Behavioral/Systems/Cognitive

Distinct Roles of Three Frontal Cortical Areas in Reward-Guided Behavior

M. P. Noonan,[1,2] **R. B. Mars,**[1,2] **and M. F. S. Rushworth**[1,2]

[1]Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, United Kingdom, and [2]Centre for Functional Magnetic Resonance Imaging of the Brain, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU, United Kingdom

Functional magnetic resonance imaging was used to measure activity in three frontal cortical areas, the lateral orbitofrontal cortex (lOFC), medial orbitofrontal cortex (mOFC)/ventromedial frontal cortex (vmPFC), and anterior cingulate cortex (ACC), when expectations about type of reward, and not just reward presence or absence, could be learned. Two groups of human subjects learned 12 stimulus–response pairings. In one group (Consistent), correct performances of a given pairing were always reinforced with a specific reward outcome, whereas in the other group (Inconsistent), correct performances were reinforced with randomly selected rewards. The mOFC/vmPFC and lOFC were not distinguished by simple differences in relative preference for positive and negative outcomes. Instead lOFC activity reflected updating of reward-related associations specific to reward type; lOFC was active whenever informative outcomes allowed updating of reward-related associations, regardless of whether the outcomes were positive or negative, and the effects were greater when consistent stimulus-outcome and response-outcome mappings were present. A psychophysiological interaction analysis demonstrated changed coupling between lOFC and brain areas for visual object representation, such as perirhinal cortex, and reward-guided learning, such as the amygdala, ventral striatum, and habenula/mediodorsal thalamus. In contrast, mOFC/vmPFC activity reflected expected values of outcomes and occurrence of positive outcomes, regardless of consistency of outcome mappings. The third frontal cortical region, the ACC, reflected the use of reward type information to guide response selection. ACC activity reflected the probability of selecting the correct response, was greater when consistent outcome mappings were present, and was related to individual differences in propensity to select the correct response.

Introduction

Several associations guide response selection. Emphasis is often placed on those between stimuli and responses (S–R association). Ventrolateral prefrontal cortex (vlPFC) is implicated in learning arbitrary S–R associations (Wise and Murray, 2000; Bunge and Wallis, 2008). If, however, a stimulus is indicative of availability of only one reward outcome type, then an organism may also learn associations between that stimulus and outcome (S–O association) and between the outcome and response to be made (O–R association). When each stimulus is associated with different outcomes, then response selection can occur either via direct S–R association or via an indirect S–O association that elicits a representation of a particular outcome that in turn guides selection of the response via an O–R association.

The presence of S–O and O–R associations is thought to underlie the differential outcomes effect (DOE), the facilitation of learning in the presence of different outcomes, in monkeys and

rats (Jones and White, 1994; Savage, 2001; Easton and Gaffan, 2002). Whether such effects occur in adult humans has been unclear (Easton, 2004; Estevez et al., 2007) and may depend on task difficulty (Plaza et al., 2011). Our first aim was, therefore, to test whether a DOE exists in humans.

A task in which type of reward outcome, and not just presence or amount of reward, could influence action selection might also help distinguish the functions of three frontal areas, the anterior cingulate cortex (ACC), medial orbitofrontal cortex (mOFC)/ventromedial frontal cortex (vmPFC), and lateral orbitofrontal cortex (lOFC), whose roles in reward-guided learning and decision making remain unclear (Hare et al., 2008; Rangel and Hare, 2010). An influential account holds that lOFC and mOFC/vmPFC are distinguished by responsiveness to positive and negative outcomes, respectively (Kringelbach and Rolls, 2004). Recently, however, it has been suggested that macaque lOFC and mOFC/vmPFC differ in other ways (Noonan et al., 2010; Walton et al., 2010; Rudebeck and Murray, 2011; Rushworth et al., 2011); lOFC is critical when learning and assigning credit for reward (or error) occurrence to a specific stimulus, whereas mOFC/vmPFC is important when representing expected outcome values to guide choice. Our second aim was, therefore, to examine whether lOFC might also be implicated in learning associations between specific stimuli and specific types of reward. Functional magnetic resonance imaging (fMRI) scans were collected while two groups of subjects learned to select responses either in the context of consistent differential or nondifferential outcomes (referred to as

the "Consistent" and "Inconsistent" groups, respectively). If lOFC is important for learning associations between specific stimuli and outcome types, then lOFC activity will be greater in the Consistent group whenever an outcome is delivered that informs subjects about those associations.

The differential outcome procedure allowed testing not only whether mOFC/vmPFC activity reflected how informative an outcome was for updating reward-related associations, but whether it reflected the values of outcomes (Plassmann et al., 2007; Lebreton et al., 2009). Finally the role of ACC in reward-guided action selection was tested. If ACC is critical for reward-guided action selection (Rudebeck et al., 2008), then its activity should reflect the changing probability of selecting correct responses during learning in the context of the DOE.

Materials and Methods

Subjects

All subjects gave informed consent to participate in the investigation, which was approved by the Central Office for Research Ethics Committee (reference number 05/Q1606/96). Thirty-six right-handed subjects, 16 of whom were men, completed the experiment. Half of the subjects ($n = 18$) were randomly assigned to the Consistent condition, and half ($n = 18$) to the Inconsistent condition. Collectively, the two groups had a mean age (and SD) of 25.14 (4.20) years. All had normal or corrected-to-normal vision and indicated no family history of psychiatric or neurological disease.

Task and procedure

The aim of the experiment was to compare learning of S–R pairings in two situations. In one situation, it was intended that response selection could be mediated only via learned S–R associations. In the other situation it was intended that response selection could be mediated via learned S–O and O–R associations in addition to S–R associations. Response selection via S–O and O–R associations is possible if consistent S–O and O–R mappings exist, but it is impossible if such mappings do not exist. We therefore trained two groups of subjects on similar S–R pairing tasks that either contained or omitted consistent S–O and O–R mappings. The groups are therefore referred to as the Consistent and Inconsistent groups, respectively. Although the focus of the report is on the learning of the S–R pairing task when MRI scans were collected (Stage 5, described below) it was first necessary to give the subjects preliminary experience in learning stimulus–reward outcome and response–reward outcome associations (Stages 1, 2, and 4, described below). By comparing S–R pair learning (at Stages 3 and 5) in the two groups, we were able to test for the existence of any DOE. Training for the task was administered in several stages on the day before MRI scanning.

Task designs are shown graphically in Figure 1. In addition, conditional group training schedules are summarized in Figure 2. In brief, Day 1 consisted of four discrete blocks of behavioral training: subjects were first taught stimulus–outcome associations (Stage 1), then response–outcome associations (Stage 2), and then were retested on the stimulus–reward associations (Stage 1R), before finally trying to learn correct S–R pairs (Stage 3). On Day 2, subjects were first retested outside of the scanner on the S–R pairs learned during Stage 3 before MRI scanning (Stage 3 recall). In addition they were taught new stimulus–reward associations (Stage 4). In the scanner (Stage 5), they were then required both to learn a new set of S–R pairs using novel stimuli ("New" learning task) and also to recall the previous S–R associations ("Old" recall task).

Reinforcement outcomes took the form of visually presented pictures of gift tokens that were converted into a payment of the corresponding gift vouchers at the end of the experiment. Each type of gift token could be used for making purchases at a particular type of retail outlet, for example, a store primarily selling music and videos, a store primarily selling books, a grocery store, or a café. Each subject was asked to rate and choose four distinct gift vouchers from a set of 12, and these four stimuli were then used as reward outcomes for that subject in the subsequent experiment.

Stage 1: stimulus-outcome learning task

The first stage of the experiment for subjects was a concurrent visual discrimination learning task consisting of 12 pairs of visual stimuli presented on a computer screen. Each stimulus pair was presented 11 times so that the number of trials in this phase was 264. One stimulus in each pair was designated the correct option and choices of it were rewarded with presentation of a reward token. Stimuli were presented on screen for 1 s, and subjects had this time and an additional 500 ms in which to respond. Subjects made stimulus selections with two fingers on their left hand on adjacent keys on a keyboard. Correct choices of a given stimulus were always rewarded with a particular token outcome for subjects assigned to the Consistent group. In the Inconsistent group, however, correct choices were rewarded with one of the four tokens selected at random (Fig. 1A). At the end of the final stage of the experiment (see below), the subjects were compensated for participating in the experiment by receiving actual tokens of the type indicated during task performance. If a subject selected the incorrect target or responded outside the time window, a red cross would appear on screen. Feedback was presented for 1.5 s, after which the screen was blank for an intertrial interval of 1.5 s. On completion of the 264 trials of Stage 1, subjects in both the Consistent and Inconsistent groups had learned 12 visual discrimination problems and had learned that 12 visual stimuli were associated with reward. Only subjects in the Consistent group, however, would have been able to form an association between each stimulus and one of four types of reward.

Stage 2: response-outcome learning task

In the second stage of the experiment, subjects explored four button press responses, which were either paired consistently with specific token reward outcomes (Consistent subjects), or random (Inconsistent subjects) token reward outcomes (Fig. 1B). Options (R1, R2, R3, and R4) were represented as four red squares on the computer screen for 1 s and corresponded to four adjacent keys on a keyboard, which the subjects could select with one of four fingers of their right hand. Subjects had this time and an additional 500 ms in which to respond. To ensure that subjects explored all four buttons, different numbers and buttons were indicated as "active" for selection. If a button was active, a yellow square outline surrounded the corresponding red square on the computer. Selection of a response option led to the presentation of a gift voucher token. If a subject happened to select a nonactive button they would be presented with the gift voucher feedback, but it would not be added to their voucher earnings. If a subject responded outside the time window, a red cross would appear on screen. Feedback was presented for 1.5 s, after which the screen was blank for an intertrial interval of 1.5 s. There were a total of 96 trials in this experimental stage. On completion of Stage 2, subjects in both Consistent and Inconsistent groups had learned that all four actions were rewarded when made, but only subjects in the Consistent mapping group would have been able to form an association between each response and one of the four types of reward.

Stage 1R: reminder of Stage 1

Subjects then completed 24 reminder trials of the S–O learning trials (learned at Stage 1) before proceeding to Stage 3.

