Behavioral/Cognitive

Ventromedial Prefrontal and Anterior Cingulate Cortex Adopt Choice and Default Reference Frames during Sequential Multi-Alternative Choice

Erie D. Boorman,[1,2,3] **Matthew F. Rushworth,**[1,2] **and Tim E. Behrens**[2,4]

[1]Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, United Kingdom, [2]Centre for Functional Magnetic Resonance Imaging of the Brain, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU, United Kingdom, [3]Computation and Neural Systems, California Institute of Technology, Pasadena, California 91125, and [4]Wellcome Trust Centre for Neuroimaging, London WC1N 3BG, United Kingdom

Although damage to the medial frontal cortex causes profound decision-making impairments, it has been difficult to pinpoint the relative contributions of key anatomical subdivisions. Here we use function magnetic resonance imaging to examine the contributions of human ventromedial prefrontal cortex (vmPFC) and dorsal anterior cingulate cortex (dACC) during sequential choices between multiple alternatives—two key features of choices made in ecological settings. By carefully constructing options whose current value at any given decision was dissociable from their longer term value, we were able to examine choices in current and long-term frames of reference. We present evidence showing that activity at choice and feedback in vmPFC and dACC was tied to the current choice and the best long-term option, respectively. vmPFC, mid-cingulate, and posterior cingulate cortex encoded the relative value between the chosen and next best option at each sequential decision, whereas dACC encoded the relative value of adapting choices from the option with the highest value in the longer term. Furthermore, at feedback we identify temporally dissociable effects that predict repetition of the current choice and adaptation away from the long-term best option in vmPFC and dACC, respectively. These functional dissociations at choice and feedback suggest that sequential choices are subject to competing cortical mechanisms.

Introduction

An abundance of research has begun to reveal the computational and neural mechanisms governing binary choice in situations where two options are presented anew on each trial (Platt and Huettel, 2008; Kable and Glimcher, 2009; Rangel and Hare, 2010). However, many real-world choices are made between multiple alternatives, and in situations where one option is in a privileged position. When shopping for cereal, for example, rather than reconsider every brand anew, we select our favorite brand, unless another seems temporarily more attractive. This strategy is computationally appealing as it obverts continual complex multi-alternative comparisons. Such sequential multi-alternative choices are ubiquitous in the real world, but their underlying neural substrates are poorly understood (but see Daw et al., 2006; Pearson et al., 2009).

The ventromedial prefrontal cortex (vmPFC) and dorsomedial frontal cortex (DMFC) have featured prominently in value-based binary choice studies (Rangel and Hare, 2010; Fellows, 2011; Rushworth et al., 2011; Rudebeck and Murray, 2011a). In several studies the blood oxygenation level-dependent (BOLD) response in the vmPFC and DMFC has correlated with the relative value between two decision options. Notably, however, as the difference between values of chosen and unchosen, or attended and unattended, options increases, the vmPFC signal increases whereas the DMFC signal decreases (Boorman et al., 2009; FitzGerald et al., 2009; Wunderlich et al., 2009; Lim et al., 2011). A major challenge is therefore understanding why opposite BOLD signals are frequently recorded in the two regions.

A separate literature has implicated dorsal anterior cingulate cortex (dACC) in behavioral adaptation. Both single cells and BOLD signals in dACC are particularly active when subjects receive information that leads to changes in beliefs or behavior (Behrens et al., 2007; Quilodran et al., 2008; Hayden et al., 2009, 2011b; Jocham et al., 2009; Wessel et al., 2012). Although such effects have been conceptualized as outcome monitoring signals, similar activity has recently been observed in monkey dACC neurons during foraging-style choices (Hayden et al., 2011a). Such an adaptation signal, while irrelevant for nonsequential binary decisions often studied in the laboratory, might be a key determinant of the sequential multi-alternative decisions commonly faced by humans and foraging animals, as it may instruct a change from a long-term or default position.

A recent functional magnetic resonance imaging (fMRI) study (Kolling et al., 2012) showed that dACC activity incorporated the average value of searching the environment, relative to

engaging with known options, and associated search costs—key variables for foraging—whereas vmPFC activity reflected the relative chosen value between two well defined options. In addition to these variables, however, behavioral ecologists emphasize that ecological choices are generally made sequentially (Freidin and Kacelnik, 2011), a crucial feature of foraging not investigated by Kolling and colleagues (2012). Furthermore, by decorrelating tractable reward probabilities and randomly generated reward magnitudes, we were able to dissociate short-term best options (those with highest expected values (reward probability × reward magnitude) from long-term best options (those with highest reward probabilities but not necessarily highest expected values). This manipulation enabled us to investigate for the first time the extent to which both choice and feedback signals in these regions might reflect default positions established over several trials or current choices.

We reasoned that examining vmPFC and dACC activity at both choice and feedback during sequential multi-alternative choice might further elucidate their respective contributions to ecological choice and bridge the parallel literatures on relative value and behavioral adaptation in dACC outlined above. We therefore designed an fMRI experiment that dissociated short-term best options from long-term best options and examined activity in vmPFC and dACC during multi-alternative and sequential choices that included or excluded the default option, and at feedback preceding choices in distinct reference frames.

