activity when the visual stimulus is shown informing the macaque about whether reward is available on that trial, it is in fact not surprising that the fMRI correlation analyses do pick up signals during trials that can be interpreted as TD error signals. Whether these fMRI correlations with a TD error reflect more than the activity of neurons that respond to a signal that informs the subject about the reward outcome (magnitude) to be obtained (equivalent in this study to the display at time $t = 10$ s which informs the subject about how much reward will be obtained on that trial) rather than a phasic TD error present only at the transition between an EV and a RM period remains to be shown.

One of the strongest fMRI activations that is found in the ventral striatum is produced by TD reward/punishment prediction errors in Pavlovian (i.e., classical conditioning) tasks (McClure et al. 2003; O'Doherty et al. 2003, 2006; Seymour et al. 2004). For example, in a classical conditioning task, a first visual stimulus probabilistically predicted high versus low pain and a second visual stimulus perfectly predicted whether the pain would be high or low on that trial. Activation of the ventral striatum and a part of the insula was related to the TD error, which arose, for example, at the transition between the first and second visual stimulus if the first visual stimulus had predicted low pain but the second informed the subject that the pain would be high (Seymour et al. 2004). In the study described here, we have extended the TD error approach beyond classical conditioning to a monetary reward decision task and have

shown how TD signals may be related to some of the underlying processing. In an instrumental juice reward task, it has been found previously that activation in the ventral striatum is correlated with reward prediction error (O'Doherty et al. 2004), but that study did not analyze how this correlation might arise as a consequence of a correlation with RM but not with EV, as illustrated in Figure 4.

A positive TD error correlation was also found in the left inferior frontal gyrus (in or near cortical area 44, Broca's area) (as shown in Fig. 7), but here the TD correlation arose because the activation became low when the subject was informed that no reward was obtained on a trial, and it appeared that the area was activated especially when the decision was difficult, in that it was between 2 approximately equal values of the EV (EV9 on the left vs. EV10 on the right). (In particular, EV10 is a more difficult choice than EV27 and EV5, and the activations are greater for EV10 than EV27 and EV5, as shown in Fig. 7). As this area was activated especially when the decision was difficult, between 2 approximately equal EVs, we believe that its activation may be related to the engagement of planning using verbal processing to supplement more direct decision making based on direct (implicit) estimates of rewards and punishers by an emotional system (Rolls 2005).

In a part of the midbrain near the dopamine neurons at [14, −20, −16], there was also a positive correlation with the TD error [14, −20, −16] ($P$ = 0.032 SVC) (also found in classical conditioning; O'Doherty et al. 2006), but here this was related to a negative correlation between the BOLD signal in the EV period of each trial and the EV. (The TD error was thus positive, e.g., whatever reward became available if it was a low EV trial set for choices on the right; and the TD error was negative if it was a high EV trial set).

In some brain regions, a negative correlation with TD error was found. For example, a negative correlation with TD error was found in the anterior cingulate cortex [8, 36, 42] in a dorsal part, and this could be a region that is activated when an anticipated reward is not obtained. Consistent with this, a nearby region of the dorsal anterior cingulate cortex is activated on reversal trials (when expected rewards are not being obtained) in a reversal task (Kringelbach and Rolls 2003) and in a number of other error-producing tasks (Rushworth et al. 2004).

In conclusion, the principal findings of this decision-making study include the following. First, in decision making, some areas, such as the medial orbitofrontal cortex, represent both RM and EV (as shown, e.g., by the conjunction analysis, Fig. 3c). That is, the medial orbitofrontal cortex may take into account the probability as well as the magnitude of a reward to represent EV. To compute this, it would need to know the RM associated with a given choice, and this is represented in the same region, as shown here. The mechanism is suggested by the finding that neurons in the orbitofrontal cortex learn to respond to visual stimuli associated with rewards (Thorpe et al. 1983; Rolls et al. 1996) and represent the RM of these stimuli (as shown, e.g., by reward devaluation including sensory-specific satiety studies) (Critchley and Rolls 1996; Rolls 2005, 2008), and these learned representations could be influenced by the probability of the reward. These activations reflect computations that are causally important in probabilistic decision making, for patients with damage to the orbitofrontal cortex perform poorly when the contingencies change in a probabilistic monetary reward/loss decision task (Hornak et al. 2004). Second, different parts of the ventral striatum had activations related to RM and to EV, and