Stage 3: stimulus–response learning task

Subjects then learned to select one of the four responses from Stage 2 each time they saw one of the rewarding stimuli from Stage 1 (Fig. 1C). Individual stimuli were presented on screen for 1 s, and subjects had this time and an additional 500 ms in which to respond. Choice of the correct response for a given S–R pairing was either consistently rewarded with a given reward token outcome (Consistent condition) or randomly rewarded with any of the four possible reward token outcomes (Inconsistent condition). Feedback was presented for 1.5 s, after which the screen was blank for an intertrial interval of 1.5 s. For subjects in the Consistent condition, a given stimulus was always rewarded with the same token outcome both at Stage 1 and at Stage 3, and a given response was always rewarded with the same token outcome both at Stage 2 and Stage 3. In other words, the component parts of a correctly performed S–R pair that was rewarded by, for example, a Blackwell's book token at Stage 3 would also both have been rewarded by a Blackwell's book token in Stages 1 and
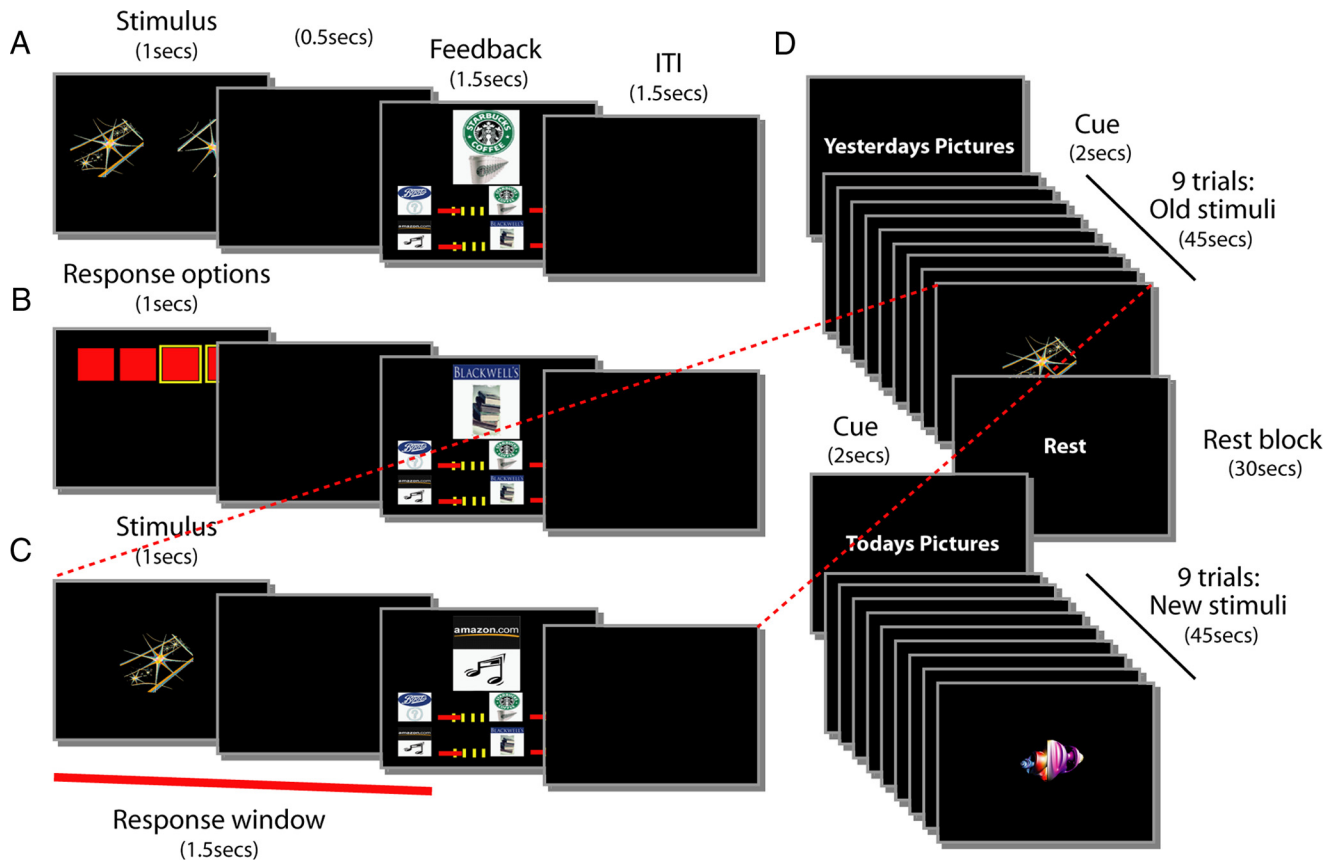
Figure 1. ***A–C***, Schematic representation of the trial events for each stage of task learning and performance. Timing of events was analogous for all phases of the experiment. ***A***, S–O learning task (Stage 1). A pair of stimuli were presented on the left and right of the screen for 1500 ms on each trial, and subjects attempted to identify the stimulus associated with reward. Subjects pressed one of two buttons that corresponded to the two sides of the screen. Subjects had to respond within 500 ms of stimulus offset. The presentation of a token constituted the outcome on correctly performed trials, while the presentation of a red cross indicated an error either of choice or timing. It was presented above four bars that reflected the amount of each token type that had been accumulated. The feedback screen was visible for 1500 ms and was then immediately replaced by an intertrial interval (ITI) of 1500 ms. ***B***, Response-outcome learning (Stage 2). Four red squares corresponded to the positions of four possible button responses subjects could make. The timing was similar to that used in Stage 1; initial visual presentation of the squares lasted 1500 ms on each trial, and subjects had to respond within 500 ms of stimulus offset. As in Stage 1, the presentation of a token constituted the outcome on correctly performed trials, while the presentation of a red cross indicated an error of timing. It was presented above four bars that reflected the amount of each token type accumulated. As in Stage 1, the feedback screen was visible for 1500 ms and then immediately followed by an ITI of 1500 ms. ***C***, S–R pair learning task (Stage 3). A single stimulus was presented and subjects had to identify which of the four possible responses was rewarded in association with each stimulus. As in Stages 1 and 2, stimulus presentation lasted 1500 ms, and subjects had to respond within 500 ms of stimulus offset; presentation of a token constituted the outcome on correctly performed trials. The presentation of a red cross indicated an error either of choice or timing. Tokens were presented above four bars that reflected the amount of each token type accumulated. As before, the feedback screen was visible for 1500 ms and then immediately followed by an intertrial interval (ITI) of 1500 ms. ***D***, Schematic representation of the order of events in the fMRI phase of the experiment (Stage 5). Subjects learned a new set of S–R pairs (New learning task) and performed a previously learned set of S–R pairs (Old recall task) in a pseudorandom order in miniblocks of nine trials. Each miniblock lasted 45 s and was cued for 2 s with "Yesterday's Pictures" or "Today's Pictures" for old and new stimuli, respectively. Thirty second rest blocks were also interspersed within the experiment. All other events and durations were the same as in Stage 3 (***C***). A given stimulus or response was always associated with a specific outcome on correctly performed trials in the Consistent group. In the Inconsistent group, correctly performed trials were reward with a randomly selected outcome from the set of four possible outcomes.

2. For subjects in the Inconsistent group, response selection could be achieved only via a direct S–R association. For subjects in the Consistent group, response selection could also be based on an indirect link mediated by an S–O association and an O–R association. The S–R learning task ended when a criterion of 95% correct performance over 24 trials was reached.

Stage 3 recall: old stimulus–response pair task recall
The following day (which was the day of MRI scanning), subjects returned to the laboratory and were given an opportunity to remind themselves of the task they had learned in Stage 3. They performed sufficient trials to reach the same criterion level of performance, 95% correct for 24 trials, using the same stimulus and response pairings that had been learned on the previous day.

Stage 4: new stimulus sets
Subjects were then also asked to learn a new version of the Stage 1 task involving a new set of visual stimuli but the same reward outcomes. Once again, only the Consistent group subjects learned consistent outcome

mappings. The subjects also explored response-outcome associations in a repeated version of the Stage 2 task involving the same response buttons and, for Consistent group subjects, the same O–R associations. The subjects therefore had repeated sessions of Stages 1 and 2 to be ready to learn a new version of the S–R pairing task on entering the MRI scanner in the final stage of the task (Stage 5 below).

Stage 5: scanning
The final stage of the experiment was performed in a Siemens 3T scanner. Miniblocks of the previously learned S–R pairing task (Old recall task, learned previously in Stage 3) were interspersed pseudorandomly with blocks of a new S–R pairing task that used the new stimuli learned at Stage 4 (New learning task) and blocks of rest (Fig. 1*D*). Each experimental "miniblock" consisted of nine stimuli and lasted in total 40.5 s. Subjects were therefore required to learn the new S–R pairings, by trial-and-error, when the scans were taken. During the S–R learning phase, as before, reward identities were either consistent for a given S–R pair (Consistent group) or random (Inconsistent group). Therefore, although both groups
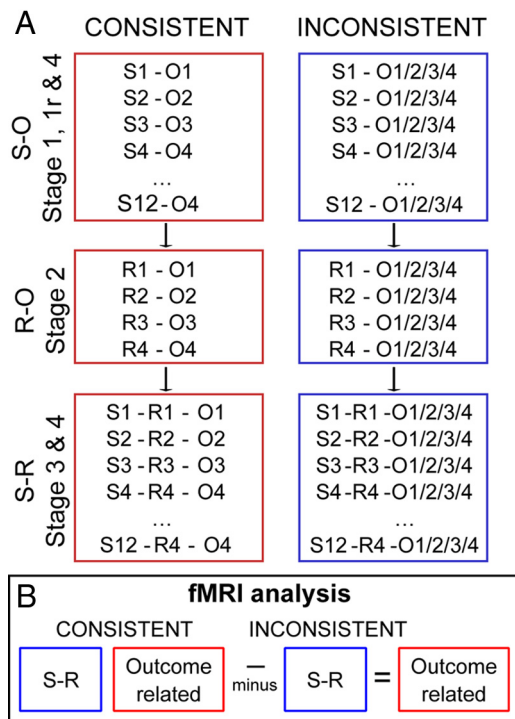
A

Figure 2. *A*, In each phase of the experiment, subjects had the opportunity to earn reward tokens for correct responses. In the Consistent group, each correct stimulus selection (Stages 1 and 1R) or response selection (Stages 2, 3, and 5) was always rewarded with the same one of four gift voucher outcomes. In contrast, when Inconsistent group subjects were rewarded, the token outcome was selected randomly. S1–S12 indicate the 12 different stimuli associated with reward. R1, R2, R3, and R4 indicate the four responses. O1, O2, O3, and O4 indicate the four token outcomes. *B*, Schematic representation of the experimental design. Both Consistent and Inconsistent group subjects could learn S–R associations (blue box). However, critically, only the Consistent group subjects had access to outcome-specific expectations (red box) that can also be exploited during action selection. By subtracting fMRI effects of interest, for example, fMRI effects associated with error occurrence, in the Inconsistent group from the same effects in the Consistent group, it is possible to isolate outcome-specific effects.

had the opportunity to learn S–R associations during Stage 5, only the Consistent group subjects could make use of outcome-specific learning mechanisms because only they had acquired S–O and O–R associations in Stages 1 and 2. Stimulus event timings were the same as the previous stages of learning. Before each experimental miniblock began, the subject was given a cue indicating the next condition: "today's pictures" (New learning task) or "yesterday's pictures" (Old recall task). There were a total of 216 trials per experimental condition. Rest blocks were 30 s in length.

Behavioral data acquisition

Stimuli were presented using Presentation (Neurobehavioral Systems; version 0.53, build 12.05.02) on a Windows 2007 operating system on a MacBook computer (Apple).

Image acquisition

Blood oxygenation level-dependent (BOLD) fMRI images and T1-weighted anatomical images were acquired on a 3T Siemens TRIO MR scanner with a maximum gradient strength of 40 mT · m^{-1} at the Oxford University Centre for Clinical Magnetic Resonance Imaging.

BOLD fMRI data were acquired with a voxel resolution of $3 \times 3 \times 3$ mm^3, a repetition time (TR) of 3 s, an echo time (TE) of 30 ms, and a flip angle of 87°. The slice angle was set to 15°, and a local z-shim was applied around the orbitofrontal cortex to minimize signal dropout in this region (Deichmann et al., 2003), which has been implicated previously in other aspects of decision making. The mean number of volumes acquired was 935, giving a mean total experiment time of ~47

min. T1-weighted structural images were also acquired for subject alignment using an MPRAGE sequence with the following parameters: voxel resolution $1 \times 1 \times 1$ mm^3 on a $176 \times 192 \times 193$ grid; TE, 4.53 ms; inversion time, 900 ms; TR, 2200 ms.

Behavioral data analysis

Accuracy data were analyzed using Excel (Microsoft Office Excel 2003), Matlab 6.5 (MathWorks), and SPSS software (version 14.0). In general, all analyses were conducted with standard parametric ANOVAs, that used the Huynh–Feldt correction where appropriate, and standard independent sample *t* tests (Figs. 3–9).

Stages 1–3

The number of correct trials performed by the Consistent and Inconsistent groups in the Stage 1 S–O learning phase was compared in an independent sample *t* test. Response choice was analyzed for the Stage 2 response-outcome phase using a repeated-measures one-way ANOVA with response (four levels; four responses) and the between-subjects factor of group (two levels, Consistent and Inconsistent). The influence of the DOE on Stage 3 S–R pair learning was tested by comparing the percentage of trials performed correctly for the first three blocks, each consisting of 36 trials (i.e., four miniblocks), by the Consistent and Inconsistent groups, with independent sample *t* tests.