Materials and Methods

Subjects. Twenty-two healthy volunteers participated in the fMRI experiment. Two volunteers failed to use either the reward probabilities or reward magnitudes in the task, and one failed to use reward probability, so their data were discarded from all analyses. The remaining 19 participants (10 women, mean age 25.2 years) were included in all further analyses. All participants gave informed consent in accordance with the National Health Service Oxfordshire Central Office for Research Ethics Committees (07/Q1603/11).

Experimental task. In our fMRI paradigm, participants decided repeatedly between three stimuli based on their reward expectation and the number of points associated with each stimulus option (see Fig. 1). Although the number of points was generated randomly (uniform distribution) and displayed on the screen, the probability had to be estimated from the recent outcome history. The true reward probabilities associated with each stimulus type varied independently from one trial to the next over the course of the experiment at a rate determined by the volatility, which was fixed in the current experiment. More specifically, the true reward probability of each stimulus was drawn independently from a β distribution with a fixed variance and a mean that was determined by the true reward probability of that stimulus on the preceding trial (see Fig. 2).To decorrelate the reward probabilities, reward magnitudes, and expected values associated with each option, and between simulated chosen and unchosen options, we used a simple Monte Carlo algorithm with 1000 simulations that aimed to reduce the summed absolute value of correlations between our variables of interest. We simulated choices using both greedy and softmax action selection rules (Sutton and Barto, 1998), inserting into the softmax function the inverse temperature fitted to subject choices from a previous study (Boorman et al., 2009). This procedure gave us some insight into the extent to which chosen, best unchosen, and worst unchosen variables would also be decorrelated. The true reward probabilities that subjects tracked and the estimated reward probabilities generated by the Bayesian learner are shown in Figure 2A, and the expected values (estimated reward probabilities × reward magnitudes) are shown in Figure 2B. The reward magnitudes and reward schedule that resulted from the selected reward probabilities were identical for each subject.

The three stimuli between which subjects chose were pictures of a real face, whole body, and house (Fig. 1). The identities of the face, body, and

house were fixed for the duration of the experiment and across participants (i.e., the same three stimuli were shown throughout the experiment). During the first decision-making phase, the three options and their associated points were displayed at three locations on the screen: left, upper middle, and right. The location at which each stimulus was displayed was randomized across trials. When the yellow question mark appeared in the center of the screen, subjects indicated their choices with right-hand finger responses on a button box corresponding to the location of each stimulus. Immediately after subjects indicated their choice, the first feedback phase was presented: the selected option was highlighted by a red rectangle that framed the chosen stimulus and the chosen outcome (reward or no reward) was presented. If the participant's choice was rewarded, a green tick appeared in the center of the screen, and the red prize bar also updated toward the gold rectangular target in proportion to the amount of points won on that trial. Each time the prize bar reached the gold target, participants were rewarded with £2. If the subject's choice was not rewarded, a red X appeared in the center of the screen. These initial decision-making and chosen feedback phases were presented on every trial in the experiment.

After presentation of the chosen feedback, one of three interleaved conditions followed in pseudorandom order. In condition 1 the outcomes for the two remaining unchosen options were presented. A green tick or a red X appeared on the left of the two options that were unchosen during the first decision-making phase, depending on whether they were rewarded or unrewarded. The red prize bar did not move. This event was followed by presentation of the next trial. In conditions 2 and 3, participants had the opportunity to choose between the two remaining options that were unselected by the participant at the first decision. These two remaining stimuli maintained their spatial locations on the screen. In condition 2 the option reward probabilities and points associated with the two options remained identical to the first decision (Fig. 1). This condition improved our ability to rank the two unchosen options at the first decision on the basis of expected value. However, in condition 3 only the reward probabilities remained the same; the points for both remaining options were changed to 50 (Fig. 1). This condition improved our ability to rank the two unchosen options at the first decision on the basis of reward probability. For both conditions 2 and 3, participants once again indicated their choice after a yellow question mark appeared. This was followed by simultaneous feedback for the chosen and unchosen options from the second decision. During this second feedback phase, a red rectangle framed the selected option and a green tick or red X was presented to the left of the chosen and unchosen options, depending on whether these options were rewarded or unrewarded. If the choice at the second decision was rewarded, the red prize bar updated in proportion to the number of points won. This event was followed by presentation of the next trial. There was no intertrial interval in any condition. Each event was jittered between 2.5 and 5.5 s (uniform distribution). There were 60 trials in each condition, making 180 trials in total. Conditions were pseudorandomly interleaved and were uncued. Participants earned between £20 and 28 on the task, depending on their performance.

Behavioral model. We used a previously described Bayesian reinforcement-learning algorithm (Behrens et al., 2007) to model subject estimates of the reward probabilities. Because feedback was provided on each option (at some stage) in every trial in our task, we assumed that beliefs concerning the reward probabilities associated with each option were updated equally, as is optimal. To assess this assumption, we constructed an additional model in which separate learning rates scaled chosen and unchosen prediction errors (Boorman et al., 2011), which yielded similar maximum likelihood estimates for chosen and unchosen learning rates ($t_{(18)} < 0.25$, $p < 0.4$).