these could reflect inputs from the orbitofrontal cortex (Rolls 2005). Third, the anterior insula has more activation when subjects are making choices with a relatively low (though still positive) EV, and this appears to be consistent with findings in other studies that similar regions of the insula are activated by predicted loss, expressions of disgust, and also in relation to autonomic activity. We also show that the anterior insula is activated under conditions of reward uncertainty, that is, when the RM variance for a particular choice is high. Fourth, a part of the inferior frontal gyrus (area 44) was activated especially when the decision was difficult, that is, when the subject was making a choice between an EV of 9 pence and 10 pence, and this activation could reflect engagement of a more calculating than emotional basis for the decision making (Rolls 2005). Fifth, TD error-related activations in functional neuroimaging were illuminated by showing how a given brain area represents EV, RM, or both. Finally, as shown in Figure 2b, the desirability of an action (the expected utility in microeconomics) does reflect the EV in this decision-making experiment, and thus, the activations described as being related to EV in this investigation are in fact also related to the desirability of the choices being made in this decision task.

Overall, the results show how some brain areas (such as the medial orbitofrontal cortex) not only respond to reward outcomes (i.e., RM) but also respond to a prediction that reflects the probability of how much reward will be obtained for a particular choice in a monetary reward decision task (i.e., EV). Both signals are crucial for decision making and could contribute to the calculation of reward prediction (or TD) error useful in reinforcement learning. The results also show how activations in some other brain areas such as the ventral striatum and midbrain are correlated in a probabilistic decision-making task with TD error and how activations in the insula may also be related to decision making by reflecting low EV.

## Notes

*Conflict of Interest*: None declared.

Address correspondence to email: edmund.rolls@psy.ox.ac.uk; web address: http://www.cns.ox.ac.uk.

## References

Abler B, Walter H, Erk S, Kammerer H, Spitzer M. 2006. Prediction error as a linear function of reward probability is coded in human nucleus accumbens. Neuroimage. 31:790-795.

Bechara A, Damasio H, Damasio AR, Lee GP. 1999. Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. J Neurosci. 19:5473-5481.

Bernoulli J. 1738/1954. Exposition of a new theory on the measurement of risk. Econometrica. 22:23-36.

Breiter HC, Aharon I, Kahneman D, Dale A, Shizgal P. 2001. Functional imaging of neural responses to expectancy and experience of monetary gains and losses. Neuron. 30:619-639.

Collins DL, Neelin P, Peters TM, Evans AC. 1994. Automatic 3D intersubject registration of MR volumetric data in standardized talairach space. J Comput Assist Tomogr. 18:192-205.

Critchley HD, Rolls ET. 1996. Hunger and satiety modify the responses of olfactory and visual neurons in the primate orbitofrontal cortex. J Neurophysiol. 75:1673-1686.

Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ. 2006. Cortical substrates for exploratory decisions in humans. Nature. 441:876-879.

Dayan P, Abbott LF. 2001. Theoretical neuroscience. Cambridge: MIT Press.

de Araujo IET, Kringelbach ML, Rolls ET, Hobden P. 2003. The representation of umami taste in the human brain. J Neurophysiol. 90:313–319.

de Araujo IET, Kringelbach ML, Rolls ET, McGlone F. 2003. Human cortical responses to water in the mouth, and the effects of thirst. J Neurophysiol. 90:1865–1876.

de Araujo IET, Rolls ET, Kringelbach ML, McGlone F, Phillips N. 2003. Taste-olfactory convergence, and the representation of the pleasantness of flavour, in the human brain. Eur J Neurosci. 18:2374–2390.

de Araujo IET, Rolls ET, Velazco MI, Margot C, Cayeux I. 2005. Cognitive modulation of olfactory processing. Neuron. 46:671–679.

Dorris MC, Glimcher PW. 2004. Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. Neuron. 44:365–378.

Doya K. 1999. What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? Neural Netw. 12:961–974.

Fellows LK, Farah MJ. 2003. Ventromedial frontal cortex mediates affective shifting in humans: evidence from a reversal learning paradigm. Brain. 126:1830–1837.

Fiorillo CD, Tobler PN, Schultz W. 2003. Discrete coding of reward probability and uncertainty by dopamine neurons. Science. 299:1898–1902.

Friston KJ, Glaser DE, Henson RN, Kiebel S, Phillips C, Ashburner J. 2002. Classical and Bayesian inference in neuroimaging: applications. Neuroimage. 16:484–512.

Friston KJ, Worsley KJ, Frackowiak RSJ, Mazziotta JC, Evans AC. 1994. Assessing the significance of focal activations using their spatial extent. Hum Brain Mapp. 1:214–220.