Stage 3 recall

The retention of S–R pairings learned on Day 1 was measured by calculating the percentage of trials performed correctly for the first 24 trials of the prescanning recall session 24 h after the initial training. One-tailed independent sample *t* tests were used to compare the two groups of subjects in the Consistent and Inconsistent conditions.

Accuracy in the first phase of New task learning (S–O learning phase at Stage 4) and response choice in the second phase (response-outcome learning phase at Stage 4) were calculated in the same manner as described above for the same task phases in Stages 1 and 2 on Day 1. Performance accuracies for each of the Old and New S–R pairing tasks in the MRI scanner (Stage 5) were calculated by determining the percentage of correct trials in six blocks, each composed of 36 trials. The log-transformed averages for the Old and New tasks, for both groups of subjects, were subjected to a three-way ANOVA of task (Old, New) by time (six blocks of 36 trials), with group (Consistent, Inconsistent) as a between-subjects factor. Independent sample *t* tests were used to further investigate the effects.

Neuroimaging data analysis

Preprocessing

Preprocessing was identical for all of the following general linear model (GLM)-based analyses described below. fMRI data were analyzed using tools from the FMRIB Software Library (www.fmrib.ox.ac.uk/fsl). At the first level (single subjects), preprocessing involved several stages. Non-brain structures were removed using Brain Extraction Tool (Smith, 2002). Motion artifacts were removed using independent component analysis Melodic software (Jenkinson and Smith, 2001), and the data were spatially smoothed using a 5 mm Gaussian kernel of full width at half maximum. Low-frequency drifts were removed with high-pass temporal filtering with a 90 s cutoff. The time series data were analyzed using a GLM approach. Registration to standard space was performed using FLIRT (Jenkinson and Smith, 2001). Statistical analysis was performed in FEAT (version 5.63) using FILM with local autocorrelation correction (Woolrich et al., 2001). The hemodynamic response function was modeled as a γ function, a normalization of the probability density function of the γ distribution with zero phase, SD of 3 s, and a mean lag of 6 s.

Primary general linear model

For the primary GLM analysis, the regressors or explanatory variables (EVs) were convolved with the hemodynamic response function. The GLM also incorporated EVs that were the temporal derivatives of all the task-related EVs. Four EVs, modeled in an event-related manner, were entered into the GLM at the first level: (1) Errors were modeled from the time of onset of the incorrect feedback for 1.5 s. (2) "First correct" trials

included any trial in which the subject correctly paired a specific stimulus and associated response for the first time, or after the occurrence of at least one error trial when responding to the same stimulus. For example, consider the case of a participant making a correct response to Stimulus 1 but then making an error when they next encountered Stimulus 1 on a subsequent trial. In that case, if the subject then responded correctly on the next trial to feature Stimulus 1, then that trial would be another first correct trial and would be included in the regressor along with the original first correct trial. First correct trials were also modeled from the time of onset of the feedback for 1.5 s. (3) Probability of Correct Response selection was modeled parametrically, according to the running average of the accuracy of the previous nine condition-specific (New and Old stimuli) trials, from the time of the trial onset for 1.5 s, normalized between $+1$ and -1. (4) Every response was also modeled from the time of its onset until the beginning of the feedback epoch. Six motion correction EVs were also included in this analysis. FEAT was used to fit this model to the data to generate parameter estimates for each of the EVs and to contrast these parameter estimates against one another. The same regressors were used to identify trials from both the New learning and the Old task recall blocks; the aim of the analysis was not to make a categorical comparison between New and Old task blocks, but rather to look at brain activity associated with events such as Error outcomes and First Correct outcomes, or the Probability of Correct Response selection.

Mixed effects analyses (FLAME 1 and 2) were applied to the whole-brain group data in MNI space to generate statistical activation maps for each of the contrasts and to test for an effect of group. Group Z (Gaussianized t) statistic images were thresholded using clusters determined by $Z > 2.3$ and a corrected cluster extent significance threshold of $p = 0.05$. There were four contrasts of interest at the second level that allowed us to look at effects as a function of group membership. These were Consistent (first level contrast effects that were present in the Consistent group), Inconsistent (first level contrast effects that were present in the Inconsistent group), Consistent minus Inconsistent (first level contrast effects that were greater in the Consistent group than in the Inconsistent group), and Inconsistent minus Consistent (first level contrast effects that were greater in the Inconsistent group than the Consistent group). In summary, both groups of subjects have the ability to learn through S–R associations during the scanning stage of the experiment. Only the Consistent group subjects, however, are able additionally to exploit outcome-specific learning mechanisms in the scanning stage of the experiment. Activity associated with outcome-specific learning mechanisms can therefore be isolated by subtracting first level contrast effects in the Inconsistent subjects from the same effects in the Consistent subjects (Fig. 2B).

Additional GLM of Consistent group subjects

While the primary analysis allowed examination of the brain areas with activity that was related to how informative an outcome was for the updating of reward outcome-related associations, it did not examine the value of outcomes. Because each token outcome had a subjective value that the subjects had reported at the beginning of the experiment, it was possible to construct a regressor that identified brain activity that was related to the expected value of the outcome. A parametric value expectation-related regressor was constructed from the rating (between 0 and 10) subjects had assigned to the gift voucher token that would be received for Correct Response selection on each trial multiplied by the probability that the subject would receive that token—the probability that response selection would be correct (based on the average accuracy within each nine-trial miniblock). The regressor was modeled from the onset of the stimulus for 1.5 s. The GLM also included the same EVs as the primary analysis, but no data from the Inconsistent group were included in the analysis. Once again, the same regressors were used to identify trials from both the New learning and the Old task recall blocks.

Subjects in the Inconsistent group could not be included in the analysis because they could not anticipate the identities of reward outcomes. In theory a regressor reflecting the average value of the outcome multiplied by the Probability of Correct Response selection would have reflected the Inconsistent group subjects' outcome value expectations, but such a regressor would have been completely correlated with the Probability of Correct Response selection regressor.

Three subjects from the Consistent group were excluded from this analysis: two as a result of an experimental oversight in misplacing their reward ratings sheets and one because the subject had indicated that all their chosen gift vouchers were equally valuable. Thus, data from 15 subjects were used in this analysis.

Region of interest-based signal extraction

We anticipated using three regions of interest (ROIs) in our analyses that corresponded to the lOFC, mOFC/vmPFC, and ACC and that were selected on the basis of activation coordinates reported in previous studies. It was often unnecessary to use these predefined ROIs because the same three areas emerged in the whole-brain statistical contrast images. Nevertheless details of the ROIs are given below for the sake of completeness and because ROIs placed at "inactive" regions were sometimes used for control comparison with the results from active regions.

The lOFC ROI had an 8 mm^3 radius, and its location was based on the coordinates from the Kringelbach and Rolls (2004) meta-analysis of error outcome related activation in lOFC (-32, 42, -18). The ACC ROI also had an 8 mm^3 radius, and its location was based on the average coordinates from a number of studies (Walton et al., 2004b; Behrens et al., 2007; Behrens et al., 2008; Croxson et al., 2009) conducted in this laboratory that have found reward-related activations in the ACC (-9, 21, 37) in an area that corresponds to the rostral cingulate motor zone of Picard and Strick (2001) and cluster 4 of Beckmann et al. (2009). A similarly sized 8-mm^3-radius mOFC/vmPFC ROI was prepared at the peak coordinates (6, 26, -14) in a region reported to reflect reward expectation (as opposed to reward delivery) by Smith et al. (2010). In each case, parameter effects (PEs) were to be extracted from a final 6 mm^3 cube ROI located at the coordinates from the peak group difference of the voxels that survived small volume correction. In practice, however, as mentioned above, activity was clearly identifiable in each of the three areas of interest, lOFC, mOFC/vmPFC, and ACC, for at least one of the key contrasts of interest.

Psychophysiological interaction analysis of lOFC-related activity

BOLD signals were found in the lOFC to outcomes (first correct and error outcomes) that allowed updating of reward outcome-related associations. It was greater in the Consistent than the Inconsistent group subjects. Because these results suggested a role for lOFC in learning associations between specific stimuli and reward outcomes, we hypothesized that the lOFC might have this specialized role because of its interactions with two sets of regions: (1) those concerned with visual object representation in the temporal lobe cortex and (2) subcortical regions implicated in reward-guided learning.

Psychophysiological interaction analysis and temporal lobe visual areas. The perirhinal cortex is known to be critical for visual object recognition and discrimination (Buckley and Gaffan, 1997, 1998; Gaffan et al., 2000; Murray and Richmond, 2001; Lee et al., 2006; Aggleton et al., 2010). The perirhinal cortex and visually responsive temporal lobe areas such as TE are known to be anatomically connected to OFC in the monkey (Kondo et al., 2005; Saleem et al., 2008). We therefore looked for task related changes in functional connectivity between lOFC and perirhinal cortex using the psychophysiological interaction (PPI) analysis described by Friston et al. (1997). In addition, we examined functional connectivity between lOFC and more posterior occipitotemporal areas known to be involved in object representation. The lateral occipital (LO) and midfusiform area (mFA) were chosen because of their relatively selective activation for novel, abstract objects similar to the ones that were used in the current experiment (Grill-Spector, 2003).

The visual object temporal lobe area analysis focused on (1) two 6 mm^3 cube bilateral ROIs in perirhinal cortex (MNI coordinates: 36, -16, -24 and -36, 16, -24) as identified by Lee et al. (2006) and (2) 6 mm^3 cube ROIs derived from coordinates from the meta-analysis of object recognition in areas LO and mFA in the occipitotemporal cortex from Grill-Spector (2003). Coordinates originally reported in Talairach space were converted into MNI space using Brett's (1999) Matlab conversion Tal2MNI.m script (http://imaging.mrc-cbu.cam.ac.uk/imaging/

MniTalairach). The resulting MNI coordinates were ±40, -76, -7 for the LO and ±33, -38, -20 for the mFA in the two hemispheres.

In summary, we looked at the PPI between lOFC and three occipito-temporal and temporal lobe areas concerned with visual object recognition and particularly expected to find effects in the more anterior of these, the perirhinal cortex, which has strong connections to lOFC and is known to be critical for representation of complex visual objects (such as the stimuli used in the present experiment) rather than more basic visual features (which did not identify the stimuli used in the present experiment).

PPI analysis and subcortical reward-guided learning related areas. We also looked at three subcortical regions implicated in reward-guided learning: the amygdala, ventral striatum, and habenula. The human amygdala has been implicated in reward-guided learning and decision making in neuroimaging studies (De Martino et al., 2006; Hampton et al., 2007), but in monkeys it is especially critical when S–O associations involving different types of reward are learned (Malkova et al., 1997; Baxter et al., 2000; Baxter and Murray, 2002; Izquierdo and Murray, 2007). Ventral striatal activity has been implicated in many reward-guided learning tasks, and it is responsive when reward and error feedback are provided that will allow a subject to update their reward-related associations (O'Doherty et al., 2004; Tanaka et al., 2004; Tobler et al., 2007; Rangel et al., 2008; Murayama et al., 2010). Finally, we attempted to look at activity in the habenula because it is thought to inhibit the operation of the dopaminergic midbrain (Matsumoto and Hikosaka, 2007; Bromberg-Martin et al., 2010), and it has been suggested that a goal-based reward-learning mechanism (which might depend on OFC) might work in competition with a reward-based learning mechanism that used cached reward values (which might depend on the dopaminergic midbrain) (Daw et al., 2005). Moreover, in the absence of lOFC, the effect of each reward appears to spread forward in time so as to incorrectly reinforce the choices that are made on subsequent trials (Walton et al., 2010). If this spread of reward effect is mediated by dopamine, then it may be necessary for lOFC to curtail dopaminergic midbrain activity, and one way of doing this might be via the habenula. A BOLD response that may reflect habenula activity has been reported for error feedback (Ullsperger and von Cramon, 2003). There are direct connections between the amygdala and lOFC (Carmichael and Price, 1995). There are no direct connections between the lOFC and the most ventromedial part of the striatum, but there are connections between lOFC and adjacent striatal regions (Ferry et al., 2000). No direct connections between lOFC and habenula have been reported, but it is conceivable that OFC might influence the habenula via a number of routes, including the pallidum (Hikosaka et al., 2008).