We used a previously published Bayesian model to generate estimates of the reward probabilities. This algorithm has been documented in detail previously (Behrens et al., 2007), but we briefly describe its concept here. The model assumes that outcomes are generated with an underlying probability, r. The objective is to track r as it changes through time. The crucial question addressed by the model is how much the estimate of r should be updated when a new positive or negative outcome is observed. An unexpected event may be just chance or it may signal a change in the underlying reward probability. To know how much to update the

estimate of r on witnessing a new outcome, it is crucial to know the rate of change of r. If r is changing fast on average then an unlikely event is more likely to signify a big change in r so an optimal learner should make a big update to its estimate. The Bayesian model therefore maintains an estimate of the expected rate of change of r, referred to as the volatility, v. In a fast-changing environment, the model will estimate a high volatility and therefore each new outcome will have a large influence on the optimal estimate of the reward rate. Conversely in a slow-changing environment, the model will estimate a low volatility and each new outcome will have a negligible effect on the model's estimate of r.

To allow for the possibility of interindividual differences in how people combine reward probability and reward magnitude, we included subject-specific free parameters that can differentially weigh probability, magnitude, and their product, to derive estimates of the subjective expected values. In addition, we initially considered the possibility that participants took into account the likelihood of encountering a second decision when making initial decisions, for example, as seen in the following:

$$g_{sc_i} = \beta r_{sc_i} + \lambda m_{sc_i} + \gamma r_{sc_i} m_{sc_i} + \ldots 0.5\tau[(\beta r_{sb_i} + \lambda m_{sb_i}$$

$$+ \gamma r_{sb_i} m_{sb_i}) + (\beta r_{sw_i} + \lambda m_{sw_i} + \gamma r_{sw_i} m_{sw_i}) + (\beta r_{sb_i} + \lambda 50$$

$$+ \gamma r_{sb_i} 50) + (\beta r_{sw_i} + \lambda 50 + \gamma r_{sw_i} 50)],$$

where, g_{sc_i}, r_{sc_i}, and m_{sc_i} are the subjective value, reward probability, and reward magnitude associated with the chosen stimulus at trial i, and r_{sb_i} and m_{sb_i} are the reward probability and reward magnitude associated with the best unchosen stimulus and r_{sw_i} and m_{sw_i} are the reward probability and reward magnitude associated with the worst unchosen stimulus at trial i. Therefore τ reflects the weight attributed to the possibility of encountering a second decision. We did not find any evidence that participants' choices at the first decision were influenced by the prospect of a second decision at which reward magnitudes could either remain the same or both change to 50: estimated values of τ equaled 0 or nearly 0 in each participant. We therefore assumed that subjective value at both decisions was computed on the basis of the current decision alone as seen in the following:

$$g_{s_i} = \beta r_{s_i} + \lambda m_{s_i} + \gamma r_{s_i} m_{s_i},$$

where, g_{s_i}, r_{s_i}, and m_{si} are the subjective value, reward probability, and reward magnitude associated with the stimulus (face, house, or body) on trial i. β, λ, and γ are the only free parameters in the behavioral model. We fitted β, λ, and γ to each individual subject's behavioral data using standard nonlinear minimization procedures in MATLAB (Mathworks). Similarly, for the fMRI analysis, we can define g_{sc} as the subjective EV of the chosen stimulus, g_{sb} as the subjective EV of the unchosen option with the next highest EV, and g_{sw} as the subjective EV of the unchosen option with the lowest EV. Finally, the selector component of the model assumed that subjects chose stimulus s according to the following softmax probability distribution:

$$P(s) \frac{\exp(g_s)}{\sum_{s'=1}^{N_s} \exp(g_{s'})}$$

where g_s is the subjective expected value of the stimulus, and N_s is the total number of stimuli to choose between ($N_s = 3$ at the first decision, $N_s = 2$ at the second decision).

fMRI data acquisition and preprocessing. fMRI data were acquired on a 3 T Siemens TRIO scanner with a voxel resolution of $3 \times 3 \times 3$ mm^3, TR = 3 s, TE = 30 ms, flip angle = 87°. The slice angle was set to 15° and a local z-shim was applied around the orbitofrontal cortex (OFC) to minimize signal dropout in this region (Deichmann et al., 2003), which has previously been implicated in other aspects of decision making. The mean number of volumes acquired was 999, giving a mean total experiment time of ~50 min.

We acquired field maps using a dual echo 2D gradient echo sequence with echoes at 5.19 and 7.65 ms, and a repetition time of 444 ms. Data were acquired on a $64 \times 64 \times 40$ grid, with a voxel resolution of 3 mm

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

Data were preprocessed using the default options in FMRIB's Software Library (FSL): motion correction was applied using rigid body registration to the central volume (Jenkinson et al., 2002); Gaussian spatial smoothing was applied with a full-width half-maximum of 5 mm; brain matter was segmented from nonbrain using a mesh deformation approach (Smith, 2002); and highpass temporal filtering was applied using a Gaussian-weighted running lines filter, with a 3 dB cutoff of 100 s.