Gintis H. 2000. Game theory evolving. Princeton (NJ): Princeton University Press.

Glimcher P. 2004. Decisions, uncertainty, and the brain. Cambridge: MIT Press.

Glimcher PW. 2003. The neurobiology of visual-saccadic decision making. Annu Rev Neurosci. 26:133–179.

Hornak J, O'Doherty J, Bramham J, Rolls ET, Morris RG, Bullock PR, Polkey CE. 2004. Reward-related reversal learning after surgical excisions in orbitofrontal and dorsolateral prefrontal cortex in humans. J Cogn Neurosci. 16:463–478.

Kahneman D, Tversky A. 1979. Prospect theory: an analysis of decision under risk. Econometrica. 47:263–292.

Kahneman D, Tversky A. 1984. Choices, values, and frames. Am Psychol. 4:341–350.

Kiebel SJ, Poline JB, Friston KJ, Holmes AP, Worsley KJ. 1999. Robust smoothness estimation in statistical parametric maps using standardized residuals from the general linear model. Neuroimage. 10:756–766.

Kim H, Shimojo S, O'Doherty JP. 2006. Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. PLoS Biol. 4:e233.

Knutson B, Fong GW, Bennett SM, Adams CM, Hommer D. 2003. A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fMRI. Neuroimage. 18:263–272.

Knutson B, Taylor J, Kaufman M, Peterson R, Glover G. 2005. Distributed neural representation of expected value. J Neurosci. 25:4806–4812.

Kringelbach ML, O'Doherty J, Rolls ET, Andrews C. 2003. Activation of the human orbitofrontal cortex to a liquid food stimulus is correlated with its subjective pleasantness. Cereb Cortex. 13:1064–1071.

Kringelbach ML, Rolls ET. 2003. Neural correlates of rapid reversal learning in a simple model of human social interaction. Neuroimage. 20:1371–1383.

Kringelbach ML, Rolls ET. 2004. The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. Prog Neurobiol. 72:341–372.

McCabe C, Rolls ET. 2007. Umami: a delicious flavor formed by convergence of taste and olfactory pathways in the human brain. Eur J Neurosci. 25:1855–1864.

McClure SM, Berns GS, Montague PR. 2003. Temporal prediction errors in a passive learning task activate human striatum. Neuron. 38:339–346.

McCoy AN, Platt ML. 2005. Expectations and outcomes: decision-making in the primate brain. J Comp Physiol A. 191:201–211.

Nagai Y, Critchley HD, Featherstone E, Trimble MR, Dolan RJ. 2004. Activity in ventromedial prefrontal cortex covaries with sympathetic skin conductance level: a physiological account of a "default mode" of brain function. Neuroimage. 22:243–251.

O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ. 2004. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. Science. 304:452–454.

O'Doherty J, Kringelbach ML, Rolls ET, Hornak J, Andrews C. 2001. Abstract reward and punishment representations in the human orbitofrontal cortex. Nat Neurosci. 4:95–102.

O'Doherty J, Rolls ET, Francis S, Bowtell R, McGlone F. 2001. The representation of pleasant and aversive taste in the human brain. J Neurophysiol. 85:1315–1321.

O'Doherty J, Rolls ET, Francis S, Bowtell R, McGlone F, Kobal G, Renner B, Ahne G. 2000. Sensory-specific satiety related olfactory activation of the human orbitofrontal cortex. Neuroreport. 11:893–897.

O'Doherty J, Winston J, Critchley H, Perrett D, Burt DM, Dolan RJ. 2003. Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. Neuropsychologia. 41:147–155.

O'Doherty JP, Buchanan TW, Seymour B, Dolan RJ. 2006. Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. Neuron. 49:157–166.

O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ. 2003. Temporal difference models and reward-related learning in the human brain. Neuron. 38:329–337.

O'Doherty JP, Deichmann R, Critchley HD, Dolan RJ. 2002. Neural responses during anticipation of a primary taste reward. Neuron. 33:815–826.

Paulus MP, Rogalsky C, Simmons A, Feinstein JS, Stein MB. 2003. Increased activation in the right insula during risk-taking decision making is related to harm avoidance and neuroticism. Neuroimage. 19:1439–1448.

Phillips ML, Williams LM, Heining M, Herba CM, Russell T, Andrew C, Bullmore ET, Brammer MJ, Williams SC, Morgan M, et al. 2004. Differential neural responses to overt and covert presentations of facial expressions of fear and disgust. Neuroimage. 21:1484–1496.