The subcortical reward-guided learning analysis focused on (1) two 6 mm^3 cube bilateral ROIs in the amygdala (27, -6, -21 and -27, -6, -21) from Hampton et al. (2007), (2) two 6 mm^3 cube bilateral ROIs in the ventral striatum. Because there is some variability in the reported location of reward-related activations in the ventral striatum, we placed the ROI at the peak of the first correct–error contrast from our own experiment (12, 8, -14 and -10, 10, -14) and (3) two 6 mm^3 cube ROIs at the position for the habenula (Tal2MNI.m coordinates 6, -26, 7 and -6, -26, 7) reported by Ullsperger and von Cramon (2003).

At the single-subject level, preprocessing involved the same stages as those described above. The lOFC was chosen as a seed region for this analysis. In standard space, a 6 mm^3 cube ROI was drawn around the peak group difference in activation for the errors contrast in the lOFC (32, 42, -16). FLIRT was used to register this mask into each individual subject space (Jenkinson and Smith, 2001). For each subject, the error EV, the first correct EV, and the mean BOLD time series of the lOFC region, and the two key interactions (first correct EV by lOFC BOLD time series EV and error EV by lOFC BOLD time series EV), were entered independently into a first level GLM analysis. The EV Probability of Correct Response selection (together with its temporal derivatives) and the six motion correction EVs were also included in this analysis. FEAT was used to fit these models to the data to generate PEs for each of the EVs and the PPI. The contrast of interest was the PPI, and this was entered into a mixed effects analysis (FLAME 1 and 2), which was applied to the whole-brain group data. This analysis generated statistical activation

maps for each of the EVs and tested for an effect of experimental group. Group Z (Gaussianized t) statistic images were thresholded using clusters determined by $Z > 2.3$ and a corrected cluster extent significance threshold of $p = 0.05$. As before, there were four contrasts of interest at the second level: Consistent, Inconsistent, Consistent minus Inconsistent, and Inconsistent minus Consistent.

Correlations between ACC BOLD parameter effects and task performance
BOLD signal changes in ACC were better correlated with a greater Probability of Correct Response selection in the Consistent group than in the Inconsistent group. The extent to which each ACC BOLD PE explained between-subject differences in task performance was investigated. At the single-subject level, preprocessing involved the same stages as those described above. The aim was to identify the ACC involvement at various stages of learning of the New task and recall of the Old task. Therefore, in this analysis, there were 20 EVs at the first level: six corresponded to the blocks (36 trials each) of the New learning task, six to the Old recall task blocks, and six to the rest blocks. Each block was modeled from the onset of each experimental nine-trial miniblock or the onset of the rest periods. Each block was composed of four miniblocks of either new or old stimuli or rest. Cues and errors were also included in the model, as were six motion parameters. FEAT was used to fit these models to the data and to generate parameter estimates for each of the EVs. For the purpose of this correlational analysis, the PEs for the 12 experimental regressors (New and Old tasks) for each subject were extracted from a 6 mm^3 cube ROI based on coordinates from the ACC activation identified by the Consistent $-$ Inconsistent group difference effect of Probability of Correct Response selection (-8, 20, 32). The correlation between the extracted PEs for a given block and accuracy of task performance for the same block was then examined. The two groups of subjects (Consistent and Inconsistent) were compared in an analysis of covariance with a time-varying covariate. This analysis allowed testing of prediction of subjects' performances from their BOLD activity, while considering other nonconstant factors such as time (six blocks) and task (Old and New tasks). Such an analysis of covariance allows determination of the impact of several covariates (in this case, the ACC BOLD PEs for each block) on several variables (in this case, the behavioral performance in each block). We extracted measures of correlation for group, task, and time and for the interactions of the factors group by BOLD and group by task by time by BOLD.

Results
Behavioral results
Subjects learned two stimulus response association tasks on days 1 (Stage 3) and 2 (Stage 5) of testing, and in each case they had first performed simple, initial stimulus-outcome learning tasks and response-outcome learning tasks (Stages 1, 2, and 4). In summary, the presence of consistent reward outcomes had no impact on how subjects initially learned about stimuli or on how they explored possible responses (Stages 1, 2, and 4), but it did affect subsequent S–R association learning and recall (Stages 3 and 5). The impact was particularly strong for the second S–R learning task that was learned while subjects were in the MRI scanner (Stage 5). The results are reviewed in detail below.

Stages 1 and 2
Subjects in the Consistent and Inconsistent groups were equally accurate in selecting the rewarding stimulus of a pair in the initial S–O learning task phase ($t_{(34)} = 0.71$; $p = 0.482$) and the recall test ($t_{(34)} = 0.25$; $p = 0.640$). Investigation of the distribution of response choices in the response-outcome learning phase also did not suggest a group difference ($F_{(1,34)} = 0.12$; $p = 0.746$) or an interaction with a particular choice of response ($F_{(1,34)} = 102 = 0.661$; $p = 0.422$). However, there was a significant effect of response ($F_{(3,102)} = 6.89$; $p < 0.001$), which reflected members of both groups having a linear preference for the response options from R1 to R4.
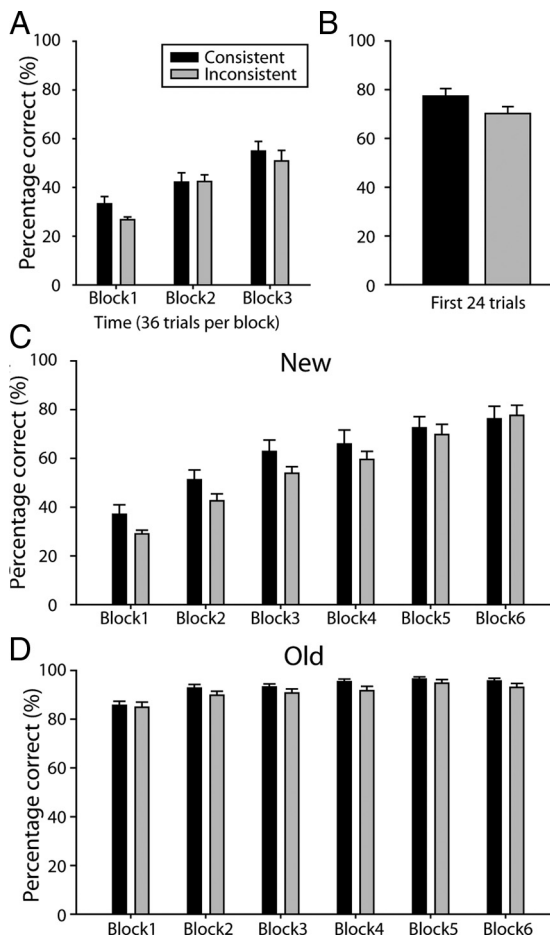
Figure 3. *A*, Mean percentage correct performance over the first three blocks (36 trials per block) of S–R pair learning on day 1. *B*, Mean percentage correct performance for the first 24 trials of S–R pair recall on Day 2. *C*, *D*, Mean percentage correct performance on day 2 for the six blocks (36 trials per block) of S–R pair learning for the New (*C*) and Old (*D*) tasks. Consistent and Inconsistent group subjects are indicated by black and gray bars, respectively, throughout (N = 36). Error bars indicate SEM.

Stage 3

The Consistent and Inconsistent groups did, however, differ in their learning of the S–R task (Fig. 3). There was evidence of DOE; subjects in the Consistent condition were more accurate than subjects in the Inconsistent group in the first 36 trial block ($t_{(34)} = 2.07$, $p = 0.046$; Fig. 3A). By the second and third blocks, the groups were equally accurate (Block 2, $t_{(34)} = -0.033$, $p = 0.974$; Block 3, $t_{(34)} = 0.681$, $p = 0.501$).

If learning of S–R associations was initially facilitated by knowledge of specific reward expectations, it might be hypothesized that memory recall would also be improved (Fig. 3B). Accuracy on the Stage 3 recall test block on day 2 (the day after initial learning) was significantly higher for subjects in the Consistent group compared to the Inconsistent group (one-tailed $t_{(34)} = 1.74$; $p = 0.046$).

Stage 4

On the second day of testing, before entry into the scanner, subjects learned a new set of stimuli and encountered the same response options as the previous day. But, as was the case in Stages 1 and 2, the consistency of the reward mapping had no effect. There were also no differences between Consistent and Inconsistent groups when they learned the new stimulus sets on day 2

($t_{(34)} = 0.25$; $p = 0.640$) or the subsequent recall test ($t_{(34)} = 0.30$; $p = 0.898$). The distribution of response choices in the response-outcome learning phase, again, did not suggest a group difference ($F_{(3,102)} = 0.20$; $p = 0.656$) or an interaction with a particular choice of response ($F_{(1,34)} = 0.69$; $p = 0.539$). There was once again a significant effect of response ($F_{(3,102)} = 7.05$; $p = 0.001$), reflecting a linear preference for the response options, from R1 to R4, for both groups.

Stage 5

As with the first S–R pairing task (Stage 3) performance accuracy in the second, new S–R pair learning task was better when subjects had been trained with Consistent as opposed to Inconsistent reward outcome mappings. The Consistent − Inconsistent effects were particularly clear during New task learning during MRI scanning in stage (Fig. 3C). The data for the learning of the S–R association task were analyzed together with the performance data from the previously learned S–R association (Old recall task) because they were performed together in the MRI scanner, and so the two data sets pertain to the fMRI data that are shown below. An ANOVA of task (Old, New) by time (six blocks of 36 trials) by group (Consistent, Inconsistent) revealed a three-way interaction ($F_{(5,34)} = 2.67$; $p = 0.043$). Subsequent two-factor ANOVAs revealed that there was a significant time by group interaction ($F_{(5,34)} = 2.67$; $p = 0.039$) during the learning of the New task. There was, however, no such interaction in the Old task data set, recorded while subjects were in the MRI scanner (time by group, $F_{(5,34)} = 0.41$, $p = 0.803$). The significant time by group interaction in the New task was due to better performance of the Consistent group subjects, particularly in the first half of the task; there was a significant interaction between the effect of group and the linear effect of time ($F_{(1,34)} = 6.00$; $p = 0.020$), and when the task was divided into two sets of trials, the first half of trials and the second half of trials, the Consistent group performed better than the Inconsistent group in the first half of the task ($t_{(34)} = 2.04$; $p = 0.049$) but not the second half of the task ($t_{(34)} = 0.240$; $p = 0.812$).

Neuroimaging results

Updating reward-related associations

To identify the neural mechanisms involved in updating reward-related associations in the context of consistent reward outcome mappings, we first looked for BOLD activity that was modulated by error feedback in the two groups of subjects.

Activity related to error feedback was significantly increased in the lOFCs of Consistent group subjects compared with those of Inconsistent group subjects (Fig. 4A). Activation in the lOFC (peak z = 3.66; MNI coordinates −32, 48, −22) survived thresholding at $p = 0.001$ (uncorrected). Activity associated with the same contrast also survived correction for small volume at $p = 0.05$, with an 8-mm³-radius sphere centered on coordinates derived from Kringelbach and Rolls (2004) meta-analysis of fMRI activations classified as "punishers leading to a behavioral change" (−32, 42, −18).