fMRI data analysis. fMRI analysis was performed using FSL (Jenkinson et al., 2012) using the default settings. A general linear model (GLM) was fit in prewhitened data space. Twenty regressors were included in the GLM: the main effect of the first decision-making phase, the main effect of the first feedback phase, the main effect of the foregone outcome phase (condition 1), the main effect of the second decision-making phase (conditions 2 and 3), the main effect of the second feedback phase (conditions 2 and 3), the interaction between subjective EV associated with the chosen stimulus (g_{sc} from here on, value) and the first decision-making phase, best unchosen value (g_{sb}) as determined by the model in conditions 1 and 3 and the first decision-making phase, best unchosen value as determined by subject choices in condition 2 and the first decision-making phase, worst unchosen value as determined by the model in conditions 1 and 3 and the first decision-making phase, worst unchosen value (g_{sw}) as determined by subject choices in condition 2 and the first decision-making phase, chosen value from condition 2 and the second decision-making phase, chosen value from condition 3 and the second decision-making phase, unchosen value from condition 2 and the second decision-making phase, unchosen value from condition 3 and the second decision-making phase, and six motion regressors produced during realignment. There were no notable differences between the z-statistic maps based on the model or subject choices, so we defined additional contrasts of parameter estimates for the best unchosen and worst unchosen values as the sum of regressors based on the model and subject choices. Similarly, we defined the chosen and unchosen value at the second decision as the sum of regressors based on conditions 2 and 3. Aside from the motion regressors, all regressors were convolved with the FSL default hemodynamic response function (gamma function, delay = 6 s, SD = 3 s), and filtered by the same highpass filter as the data.

Region of interest analysis. To identify regions of interest (ROIs) for further time course analysis, we performed an initial group analysis of chosen value at decision, which revealed positive effects ($Z > 3.1$, $p < 0.001$ uncorrected, extent > 10 voxels) in regions within vmPFC, as well as mid-cingulate (caudal cingulate zone) and posterior cingulate cortex (PCC), and negative effects in dACC (rostral cingulate zone). These criteria were simply used to identify vmPFC and dACC regions, as well as any additional potential regions of interest for subsequent analyses. We then used a leave-one-out extraction procedure to provide an independent criterion for voxel selection from these regions (Kriegeskorte et al., 2009): in each participant, BOLD signal was extracted from a sphere (radius = 3 mm) centered on the peak voxel within the ROI for the contrast of chosen expected value at decision in a group model that excluded that participant. This procedure obviates questions of multiple comparisons, as peaks are selected from one set of data ($n - 1$ subjects) and tested in the independent left-out dataset. We then performed independent statistical tests (see description below) and characterized the time course of BOLD fluctuations in these regions.