Phillips ML, Young AW, Scott SK, Calder AJ, Andrew C, Giampietro V, Williams SC, Bullmore ET, Brammer M, Gray JA. 1998. Neural responses to facial and vocal expressions of fear and disgust. Proc R Soc Lond B Biol Sci. 265:1809–1817.

Platt ML, Glimcher PW. 1999. Neural correlates of decision variables in parietal cortex. Nature. 400:233–238.

Rogers RD, Ramnani N, Mackay C, Wilson JL, Jezzard P, Carter CS, Smith SM. 2004. Distinct portions of anterior cingulate cortex and medial prefrontal cortex are activated by reward processing in separable phases of decision-making cognition. Biol Psychiatry. 55:594–602.

Rolls ET. 2004. The functions of the orbitofrontal cortex. Brain Cogn. 55:11–29.

Rolls ET. 2005. Emotion explained. Oxford: Oxford University Press.

Rolls ET. 2006. Brain mechanisms underlying flavour and appetite. Phil Trans R Soc B. 361:1123–1136.

Rolls ET. 2008. Memory, attention, and decision-making: a unifying computational neuroscience approach. Oxford: Oxford University Press.

Rolls ET, Baylis LL. 1994. Gustatory, olfactory, and visual convergence within the primate orbitofrontal cortex. J Neurosci. 14:5437–5452.

Rolls ET, Critchley HD, Mason R, Wakeman EA. 1996. Orbitofrontal cortex neurons: role in olfactory and visual association learning. J Neurophysiol. 75:1970–1981.

Rolls ET, Hornak J, Wade D, McGrath J. 1994. Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. J Neurol Neurosurg Psychiatry. 57:1518–1524.

Rolls ET, Kringelbach ML, de Araujo IET. 2003. Different representations of pleasant and unpleasant odors in the human brain. Eur J Neurosci. 18:695–703.

Rolls ET, Sienkiewicz ZJ, Yaxley S. 1989. Hunger modulates the responses to gustatory stimuli of single neurons in the caudolateral orbitofrontal cortex of the macaque monkey. Eur J Neurosci. 1:53–60.

Rushworth MF, Walton ME, Kennerley SW, Bannerman DM. 2004. Action sets and decisions in the medial frontal cortex. Trends Cogn Sci. 8:410–417.

Schultz W. 2006. Behavioral theories and the neurophysiology of reward. Annu Rev Psychol. 57:87–115.

Seymour B, O'Doherty JP, Dayan P, Koltzenburg M, Jones AK, Dolan RJ, Friston KJ, Frackowiak RS. 2004. Temporal difference models describe higher-order learning in humans. Nature. 429:664–667.

Sugrue LP, Corrado GS, Newsome WT. 2004. Matching behavior and the representation of value in the parietal cortex. Science. 304:1782–1787.

Sugrue LP, Corrado GS, Newsome WT. 2005. Choosing the greater of two goods: neural currencies for valuation and decision making. Nat Rev Neurosci. 6:363–375.

Sutton RS, Barto AG. 1998. Reinforcement learning. Cambridge: MIT Press.

Thorpe SJ, Rolls ET, Maddison S. 1983. Neuronal activity in the orbitofrontal cortex of the behaving monkey. Exp Brain Res. 49:93–115.

Tobler PN, O'Doherty JP, Dolan RJ, Schultz W. 2007. Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. J Neurophysiol. 97:1621–1632.

Tversky A, Kahneman D. 1986. Rational choice and the framing of decisions. J Bus. 59:251–278.

von Neumann J, Morgenstern O. 1944. The theory of games and economic behavior. Princeton (NJ): Princeton University Press.

Williams GV, Rolls ET, Leonard CM, Stern C. 1993. Neuronal responses in the ventral striatum of the behaving macaque. Behav Brain Res. 55:243–252.

Wilson JL, Jenkinson M, Araujo IET, Kringelbach ML, Rolls ET, Jezzard P. 2002. Fast, fully automated global and local magnetic field optimisation for fMRI of the human brain. Neuroimage. 17:967–976.

Worsley KJ, Marrett P, Neelin AC, Friston KJ, Evans AC. 1996. A unified statistical approach for determining significant signals in images of cerebral activation. Hum Brain Mapp. 4:58–73.

Yacubian J, Glascher J, Schroeder K, Sommer T, Braus DF, Buchel C. 2006. Dissociable systems for gain- and loss-related value predictions and errors of prediction in the human brain. J Neurosci. 26:9530–9537.