To test whether the lOFC is only concerned with error feedback or whether it has a more general role in updating reward-related associations, PEs for the first correct regressor were extracted for the same ROI. As explained in Materials and Methods, the first correct regressor identified outcomes that were positive on stimulus–response trials that were performed correctly for the first time or that were performed correctly after the occurrence of an error. Like error outcomes, these first correct outcomes could be called "informative" outcomes because they
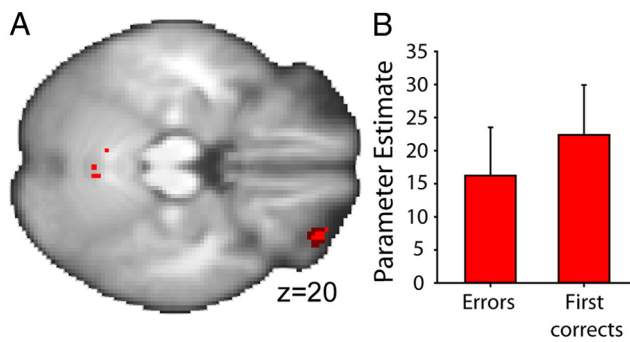
A

z=20

B

Figure 4. *A*, Voxels in lOFC ($z = 4.15$; coordinates $-32, 48, -22$) survived (1) thresholding at $p = 0.001$ (uncorrected, light red) and (2) correction for small volume at $p = 0.05$ with an 8 mm^3 sphere centered on lOFC coordinates [dark red, derived from Kringelbach and Rolls (2004)] for error feedback-related activity that differed between Consistent and Inconsistent groups. Images presented on an averaged brain. While a number of effects were revealed in this manner at the whole-brain level, only uncorrected images are presented for illustrative purposes ($N = 36$). *B*, PEs for informative feedback. First correct feedback and error feedback, from an lOFC ROI, centered over the error-related activation peak for Consistent group subjects are shown (red; $N = 18$). Error bars indicate SEM.

inform subjects about the associations in operation in the task. Both first correct and error outcomes are associated with activity in some frontal cortical areas (Walton et al., 2004; Quilodran et al., 2008), although the areas' role in representing expectations about types of reward has not been investigated. There was significant activation associated with the first correct regressor in lOFC region in the Consistent group subjects (one-sample t test; $t_{(17)} = 2.98$; $p = 0.008$). There was even a tendency for the PE for the first correct regressor to be larger than those for the error regressor.

The lOFC region that we have discussed so far is the one that was identified by the comparison of the error effects in the Consistent and Inconsistent group subjects, and it might be thought that its similar responsiveness to both errors and first correct outcomes is an unusual feature of this particular fraction of the lOFC. We therefore also examined the PEs for the first correct and error regressors in the region where the error effect was maximal when the Consistent group subjects were compared against baseline rather than against the Inconsistent group (MNI coordinates, $-26, 50, -20$) (Fig. 4*B*). This seemed a very direct test of whether the first correct and error outcomes were associated with similar responsiveness in lOFC. Both error and first correct feedback PEs were significantly greater than zero ($t_{(17)} = 2.22$, $p = 0.004$ and $t_{(17)} = 2.96$, $p = 0.009$, respectively), and there was no significant difference in their size (paired sample t test; $t_{(17)} = -0.88$; $p = 0.393$).

In summary, lOFC activity occurs to any informative feedback, whether a first correct outcome or an error outcome that allows updating of reward-related associations. It is most prominent when subjects learn associations with specific reward outcomes in the Consistent group.

To test how specific such a pattern of activation is to lOFC as opposed to mOFC/vmPFC and ACC, the PEs for both informative outcomes, first correct and errors, were compared across ROIs placed in each region (see Materials and Methods, ROI-based signal extraction). There were significant interactions between brain region (lOFC, mOFC/vmPFC, and ACC) and outcome type (first correct, error) that suggested that the areas' responses to feedback did indeed differ ($F_{(2,68)} = 28.79$; $p < 0.001$). The ACC responded to both types of informative outcomes (First Corrects and Errors). Unlike lOFC, however, ACC responses did

not depend on whether or not it was possible for subjects to learn associations between specific stimuli and specific outcomes (interaction of brain region and group; $F_{(1,34)} = 7.91$; $p = 0.008$). Unlike lOFC, the mOFC/vmPFC only responded to positive outcome events (interaction of brain region and outcome type region; $F_{(1,34)} = 22.81$; $p < 0.001$). A more detailed comparison between mOFC/vmPFC and lOFC activity patterns is presented below in the section Contrasting roles of mOFC/vmPFC and lOFC, and the ACC is discussed in detail below in the sections Selecting the correct response and Correlations of ACC BOLD activity and learning performance.

Outside of the three frontal areas of interest, the contrast also revealed activation that survived thresholding at $p = 0.05$ (corrected) in the left insula ($z = 4.57$; MNI coordinates, $-38, 0, 16$) close to the error-responsive region described by Klein et al. (2007), globus pallidus ($z = 3.00$; MNI coordinates, $-32, -10, 8$), premotor cortex ($z = 3.05$; MNI coordinates, $-62, -10, 32$) superior temporal gyrus ($z = 3.23$; MNI coordinates, $-68, -12, 10$), and cerebellum ($z = 3.96$; MNI coordinates, $8, -54, -14$). These regions were found to be more active in Consistent subjects and are therefore potentially contributing to a heightened error-related signal that facilitates overall learning and performance.

Functional connectivity of lOFC during updating of reward-related associations
A PPI was used to assess whether lOFC's role in the updating of reward-related associations reflected interactions with three temporal lobe areas (LO, mFA, and perirhinal cortex) involved in visual form and object representation and with three subcortical areas (amygdala, ventral striatum and habenula) involved in reward-guided learning. The PPI examined whether activity in these areas was modulated as a function of lOFC activity and receipt of first correct or error feedback (for further details, see Materials and Methods, Psychophysiological interaction analysis of lOFC-related activity).

An ANOVA was used to compare the PEs condition-dependent functional coupling between lOFC and the three visual temporal cortex areas (perirhinal cortex, mFA, and LO) that were identified in the PPI analysis. Functional coupling was examined in the context of error feedback and first correct feedback, and so the ANOVA used a two-level factor of feedback. Because coupling with the three visual temporal cortex areas was examined in two hemispheres, the ANOVA also used a three-level factor of area and a two-level factor of hemisphere. Last, because coupling was to be examined as a function of membership of either the Consistent or Inconsistent groups, the ANOVA also used a two-level factor group. There was a significant interaction of region and group ($F_{(2,68)} = 4.14$; $p = 0.029$).

From Figure 5 it is clear that the interaction effects emerged because lOFC activity in the context of receipt of first correct feedback was associated with group differences in the modulation of activity in the perirhinal cortex, an area directly connected with lOFC in the monkey and known to be important for recognizing not just visual features but whole complex objects (Kondo et al., 2005; Buckley and Gaffan, 2006; Murray et al., 2007; Saksida et al., 2007; Saleem et al., 2008). Although lOFC activity during receipt of informative feedback (first correct and error feedback) was also associated with modulation of activity in more posterior LO and mFA temporal lobe regions, it was not modulated in a different way in the Consistent and Inconsistent groups. This impression was confirmed by analyzing data from only the perirhinal cortex. In the perirhinal cortex, there was a significant two-way interaction of feedback and group ($F_{(1,34)} = 9.65$; $p =$
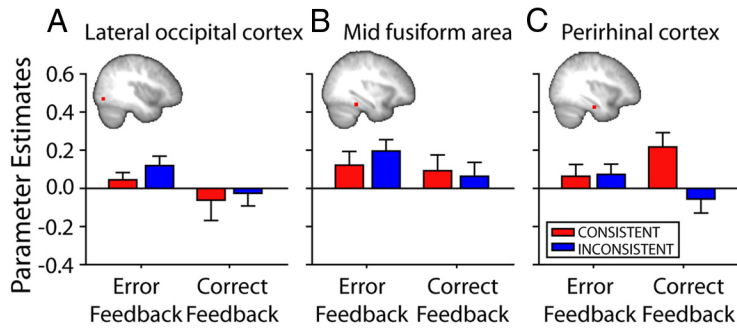
Figure 5. *A–C*, Changes in functional connectivity between left lOFC and LO (*A*), mFA (*B*), and perirhinal cortex (*C*) for Consistent (red) and Inconsistent (blue) subjects at the time of informative feedback (left, first correct feedback; right, error feedback). PEs were extracted for bilateral temporal lobe ROIs, although only the right hemisphere is illustrated. Although lOFC activity during receipt of informative feedback was associated with modulation of activity in more posterior temporal lobe areas, LO and mFA, it was not modulated in a different way in the two groups. In contrast, functional connectivity within perirhinal cortex, which has strong connections to lOFC and is important for representing complex visual stimuli like the ones used in the experiment, differed between Consistent and Inconsistent groups (N = 36). Error bars indicate SEM.
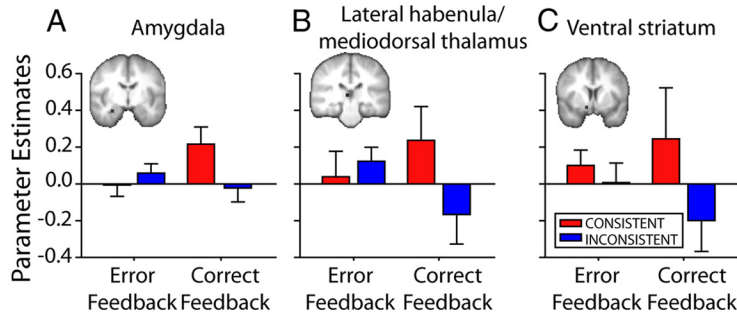

Figure 6. *A–C*, Changes in functional connectivity between left lOFC and amygdala (*A*), lateral habenula/mediodorsal thalamus (*B*), and ventral striatum (*C*) for Consistent (red) and Inconsistent (blue) subjects at the time of informative feedback (left, first correct feedback; right, error feedback). PEs were extracted for bilateral subcortical ROIs, although only right hemisphere data are illustrated. Functional connectivity with each subcortical area differed between Consistent and Inconsistent groups (N = 36). Error bars indicate SEM.

0.004) and a main effect of hemisphere ($F_{(1,34)} = 5.83$; $p = 0.021$). Independent sample t tests identified significant group differences in first correct feedback effects in the right hemisphere and a trend to significance in the left hemisphere ($t_{(34)} = 2.61$, $p = 0.013$ and $t_{(34)} = 1.91$, $p = 0.065$ respectively).

An ANOVA was used to compare the PEs indexing condition dependent functional coupling between lOFC and the three subcortical brain areas (amygdala, ventral striatum, and habenula) that were identified in the PPI analysis. Functional coupling was examined in the context of error feedback and first correct feedback, and so the ANOVA used a two-level factor of feedback. Because coupling with the three subcortical brain areas was examined in two hemispheres, the ANOVA also used a three-level factor of area and a two-level factor of hemisphere. Last, because coupling was to be examined as a function of membership of either the Consistent or Inconsistent groups, the ANOVA also used a two-level factor group. There was a significant interaction of feedback and group ($F_{(1,34)} = 5.20$, $p = 0.029$; Fig. 6). From Figure 6 it can be seen that coupling between lOFC and each of the three subcortical areas was prominent when first correct feedback was received, and whereas the coupling tended to be positive in the Consistent group, it tended to be negative in the Inconsistent group.

An important note of caution must, however, be mentioned. The habenula is a very small brain area adjacent to the mediodorsal nucleus of the thalamus, and on *post hoc* observation it was clear that the activation pattern that was present in the mediodorsal thalamus was similar to the one shown in Figure 6. It is therefore a possibility that the results in Figure 6 actually reflect mediodorsal thalamic effects instead of or in addition to habenula effects. The mediodorsal thalamus is connected to the lOFC in monkeys (Ray and Price, 1993) and probably in humans (Klein et al., 2010), and in the monkey it interacts with the OFC and amygdala during S–O learning (Gaffan and Murray, 1990; Izquierdo and Murray, 2010).

In summary, when consistent S–O mappings are present, the lOFC enters into a distinctive pattern of positive coupling with the perirhinal cortex, a brain region critical for representing complex visual objects, and subcortical areas implicated in reward-guided learning each time correct feedback for a particular S–R pair choice is received for the first time or after previous errors have been made.