Each subject's BOLD time series for a given ROI was divided into trials, which were resampled to 300 ms and truncated based on the mean trial length for each condition across trials and subjects. In the resampled time series, trial events were aligned based on their mean onset times across trials and subjects. A GLM was then fit across trials in each subject independently, which resulted in parameter estimates at each time point for each regressor included in the GLM. We then calculated group mean effect sizes at each time point and their SEs, which are plotted in Figures 3, 5, 6. To ascertain which variables were reflected in BOLD activity in a given ROI, we then fit the BOLD effect of interest with a canonical hemodynamic response function (gamma function) aligned to the onset of the event in each subject and calculated resulting t statistics and p values.
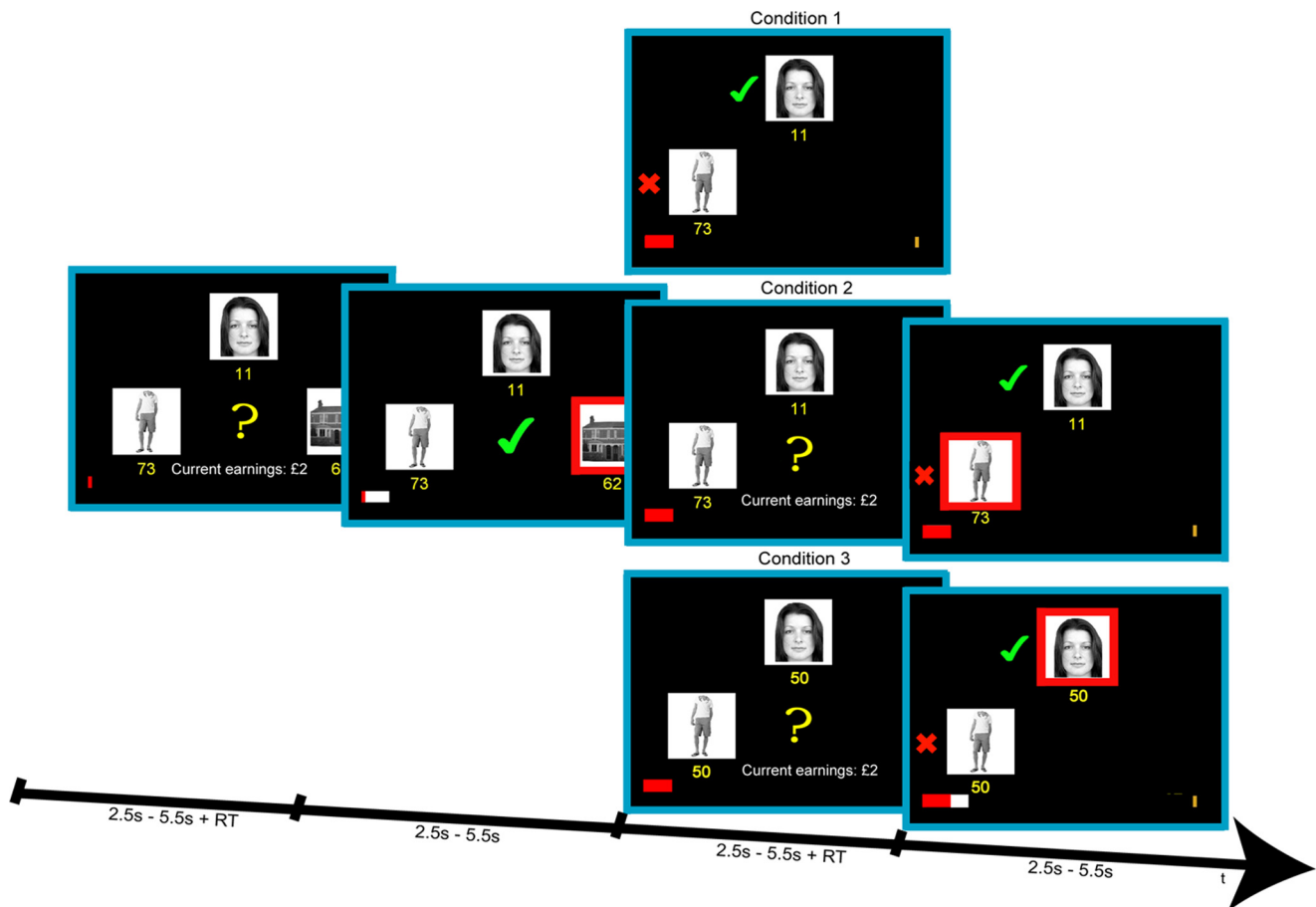
Figure 1. Experimental task. Subjects were faced with decisions between a face, whole-body, and house stimulus, whose locations on the screen were randomized across trials. Subjects were required to combine two pieces of information: the reward magnitude associated with each choice (which was generated randomly and shown in yellow beneath each stimulus) and the reward probability associated with each stimulus (which drifted independently from one trial to the next (Fig. 2A) and could be estimated from the recent outcome history). When the yellow question mark appeared, subjects could indicate their choices. The selected option was then highlighted by a red frame and the outcome was presented: a green tick or a red X indicating a rewarded or unrewarded choice. If the choice was rewarded, the red points bar at the bottom of the screen updated toward the gold rectangular target in proportion to the number of points won. Each time the bar reached the target, subjects were rewarded with £2. One of three conditions followed in pseudorandom order. In condition 1 the outcomes (rewarded or unrewarded) of the two unselected options were presented to the left of each stimulus, followed by the next trial. The points bar did not move. In conditions 2 and 3 subjects choose between the two options they had foregone at the first decision. In condition 2 the points associated with each stimulus remained the same as at the first decision. In condition 3 the points both changed to 50. In both conditions 2 and 3, when the yellow question mark appeared for the second time, subjects could indicate their second choices. This was followed immediately by feedback for both the chosen option, which was highlighted by a red frame, and the unchosen option, to the left of each stimulus. If the chosen option at the second decision was rewarded, the red points bar moved toward the target in proportion to the number of points won.

Nine GLMs were tested on ROIs identified using the procedure detailed above. The following analyses were conducted across conditions 1–3: GLM1 (conducted on vmPFC, mid-cingulate, and PCC): $y = B_0 + B_1 g_{sc} + B_2 g_{sb} + B_3 g_{sw} + e$; GLM2 (vmPFC and dACC): $y = B_0 + B1(g_{sc} - g_{sb}) + B_2(g_{d1} - g_{d2}) + e$; GLM3 (vmPFC): $y = B_0 + B_1 r_{sc} + B_2 m_{sc} + B_3 r_{sb} + B_4 m_{sb} + B_5 r_{sw} + B_6 m_{sw} + e$; GLM4 (dACC): $y = B_0 + B_1 g_{d1} + B_2 g_{d2} + B_3 g_{d3} + e$; GLM5 (dACC): $y = B_0 + B_1 r_{d1} + B_2 m_{d1} + B_3 r_{d2} + B_4 m_{d2} + B_5 r_{d3} + B_6 m_{d3} + e$; GLM6 (dACC): $y = B_0 + B1(g_{sc} - g_{sb}) + B_2(g_{d1} - g_{d2}) + B3(-|g_{sc} - g_{sb}|) + e$. The following GLM was conducted across conditions 2–3 (i.e., when there was a second decision): GLM7 (vmPFC, mid-cingulate, PCC, and dACC): $y = B_0 + B_1 g_{sc} + B_2 g_{dec2sc} + B_3 g_{dec2suc} + e$. The following GLM was conducted on trials on which the default option was still on offer at the second decision and when it was not: GLM8 (dACC): $y = B_0 + B1(g_{d1} - g_{d2}) + e$. The following GLM was also conducted across conditions 2–3: GLM9 (vmPFC and dACC): $y = B_0 + B_1 g_{sc} + B_2 g_{dec2sc} + B_3 g_{dec2suc} + B_4 ch + B_5 unch + B_6 repch + B_7 swdf + e$. g refers to expected value, m to reward magnitude, and r to reward probability. For GLMs 1–3, sc, sb, and sw refer to the chosen stimulus, the best unchosen stimulus, and the worst unchosen stimulus, respectively. For GLMs 3–6 and 8, d1, d2, and d3 refer to default options 1, 2, and 3. For GLMs 7–8, dec2sc and dec2suc refer to the chosen and unchosen stimulus at decision 2, respectively. For GLM 8, ch and unch refer to binary chosen and unchosen outcome regressors, and repch

and swdf refer to trials preceding choice repetitions and switches away from the default, respectively.