Contrasting roles of mOFC/vmPFC and lOFC

We also tested whether mOFC/vmPFC was like lOFC and equally responsive to either informative feedback that allowed updating of reward-related associations or whether it was more responsive to feedback that entailed reward. The latter hypothesis is consistent with the claim that mOFC/vmPFC encodes the value of reward outcomes as well as expectations (Sescousse et al., 2010; Smith et al., 2010). The comparison of first correct minus error feedback (informative outcomes) across the whole brain revealed an extended reward-related network of regions in both groups of subjects, including mOFC/vmPFC (Fig. 7*Ai*), similar to that reported previously when positive and negative outcomes were compared, regardless of how informative they were (Murayama et al., 2010; Sescousse et al., 2010). At the whole-brain level, there were no significant differences between the two groups (Table 1 reports those for Consistent group subjects only). Errors and first correct PEs were extracted from an ROI centered on the peak vmPFC/mOFC activation coordinates ($z = 8.22$; MNI coordinates, 0, 48, −10). As might be expected, the vmPFC/mOFC shows a striking relative activation in response to first correct feedback compared to errors (Fig. 7*Aii*). Comparison of the first correct and error PEs in the mOFC/vmPFC and lOFC area discussed above revealed a main effect of brain region ($F_{(1,17)} = 19.17$; $p < 0.001$), feedback type ($F_{(1,17)} = 20.54$; $p < 0.001$), and an interaction between the two factors ($F_{(1,17)} = 12.31$; $p < 0.05$).

An additional analysis (see Materials and Methods, Region of interest-based signal extraction) used a regressor designed to capture the expected reward value for each outcome on every trial as a function of the subjective value to each subject of the reward type expected and the probability that it would be obtained by making a correct response. The analysis focused on the mOFC/vmPFC region identified by Smith et al. (2010) as most specifi-

cally encoding value expectation (12, 58, 2). A direct comparison (Fig. 7B) between the value PEs in the mOFC/vmPFC region and the lOFC region identified earlier demonstrated they were significantly different in the two areas ($t_{(14)} = -3.41$; $p = 0.004$), with the mOFC/vmPFC being significantly different from zero ($t_{(14)} = 2.22$; $p = 0.044$) while the lOFC region is not ($t_{(14)} = -0.63$; $p = 0.541$).

Selecting the correct response
The final set of analyses looked at activity related to selection of responses when selection was mediated by S–O and O–R associations. Activity related to the Probability of Correct Response selection, a regressor that tracks each subject's knowledge of how best to perform the task, and differing between the two groups (greater in the Consistent than Inconsistent group) was found in the ACC ($z = 3.26$; MNI coordinates, −8, 20, 30; Fig. 8). Coordinates from the peak group difference extracted from within our a priori region of interest (8-mm^3-radius sphere centered on the average coordinates from reward-related activations in this laboratory; MNI coordinates, −9, 21, 37) were used to extract PEs for Probability of Correct Response selection-related activity, and it can be seen from Figure 8 that the effect was positive in the Consistent group subjects and negative in the Inconsistent group subjects.

To test the specificity of the ACC representation for guiding response selection in the Consistent condition, we compared the PEs associated with the Probability of Correct Response selection regressor in the ACC with the PEs associated with the same regressor in the lOFC and mOFC/vmPFC ROIs (see Materials and Methods, ROI-based signal extraction). The PE sizes varied significantly across region (linear main effect; $F_{(1,17)} = 5.48$; $p = 0.032$). While PEs were significantly greater than zero in ACC ($t_{(17)} = 3.53$; $p = 0.002$), that was not the case in either lOFC ($t_{(17)} = -0.48$; $p = 0.635$) or mOFC/vmPFC ($t_{(17)} = -0.98$; $p = 0.341$).

It should be noted that the ACC activation extended some distance anterior to the ACC peak activation coordinate and the a priori region of interest in the ACC; an additional subpeak was found rostral to the rostral cingulate motor zone in a region that probably corresponds to the area variously referred to as 32ac (Ongur et al., 2003), 32′ (Vogt, 2008), and cluster 4 (Beckmann et al., 2009) (MNI coordinates, 0, 40, 26; Fig. 8, axial view). In addition, activity was also identified in the adjacent presupplementary motor area ($z = 3.27$; MNI coordinates, 2, 20, 54), a posterior cingulate region ($z = 3.68$; MNI coordinates, 8, −48, 32), the bilateral temporoparietal junction/angular gyrus ($z = 3.91$, MNI coordinates, 54, −48, 14 and $z = 4.10$, MNI coordinates, −50, −66, 22, respectively), and cerebellum ($z = 3.77$, MNI coordinates, −6, −50, −16) corrected at cluster 2.3 and $p = 0.05$.

Correlations of ACC BOLD activity and learning performance
If ACC activity is important for selecting the correct response when selection is mediated via O–R associations, then individual differences in ACC activity may underlie individual differences in how successfully subjects, especially subjects in the Consistent group, select the correct response. We investigated whether individual differences in the size of ACC PEs related to the Probability
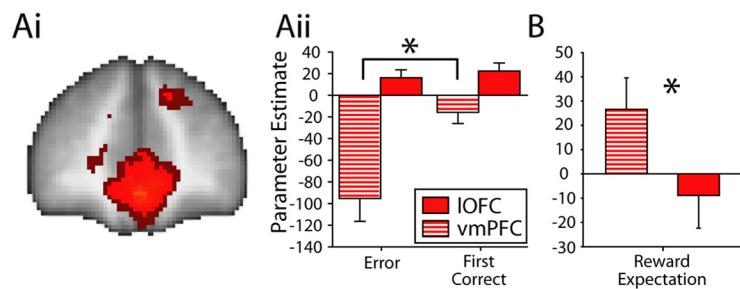

Figure 7. **Ai**, MOFC/vmPFC voxels surviving (1) thresholding at $p = 0.001$ (uncorrected; light red) and (2) thresholding with whole-brain comparison correction ($z < 2.3$; $p < 0.05$; dark red) for the contrast of first correct feedback minus error feedback in the Consistent group. **Aii**, PE sizes for Errors and First Corrects from the ROI at peak mOFC/vmPFC coordinates ($z = 7.52$; 0, 48, −10) compared with PEs from the lOFC ROI (Fig. 3). Comparison between the two regions indicates that mOFC/vmPFC reflects the occurrence of a reward outcome, whereas lOFC activity reflects the occurrence of any outcome, whether positive or negative, that enables subjects to update reward-related associations **B**, PE size for reward expectation from a bilateral ROI from Smith et al. (2010) (MNI coordinates, left: −6, 26, −14; right, 6, 26, −14) compared with the PE from the lOFC ROI (Fig. 3). Comparison between the mOFC/vmPFC and lOFC suggests that only the former region represents expectations of reward value. All images are presented on an averaged brain image ($N = 18$). Error bars indicate SEM. *$p < 0.05$.

Table 1. Activations of First Correct minus Errors

Region	z statistic	x	y	z
Ventromedial prefrontal cortex	6.36	−2	54	−6
Ventromedial prefrontal cortex	7.52	0	48	−10
Ventromedial prefrontal cortex	7.43	2	46	−10
Anterior cingulate cortex	5.82	−2	38	0
Caudate head	5.43	−10	10	−14
Putamen	5.76	6	10	−10
Caudate head	6.13	12	8	−14
Calcarine fissure	5.11	10	−65	8

Brain areas in which there was a significant effect of the contrast of first correct minus error outcome trials in the Consistent group subjects only and clusters of BOLD activation exceeding $z = 5.0$ ($N = 18$) are shown.
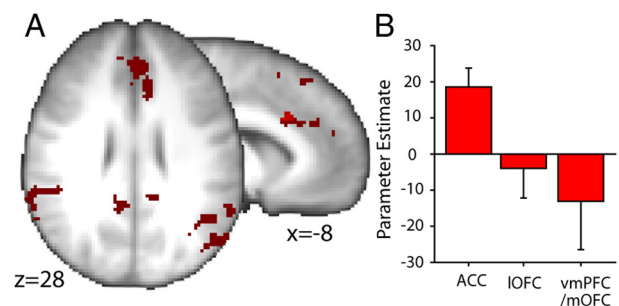

Figure 8. **A**, ACC ($z = 4.34$; coordinates, −8, 20, 30) voxels surviving (1) thresholding at $p = 0.001$ (whole-brain comparison uncorrected; light red), (2) correction for small volume at $p = 0.05$ with an 8 mm^3 sphere centered on averaged coordinates from reward-related activations in this laboratory (coordinates, −9, 21, 37), and (3) thresholding with whole-brain comparison correction ($z < 2.3$; $p < 0.05$; dark red). BOLD activity in these voxels varied with the Probability of Correct Response selection over each nine-trial miniblock and to differing degrees in the Consistent and Inconsistent groups. Images are presented on a averaged brain ($N = 36$). **B**, The PE size from an ACC ROI is compared with the PE sizes for the same contrast from ROIs in lOFC and mOFC/vmPFC in Consistent group subjects ($N = 18$). Error bars indicate SEM.

of Correct Response selection were related to individual differences in performance accuracy. A supplementary GLM (see Materials and Methods, Correlations between ACC BOLD parameter effects and task performance) was used and included regressors related to Probability of Correct Response selection for each of the six blocks (of 36 trials) of the New learning task and the six blocks of the Old recall task, each divided into six blocks of 36 trials across the scanning session of interest. In this way, it was possible to extract a PE for Probability of Correct Response
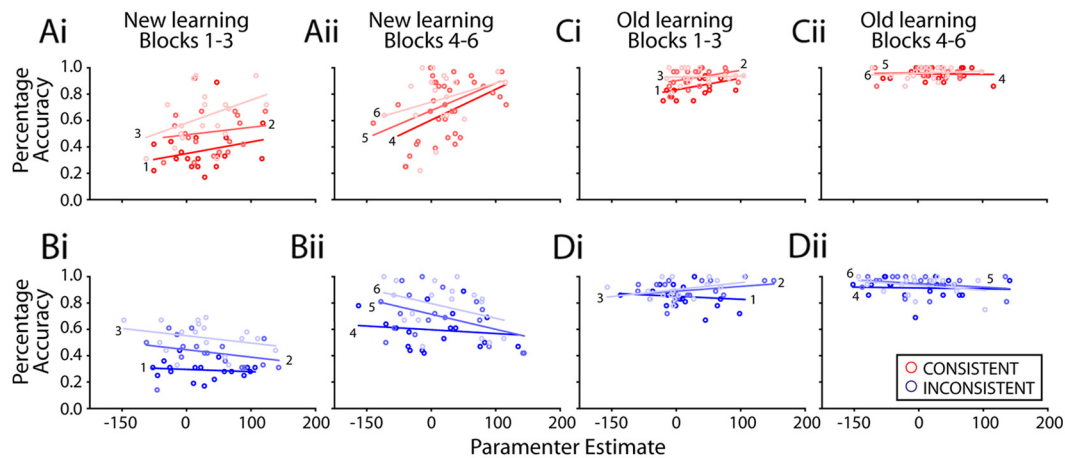
Figure 9. Correlations of BOLD PEs (extracted from a 6 mm^3 cube ACC ROI; Fig. 7) for each 36 trial block of S–R pair learning against the same subjects' average percentage accuracy scores for the same blocks. *Ai–Dii*, New (*Ai–Bii*) and Old (*Ci–Dii*) task blocks are shown moving from left to right (each task was composed of 6 blocks, and for each task, increasing lightness of shading indicates block positions 1–3 and 4 – 6). In other words, the earliest New learning blocks (Blocks 1–3) are shown on the far left of the figure, whereas the final Old task recall blocks are shown on the right of the figure, and data from three blocks are shown in each panel. Consistent (red) and Inconsistent (blue) subjects are shown at the top and bottom, respectively. There is a positive correlation between ACC PEs and behavioral accuracy scores in Consistent group subjects, especially during New task learning, suggesting that a larger ACC signal was associated with better performance. In contrast, there was a negative correlation between these two factors at similar points during New task learning in Inconsistent group subjects (N = 36).

selection for all 12 blocks of the task and to examine the correlation between the subjects' PEs and the subjects' task performance accuracies in each of the 12 blocks of Old and New task performance.