Results
Behavior
We designed an fMRI task that required participants to first choose between three options on the basis of both reward probability (which changed independently from trial to trial and had to be learned from chosen and unchosen outcomes) and reward magnitude (which was generated randomly and presented explicitly on the screen) and then choose between the remaining two foregone options (Fig. 1). To obtain estimates of subjects' trial-by-trial valuations of each of the options, we compared the fit to behavior of three competing reinforcement-learning models of subject behavior. Participants' choices were best described by a Bayesian model that updates both chosen and unchosen options (see Materials and Methods; Table 1). The focus of a previously published paper (Boorman et al., 2011) reported analyses that allowed us to determine the role of the lateral anterior prefrontal cortex and associated brain regions in counterfactual

Table 1. Behavioral model fits

Model	Parameters per subject	Log likelihood	BIC
Optimal Bayesian	3 (0 predictor)	−2080	4175.6
Experiential Bayesian	3 (0 predictor)	−2265.8	4547.1
Rescorla–Wagner	4 (1 predictor)	−2634.2	5289.1

Comparison of model fits for the complete Bayesian, experiential Bayesian, and Rescorla–Wagner models. Less negative log likelihoods and lower Bayesian information criterion (BIC) indicate superior fits. Predictor refers to the reward probability learning algorithm.

choices and learning. Here, by contrast, we present the results of a different set of analyses that allowed us to examine distinct aspects of behavior and distinct aspects of brain activity: (1) the value difference signals that should guide decisions on each trial and trial-to-trial switching tendencies and (2) the roles of the vmPFC and dACC.

We first examined to what extent reward probabilities and reward magnitudes (Fig. 2A,B) associated with each of the three options influenced behavior and whether subjects exhibited any systematic bias toward the long-term best option. In our task the reward magnitudes change randomly on each trial but *the reward probabilities* fluctuate slowly through the experiment. The long-term best option is therefore the one with the highest reward probability, which we refer to as the default option. This is of course not necessarily the option with the highest overall expected value ($EV = reward\ probability \times reward\ magnitude$), which is the short-term best option. To test for a bias in favor of the default option, we included the reward probabilities and reward magnitudes associated with the default option alongside the reward probabilities and reward magnitudes associated with the best (the option with the highest EV), mid (the option with the next highest EV), and worst (the option with the lowest EV) options in a logistic regression on choices of the short-term best option. This analysis revealed that choices of the short-term best option were driven by the reward probability and reward magnitude associated with the best (reward probability: $t_{(18)} = 6.55, p < 0.0001$; reward magnitude: $t_{(18)} = 6.03, p < 0.0001$) and next best options (reward probability: $t_{(18)} = -3.03, p < 0.005$; reward magnitude: $t_{(18)} = -4.31, p < 0.0005$), and the reward magnitude but not reward probability associated with the worst option (reward probability: $t_{(18)} = 0.74, p > 0.1$; reward magnitude: $t_{(18)} = 4.17, p < 0.0005$) (Fig. 2C). Importantly, there was also a significant negative effect of the default option's reward probability ($t_{(18)} = -1.93, p < 0.05$) but not reward magnitude ($t_{(18)} = 1.2, p > 0.1$), indicating that people were biased away from choosing the best option when the default option had a high reward probability. This effect was present even after accounting for any influence of the reward probabilities and reward magnitudes associated with the best, mid, and worst options, and even after including choice difficulty as an additional nuisance regressor in a second logistic regression (effect of default option's reward probability: $t_{(18)} = -1.85, p < 0.05$). Behavior was therefore not only guided by the reward probabilities and reward magnitudes associated with the best and mid options, as is optimal, but also biased away from the best option when there was an alternative strong default preference.

We further reasoned that subjects might exhibit a general bias toward choosing the default option. To test this possibility, we examined the proportion of choices of the short-term best, mid, and worst options when they were also the default option and when they were not (Fig. 2D). A two-way repeated ANOVA revealed a main effect of both default type (default, nondefault) ($F_{(1,18)} = 26.15, p < 0.0001$) and rank (best, mid, worst) ($F_{(2,18)} = 169.43, p < 0.00001$), but no interaction ($F_{(2,18)} = 1.7, p > 0.1$).

Importantly, the main effect of default type remained after matching both option values ($F_{(1,18)} = 10.9, p = 0.001$) and difficulty ($F_{(1,18)} = 15.06, p = 0.0002$) across default and non-default options.

Value coding in the BOLD signal

We hypothesize that separable decision processes exist in vmPFC and dACC relating, respectively, to a comparison between current options and behavioral adaptation from a long-term default option. If this is the case, then several features of our task make specific predictions of the BOLD signal in the two structures. First, vmPFC value signals will take the frame of reference of the current choice (i.e., they will reflect the chosen value relative to the value of one or both of the other options), but dACC signals will be referenced according to the default choice. Second, during choices when the default option is not present, the vmPFC but not the dACC will maintain its relative value coding. Third, when choice outcomes are observed, signals in vmPFC will influence future choices of the chosen option, but dACC signals will influence future choices of the default option.

To test these hypotheses, we needed to identify ROIs in a way that would not bias future comparisons. To ensure this was the case, we performed a leave-one-out procedure to identify ROIs in each individual that were defined from a group model incorporating all subjects except that individual (see Materials and Methods). This test obviates questions of multiple comparisons, as peaks are selected from one set of data and tested in the independent alternative dataset. This approach revealed two sets of ROIs. First, the BOLD response in a network of regions including the vmPFC (X,Y,Z; 2, 32, −10), corresponding to cluster 2 in Figure 3A (Beckmann et al., 2009), mid-cingulate [0, −10, 46, putative caudal cingulate zone as termed by Picard and Strick (2001), encompassing parts of clusters 5 and 6 in Fig. 3A (Beckmann et al., 2009)], and PCC [14, −30, 42, PCC, encompassing parts of clusters 8 and 9 in Fig. 3A (Parvizi et al., 2006; Beckmann et al., 2009; Fig. 3B; Table 2)] correlated positively with chosen expected value (chosen value). Second, a circumscribed region in dACC [−6, 24, 34, putative anterior rostral cingulate zone, corresponding to cluster 4 in Fig. 3A (Picard and Strick, 2001; Beckmann et al., 2009; Fig. 3C; Table 2)] correlated negatively with chosen value. We extracted signal from these ROIs to test the key hypotheses relating to the two putative competing choice mechanisms.

Relative value during multi-alternative choice

We performed a multiple regression on the BOLD time course from the ROIs defined above. We first looked to see whether signal fluctuations contained a reflection of all of the different values available during the choice (*GLM1*, see Materials and Methods). These analyses revealed a significant positive effect of the chosen EV (Vch; $t_{(18)} = 2.30, p < 0.05$), a negative effect of the best unchosen option's EV (V2; $t_{(18)} = -2.37, p < 0.05$), but no significant effect of the worst unchosen option's EV (V3; $t_{(18)} = -1.2, p > 0.1$) in vmPFC (Fig. 3D). There were similar effects in mid-cingulate and PCC (Vch: mid-cingulate: $t_{(18)} = 2.36, p < 0.05$; PCC: $t_{(18)} = 1.32, p = 0.1$; V2: mid-cingulate: $t_{(18)} = -1.79 p < 0.05$; PCC: $t_{(18)} = -2.18, p < 0.05$; V3: mid-cingulate: $t_{(18)} = -0.50, p > 0.1$; PCC: $t_{(18)} = -0.93, p > 0.1$).