As can be seen from Figure 9, performance accuracy in the Consistent condition is positively related to ACC PEs, whereas performance accuracy in the Inconsistent condition is negatively related to ACC PEs. These relationships are more evident in the New learning blocks (Fig. 9*Ai–Bii*) than the Old recall blocks (*Ci–Dii*). In other words, higher ACC PEs are predictive of more accurate response selection, but only when response selection is mediated by O–R associations in the Consistent group. This picture was confirmed when performance accuracies in each block in the two groups of subjects (Consistent and Inconsistent) were compared in an ANCOVA with a time-varying covariate of the ACC PE in each block and factors of group (Consistent and Inconsistent) and time on task (12 levels corresponding to the six New learning blocks and the six Old recall blocks). The analysis suggested that subjects' performances were predicted by the interaction between group membership and ACC PE ($F_{(2,317)} = 3.285$; $p = 0.039$), as well as by the interaction of group, time, and BOLD ($F_{(22,370)} = 2.707$; $p < 0.001$). In this analysis, performance was not predictable by group assignment alone ($F_{(1,37)} = 0.349$; $p = 0.558$).

Discussion

The faster learning and better recall of S–R pairing in the presence of consistent outcome mappings constitutes a demonstration of the DOE and confirms that S–O and O–R associations mediate response selection in humans. While the reported DOE effects are usually close to the $p = 0.05$ significance threshold, they are reproducible in a number of different measures. The DOE models situations that may frequently occur during goal-directed behavior (Savage, 2001); responses are not automatically made just because a stimulus is present, but instead are induced by the prospect of specific types of rewards that environmental stimuli indicate are available.

According to one influential proposal, lOFC and mOFC/vmPFC differ in responsiveness to negative and positive feedback, respectively (Kringelbach and Rolls, 2004). An alternative

proposal, based on recent investigation of macaque OFC (Noonan et al., 2010; Walton et al., 2010), maintains that lOFC is concerned with the updating of precise associations between stimuli and outcomes during learning, regardless of whether outcomes are positive or negative. In contrast, it is argued mOFC/vmPFC maintains value representations of attended stimuli used to guide choice. Only this latter view predicted lOFC would be more active when Consistent group subjects were updating associations between particular stimuli and particular reward outcomes regardless of whether feedback was positive or negative (Fig. 4). Previous studies that have not recorded lOFC activation to positive outcomes may not have done so because they did not focus on positive outcomes, such as the first positive outcome associated with a particular S–R pair, that inform the subject what reward-related associations are in operation and because informative positive outcomes are often less frequent than negative outcomes in most task designs.

An ROI-based analysis also identified similar ACC activity for first correct and error feedback, as observed previously (Walton et al., 2004; Mars et al., 2005; Matsumoto et al., 2007; Sallet et al., 2007; Quilodran et al., 2008). Unlike in lOFC, however, in ACC there was no significant difference between activity in Consistent and Inconsistent groups. The pattern of results favor the suggestion (Rushworth et al., 2007; Rushworth and Behrens, 2008) that while ACC carries reward prediction and reward prediction error-like signals, it may only be in lOFC that such signals incorporate information about type of reward. In contrast ACC, but not OFC, value signals reflect the anticipated effort entailed by the response in addition to the anticipated reward (Croxson et al., 2009; Prevost et al., 2010). Because it has been argued that some frontal areas are responsive to informative feedback but do not possess quantitative reward prediction errors (Sallet et al., 2007; Quilodran et al., 2008), we eschewed formal modeling of prediction errors.

Noonan et al. (2010) speculated that lOFC's role in learning associations between specific stimuli and outcomes might reflect connections with temporal cortical areas important for form and object representation (Kondo et al., 2005; Saleem et al., 2008).

PPI analyses in the current study examined BOLD signal changes as a function of lOFC activity and feedback and found evidence of increased coupling between lOFC and perirhinal cortex particularly when subjects received correct feedback for a particular S–R pairing for the first time. While occipitotemporal areas such as the LO and mFA are responsive to visual forms and features, the more anterior perirhinal cortex is essential for recognizing complex configurations of component features during visual object recognition in monkeys (Buckley and Gaffan, 1997, 1998; Murray and Richmond, 2001; Saksida et al., 2007) and people (Lee et al., 2006). Such coupling might grant lOFC access to stimulus representations at the time informative reward feedback is received (cf. Tsujimoto et al., 2009). Interactions with perirhinal cortex are consistent with human lOFC being more concerned with S–O, as opposed to R–O, associations, as suggested by rat and monkey lesion experiments (Ostlund and Balleine, 2007; Rudebeck et al., 2008). lOFC neurons are more active than mOFC/vmPFC neurons when visual stimuli cue reward-motivated behavior (Bouret and Richmond, 2010).

There was also evidence for similar feedback-related coupling between lOFC and three subcortical areas (amygdala, ventral striatum, and habenula/mediodorsal thalamus) when consistent S–O mappings were present. In monkeys and rats, the amygdala and its interactions with OFC and mediodorsal thalamus are especially important when specific S–O associations are learned (Malkova et al., 1997; Baxter et al., 2000; Schoenbaum and Roesch, 2005; Izquierdo and Murray, 2007, 2010). Although the role of areas such as the ventral striatum in learning are often discussed in the context of the influence of the dopaminergic midbrain, the present results suggest that coupling with lOFC may be an important determinant of its functioning when specific S–O mappings are present. Competitive interactions between frontal goal-based systems and other reward-learning systems are predicted on theoretical grounds (Daw et al., 2005).

MOFC/vmPFC activity differed significantly from lOFC and ACC activity. MOFC/vmPFC activity did not reflect whether outcomes were informative for updating associations, but whether they were positive (Fig. 7Aii). In addition, mOFC/vmPFC activation did not reflect whether subjects were learning consistent S–O mappings in the Consistent as opposed to Inconsistent task. Moreover, while mOFC/vmPFC activity was proportional to the value of outcomes expected after response selection, as reported previously (O'Doherty et al., 2002; Plassmann et al., 2007; Lebreton et al., 2009; Smith et al., 2010), this was not true in lOFC (Fig. 7B).

Activation in mOFC/vmPFC has been emphasized in investigations of goal-based action in which different rewards are associated with different options and which use reward devaluation, probing whether choices reflect changing outcome values (Valentin et al., 2007; de Wit et al., 2009). The present results suggest mOFC/vmPFC activity does not, however, reflect learning and updating of specific S–O associations. Rather, in tandem with other recent studies, they suggest mOFC/vmPFC activation in devaluation studies reflects subjects' expectations about the value of outcomes as a consequence of having learned specific S–O associations and the process of attending to those values and choosing between them; mOFC/vmPFC activity reflects the values of two different rewards that are offered to subjects and the one that they attend to or choose (Boorman et al., 2009; FitzGerald et al., 2009; Wunderlich et al., 2011).

The mOFC/vmPFC may be better suited to represent the value of reward regardless of identity because mOFC/vmPFC, but not lOFC, neurons reflect aspects of the organism's physiological state that determine a reward's value. For example mOFC/vmPFC neuron responses to fluid rewards reflect thirst (Bouret and Richmond, 2010). Such a specialization may reflect the greater bidirectional interconnection of mOFC/vmPFC with autonomic brain regions (Ongur et al., 1998; Rempel-Clower and Barbas, 1998).

It is important to distinguish between two selection processes confounded in many decisions: the selection of a reward outcome goal, which depends directly or indirectly on mOFC/vmPFC, and subsequent selection of a response to obtain the reward goal. The current study suggests that the latter process depends on ACC. ACC activity significantly differed from lOFC and mOFC/vmPFC activity in the degree that it reflected the probability of selecting the correct response. Moreover the ACC signal was more prominent in the Consistent group, suggesting that it reflected knowledge and mediation of response selection via O–R associations available to subjects in that group but unavailable to those in the Inconsistent group. In other words, although ACC activity, unlike lOFC activity, did not reflect updating of specific reward outcome-based associations, it did reflect use of such associations to guide response selection. That ACC PEs were more positively correlated with performance in the Consistent group than the Inconsistent group provided further evidence that this was the case. Activity in a network of regions including vlPFC and dorsal premotor cortex is associated with action selection via arbitrary, learned S–R associations (Passingham et al., 2000; Wise and Murray, 2000; Petrides, 2005; Bunge and Wallis, 2008). The ACC may be important when O–R associations also guide response selection.

Such a role for ACC is consistent with its possession of reward-related activity (Matsumoto et al., 2003; Amiez et al., 2006; Quilodran et al., 2008; Kennerley et al., 2009; Luk and Wallis, 2009) that is combined with information about responses (Kennerley et al., 2009; Hayden and Platt, 2010) and evidence that lesions disrupt reward-guided selection of actions (Kennerley et al., 2006; Rudebeck et al., 2008). The present results add to the emerging picture of ACC function, showing that ACC does not just decide how much or how frequently to make an action given expectations of reward magnitude or probability, but that it may also select a specific action because stimuli in the environment indicate that a particular reward is available.

While acknowledging other organizing principles underlying functional organization in OFC, such as anterior–posterior gradients related to reward complexity (Sescousse et al., 2010) and that frontal areas are sensitive to other aspects of reward (Tobler et al., 2009), the present results highlight basic distinctions between lOFC, mOFC/vmPFC, and ACC that are likely to underlie the areas' specializations in other contexts. A distributed pattern of activity spanning amygdala, perirhinal cortex, lOFC, and ACC reflects S–O and O–R associations that provide an alternative to S–R associations as a route to action selection, whereas mOFC/vmPFC represents the value of the outcomes anticipated.

References

Aggleton JP, Albasser MM, Aggleton DJ, Poirier GL, Pearce JM (2010) Lesions of the rat perirhinal cortex spare the acquisition of a complex configural visual discrimination yet impair object recognition. Behav Neurosci 124:55–68.

Amiez C, Joseph JP, Procyk E (2006) Reward encoding in the monkey anterior cingulate cortex. Cereb Cortex 16:1040–1055.

Baxter MG, Murray EA (2002) The amygdala and reward. Nat Rev Neurosci 3:563–573.

Baxter MG, Parker A, Lindner CC, Izquierdo AD, Murray EA (2000) Con-

trol of response selection by reinforcer value requires interaction of amygdala and orbital prefrontal cortex. J Neurosci 20:4311–4319.

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

Behrens TE, Woolrich MW, Walton ME, Rushworth MF (2007) Learning the value of information in an uncertain world. Nat Neurosci 10:1214–1221.

Behrens TE, Hunt LT, Woolrich MW, Rushworth MF (2008) Associative learning of social value. Nature 456:245–249.

Boorman ED, Behrens TE, Woolrich MW, Rushworth MF (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.

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

Brett M (1999) The MNI brain and the Talairach atlas. Available at http://imaging.mrc-cbu.cam.ac.uk/imaging/MniTalairach.

Bromberg-Martin ES, Matsumoto M, Hong S, Hikosaka O (2010) A pallidus-habenula-dopamine pathway signals inferred stimulus values. J Neurophysiol 104:1068–1076.

Buckley MJ, Gaffan D (1997) Impairment of visual object-discrimination learning after perirhinal cortex ablation. Behav Neurosci 111:467–475.

Buckley MJ, Gaffan D (1998) Perirhinal cortex ablation impairs visual object identification. J Neurosci 18:2268–2275.

Buckley MJ, Gaffan D (2006) Perirhinal cortical contributions to object perception. Trends Cogn Sci 10:100–107.

Bunge S, Wallis JD, eds (2008) Neuroscience of rule-guided behavior. New York: Oxford UP.

Carmichael ST, Price JL (1995) Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. J Comp Neurol 363:615–641.

Croxson PL, Walton ME, O'Reilly JX, Behrens TE, Rushworth MF (2009) Effort-based cost-benefit valuation and the human brain. J Neurosci 29:4531–4541.

Daw ND, Niv Y, Dayan P (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat Neurosci 8:1704–1711.

Deichmann R, Gottfried JA, Hutton C, Turner R (2003) Optimized EPI for fMRI studies of the orbitofrontal cortex. Neuroimage 19:430–441.

De Martino B, Kumaran D, Seymour B, Dolan RJ (2006) Frames, biases, and rational decision-making in the human brain. Science 313:684–687.

de Wit S, Corlett PR, Aitken MR, Dickinson A, Fletcher PC (2009) Differential engagement of the ventromedial prefrontal cortex by goal-directed and habitual behavior toward food pictures in humans. J Neurosci 29:11330–11338.

Easton A (2004) Differential reward outcome learning in adult humans. Behav Brain Res 154:165–169.

Easton A, Gaffan D (2002) Insights into the nature of fronto-temporal interactions from a biconditional discrimination task in the monkey. Behav Brain Res 136:217–226.

Estevez AF, Vivas AB, Alonso D, Mari-Beffa P, Fuentes LJ, Overmier JB (2007) Enhancing challenged students' recognition of mathematical relations through differential outcomes training. Q J Exp Psychol (Colchester) 60:571–580.

Ferry AT, Ongur D, An X, Price JL (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 TH, Seymour B, Dolan RJ (2009) The role of human orbitofrontal cortex in value comparison for incommensurable objects. J Neurosci 29:8388–8395.

Friston KJ, Buechel C, Fink GR, Morris J, Rolls E, Dolan RJ (1997) Psychophysiological and modulatory interactions in neuroimaging. Neuroimage 6:218–229.

Gaffan D, Murray EA (1990) Amygdalar interaction with the mediodorsal nucleus of the thalamus and the ventromedial prefrontal cortex in stimulus-reward associative learning in the monkey. J Neurosci 10:3479–3493.

Gaffan EA, Eacott MJ, Simpson EL (2000) Perirhinal cortex ablation in rats selectively impairs object identification in a simultaneous visual comparison task. Behav Neurosci 114:18–31.

Grill-Spector K (2003) The neural basis of object perception. Curr Opin Neurobiol 13:159–166.

Hampton AN, Adolphs R, Tyszka MJ, O'Doherty JP (2007) Contributions of the amygdala to reward expectancy and choice signals in human prefrontal cortex. Neuron 55:545–555.

Hare TA, O'Doherty J, Camerer CF, Schultz W, Rangel A (2008) Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. J Neurosci 28:5623–5630.

Hayden BY, Platt ML (2010) Neurons in anterior cingulate cortex multiplex information about reward and action. J Neurosci 30:3339–3346.

Hikosaka O, Sesack SR, Lecourtier L, Shepard PD (2008) Habenula: crossroad between the basal ganglia and the limbic system. J Neurosci 28:11825–11829.

Izquierdo A, Murray EA (2007) Selective bilateral amygdala lesions in rhesus monkeys fail to disrupt object reversal learning. J Neurosci 27:1054–1062.

Izquierdo A, Murray EA (2010) Functional interaction of medial mediodorsal thalamic nucleus but not nucleus accumbens with amygdala and orbital prefrontal cortex is essential for adaptive response selection after reinforcer devaluation. J Neurosci 30:661–669.

Jenkinson M, Smith S (2001) A global optimisation method for robust affine registration of brain images. Med Image Anal 5:143–156.

Jones M, White KG (1994) An investigation of the differential-outcomes effect within sessions. J Exp Anal Behav 61:389–406.

Kennerley SW, Walton ME, Behrens TE, Buckley MJ, Rushworth MF (2006) Optimal decision making and the anterior cingulate cortex. Nature Neurosci 9:940–947.

Kennerley SW, Dahmubed AF, Lara AH, Wallis JD (2009) Neurons in the frontal lobe encode the value of multiple decision variables. J Cogn Neurosci 21:1162–1178.

Klein JC, Rushworth MF, Behrens TE, Mackay CE, de Crespigny AJ, D'Arceuil H, Johansen-Berg H (2010) Topography of connections between human prefrontal cortex and mediodorsal thalamus studied with diffusion tractography. Neuroimage 51:555–564.

Klein TA, Neumann J, Reuter M, Hennig J, von Cramon DY, Ullsperger M (2007) Genetically determined differences in learning from errors. Science 318:1642–1645.

Kondo H, Saleem KS, Price JL (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 ML, Rolls ET (2004) The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. Prog Neurobiol 72:341–372.

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

Lee AC, Bandelow S, Schwarzbauer C, Henson RN, Graham KS (2006) Perirhinal cortex activity during visual object discrimination: an event-related fMRI study. Neuroimage 33:362–373.

Luk CH, Wallis JD (2009) Dynamic encoding of responses and outcomes by neurons in medial prefrontal cortex. J Neurosci 29:7526–7539.

Malkova L, Gaffan D, Murray EA (1997) Excitotoxic lesions of the amygdala fail to produce impairment in visual learning for auditory secondary reinforcement but interfere with reinforcer devaluation effects in rhesus monkeys. J Neurosci 17:6011–6020.

Mars RB, Coles MG, Grol MJ, Holroyd CB, Nieuwenhuis S, Hulstijn W, Toni I (2005) Neural dynamics of error processing in medial frontal cortex. Neuroimage 28:1007–1013.

Matsumoto K, Suzuki W, Tanaka K (2003) Neuronal correlates of goal-based motor selection in the prefrontal cortex. Science 301:229–232.

Matsumoto M, Hikosaka O (2007) Lateral habenula as a source of negative reward signals in dopamine neurons. Nature 447:1111–1115.

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

Murayama K, Matsumoto M, Izuma K, Matsumoto K (2010) Neural basis of the undermining effect of monetary reward on intrinsic motivation. Proc Natl Acad Sci U S A 107:20911–20916.

Murray EA, Richmond BJ (2001) Role of perirhinal cortex in object perception, memory, and associations. Curr Opin Neurobiol 11:188–193.

Murray EA, Bussey TJ, Saksida LM (2007) Visual perception and memory: a new view of medial temporal lobe function in primates and rodents. Annu Rev Neurosci 30:99–122.

Noonan MP, Walton ME, Behrens TE, Sallet J, Buckley MJ, Rushworth MF (2010) Separate value comparison and learning mechanisms in macaque

medial and lateral orbitofrontal cortex. Proc Natl Acad Sci U S A 107:20547–20552.

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

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.

Ongur D, An X, Price JL (1998) Orbital and medial prefrontal cortical projections to the hypothalamus in macaque monkeys. J Comp Neurol 401:480–505.

Ongur D, Ferry AT, Price JL (2003) Architectonic subdivision of the human orbital and medial prefrontal cortex. J Comp Neurol 460:425–449.

Ostlund S, Balleine BW (2007) The contribution of orbitofrontal cortex to action selection. Ann N Y Acad Sci 1121:174–192.

Passingham RE, Toni I, Rushworth MFS (2000) Specialisation within the prefrontal cortex: the ventral prefrontal cortex and associative learning. Exp Brain Res 133:103–113.

Petrides M (2005) Lateral prefrontal cortex: architectonic and functional organization. Philos Trans R Soc Lond B Biol Sci 360:781–795.

Picard N, Strick PL (2001) Imaging the premotor areas. Curr Opin Neurobiol 11:663–672.

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

Plaza V, Estévez AF, López-Crespo G, Fuentes LJ (2011) Enhancing recognition memory in adults through differential outcomes. Acta Psychol (Amst) 136:129–136.

Prevost C, Pessiglione M, Metereau E, Clery-Melin ML, Dreher JC (2010) Separate valuation subsystems for delay and effort decision costs. J Neurosci 30:14080–14090.

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

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

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

Ray JP, Price JL (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.

Rempel-Clower NI, Barbas H (1998) Topographic organization of connections between the hypothalamus and prefrontal cortex in the rhesus monkey. J Comp Neurol 398:393–419.

Rudebeck PH, Murray EA (2011) Dissociable effects of subtotal lesions within the macaque orbital prefrontal cortex on reward-guided behavior. J Neurosci 31:10569–10578.

Rudebeck PH, Behrens TE, Kennerley SW, Baxter MG, Buckley MJ, Walton ME, Rushworth MF (2008) Frontal cortex subregions play distinct roles in choices between actions and stimuli. J Neurosci 28:13775–13785.

Rushworth MF, Behrens TE (2008) Choice, uncertainty and value in prefrontal and cingulate cortex. Nat Neurosci 11:389–397.

Rushworth MF, Behrens TE, Rudebeck PH, Walton ME (2007) Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour. Trends Cogn Sci 11:168–176.

Rushworth MF, Noonan MP, Boorman ED, Walton ME, Behrens TE (2011) Frontal cortex and reward-guided learning and decision-making. Neuron 70:1054–1069.

Saksida LM, Bussey TJ, Buckmaster CA, Murray EA (2007) Impairment and facilitation of transverse patterning after lesions of the perirhinal cortex and hippocampus, respectively. Cereb Cortex 17:108–115.

Saleem KS, Kondo H, Price JL (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, Rothe M, Vezoli J, Joseph JP, Procyk E (2007) Expectations, gains, and losses in the anterior cingulate cortex. Cogn Affect Behav Neurosci 7:327–336.

Savage LM (2001) In search of the neurobiological underpinnings of the differential outcomes effect. Integr Physiol Behav Sci 36:182–195.

Schoenbaum G, Roesch M (2005) Orbitofrontal cortex, associative learning, and expectancies. Neuron 47:633–636.

Sescousse G, Redoute J, Dreher JC (2010) The architecture of reward value coding in the human orbitofrontal cortex. J Neurosci 30:13095–13104.

Smith DV, Hayden BY, Truong TK, Song AW, Platt ML, Huettel SA (2010) Distinct value signals in anterior and posterior ventromedial prefrontal cortex. J Neurosci 30:2490–2495.

Smith SM (2002) Fast robust automated brain extraction. Hum Brain Mapp 17:143–155.

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

Tobler PN, Fletcher PC, Bullmore ET, Schultz W (2007) Learning-related human brain activations reflecting individual finances. Neuron 54:167–175.

Tobler PN, Christopoulos GI, O'Doherty JP, Dolan RJ, Schultz W (2009) Risk-dependent reward value signal in human prefrontal cortex. Proc Natl Acad Sci U S A 106:7185–7190.

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

Ullsperger M, von Cramon DY (2003) Error monitoring using external feedback: specific roles of the habenular complex, the reward system, and the cingulate motor area revealed by functional magnetic resonance imaging. J Neurosci 23:4308–4314.

Valentin VV, Dickinson A, O'Doherty JP (2007) Determining the neural substrates of goal-directed learning in the human brain. J Neurosci 27:4019–4026.

Vogt B (2008) Architecture, cytology and comparative organization of primate cingulate cortex. In: Cingulate neurobiology and disease (Vogt B, ed). Oxford, UK: Oxford UP.

Walton ME, Devlin JT, Rushworth MFS (2004) Interactions between decision making and performance monitoring within prefrontal cortex. Nat Neurosci 7:1259–1265.

Walton ME, Behrens TE, Buckley MJ, Rudebeck PH, Rushworth MF (2010) Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. Neuron 65:927–939.

Wise SP, Murray EA (2000) Arbitrary associations between antecedents and actions. Trends Neurosci 23:271–276.

Woolrich MW, Ripley BD, Brady M, Smith SM (2001) Temporal autocorrelation in univariate linear modeling of FMRI data. Neuroimage 14:1370–1386.

Wunderlich K, Rangel A, O'Doherty J (2011) Economic choices can be made using only stimulus values. Proc Natl Acad Sci U S A 107:15005–15010.