To examine whether these regions might bear specific influence on default choices, we also considered an alternative coding scheme according to which options are ranked in the long term, as opposed to the current choice. As described above, the long-term best option in our task is the one with the highest reward
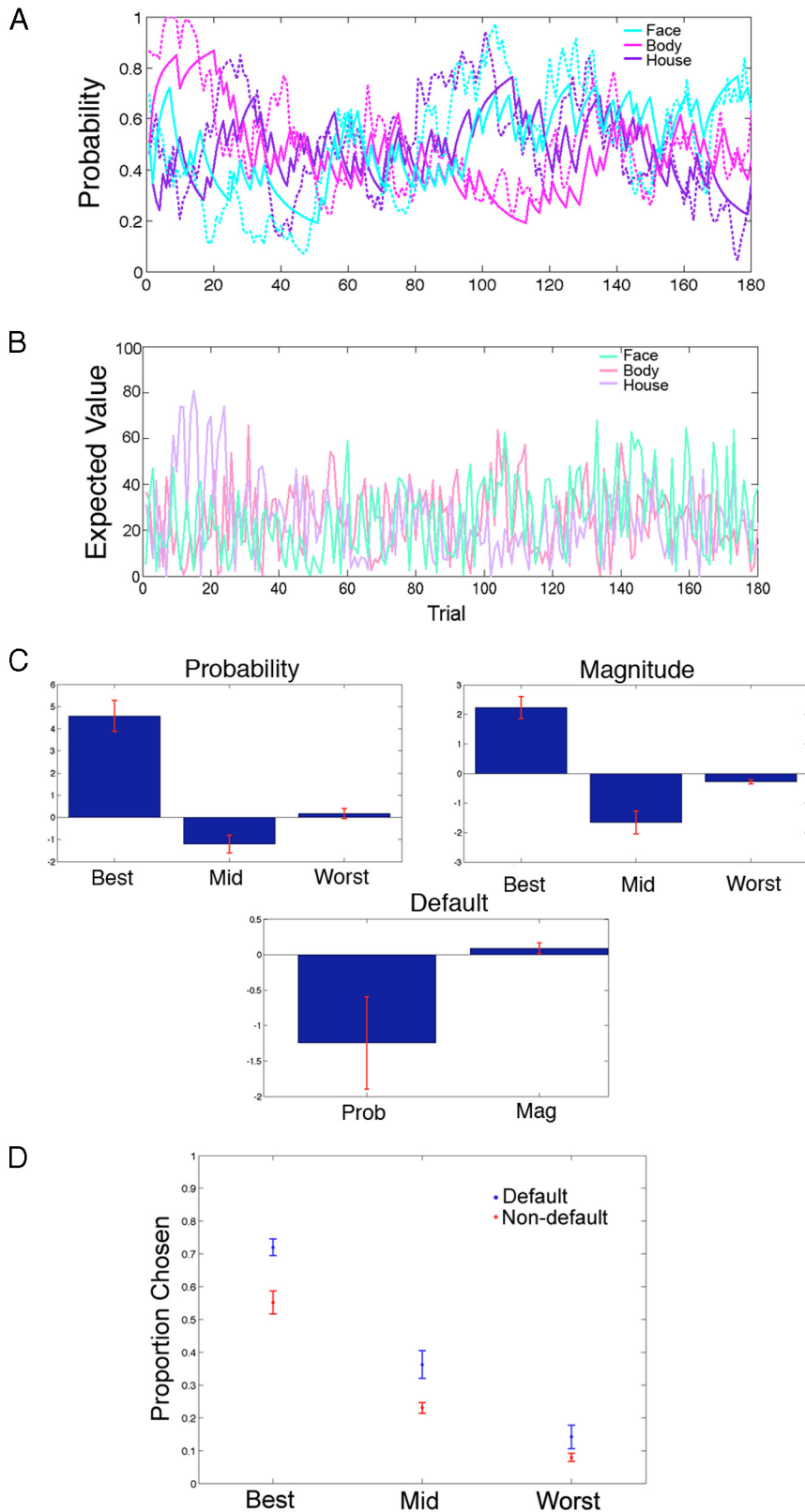
Figure 2. Reward probability, expected value, and behavior. ***A***, True reward probabilities associated with the face (cyan), body (pink), and house (magenta) stimuli across the experiment are plotted (dotted lines) alongside reward probabilities as estimated by the Bayesian learner (solid lines). ***B***, Expected values (reward magnitudes × reward probabilities as estimated by the Bayesian learner) are plotted. ***C***, Regression coefficients from a logistic regression on optimal choices (i.e., choices of the best option) for reward probabilities (top left) and reward magnitudes (top right) associated with the best, middle, and worst options (based on their expected value) and default option reward probability and magnitude (bottom). ***D***, Proportion of choices of the short-term best, mid, and worst options are plotted when they are also the default option (blue) compared with when they are not the default option (red).

probability (but not necessarily the highest *EV*). It is therefore possible to define two regressors that can be dissociated in our data: the relative chosen value (Vch − V2) and the relative default value (the most probable option's *EV* minus the next most probable option's *EV*: default V1 − default V2). When we allowed the relative chosen value and the relative default value to compete in the same GLM (*GLM2*), we found a significant positive effect of the relative chosen value in vmPFC ($t_{(18)}$ = 3.82, $p < 0.001$), mid-cingulate ($t_{(18)}$ = 2.46, $p < 0.05$), and PCC ($t_{(18)} = 1.79, p < 0.05$), but not of the relative default value (all abs ($t_{(18)}) < 1.0, p > 0.1$; Fig. 4). These relative chosen value effects reflected both relative reward probability and relative reward magnitude in vmPFC (*GLM3*; reward probability: $t_{(18)} = 2.79, p < 0.01$; reward magnitude ($t_{(18)} = 2.17, p < 0.05$), replicating previous findings (Boorman et al., 2009), and PCC (reward probability: $t_{(18)} = 2.84, p < 0.01$; reward magnitude: $t_{(18)} = 1.94, p < 0.05$), with a similar trend in mid-cingulate (reward probability: $t_{(18)} = 1.98, p < 0.05$; reward magnitude: $t_{(18)} = 1.57, p < 0.1$). vmPFC, mid-cingulate, and PCC activity thus reflects the chosen *EV* relative to the next best option's *EV* during multi-alternative choice. Importantly, activity in these regions did not reflect action (or position) values [index finger (left position), middle finger (upper center position), ring finger (right position)] or stimulus (body, house, face) values (all $t < 1.5, p > 0.05$), indicating that activity in these regions is tied to the choice, rather than the action or stimulus, in this study.

The coding scheme in dACC was notably different. Time course analyses on ROIs identified using the leave-one-out procedure described above revealed that dACC activity was indeed influenced by long-term behavioral strategies. Including both relative chosen value and relative default value in the same GLM (*GLM2*) revealed a strong significant negative effect of the relative default value (default V1 − default V2; $t_{(18)} = −3.77, p < 0.001$) and only a trend toward an effect of relative chosen value ($t_{(18)} = −1.55, p < 0.1$; Fig. 4). These relative default value effects in dACC were driven by both relative reward probabilities ($t_{(18)} = 2.03, p < 0.05$) and relative reward magnitudes ($t_{(18)} = 4.3, p < 0.0005$) between default V1 and default V2 (*GLM5*). Including default values of all three options in the same GLM (*GLM4*) revealed a strong negative effect of default V1 ($t_{(18)} = −2.84, p = 0.005$), a strong positive effect of default V2 ($t_{(18)} = 3.84, p < 0.001$), and a modest
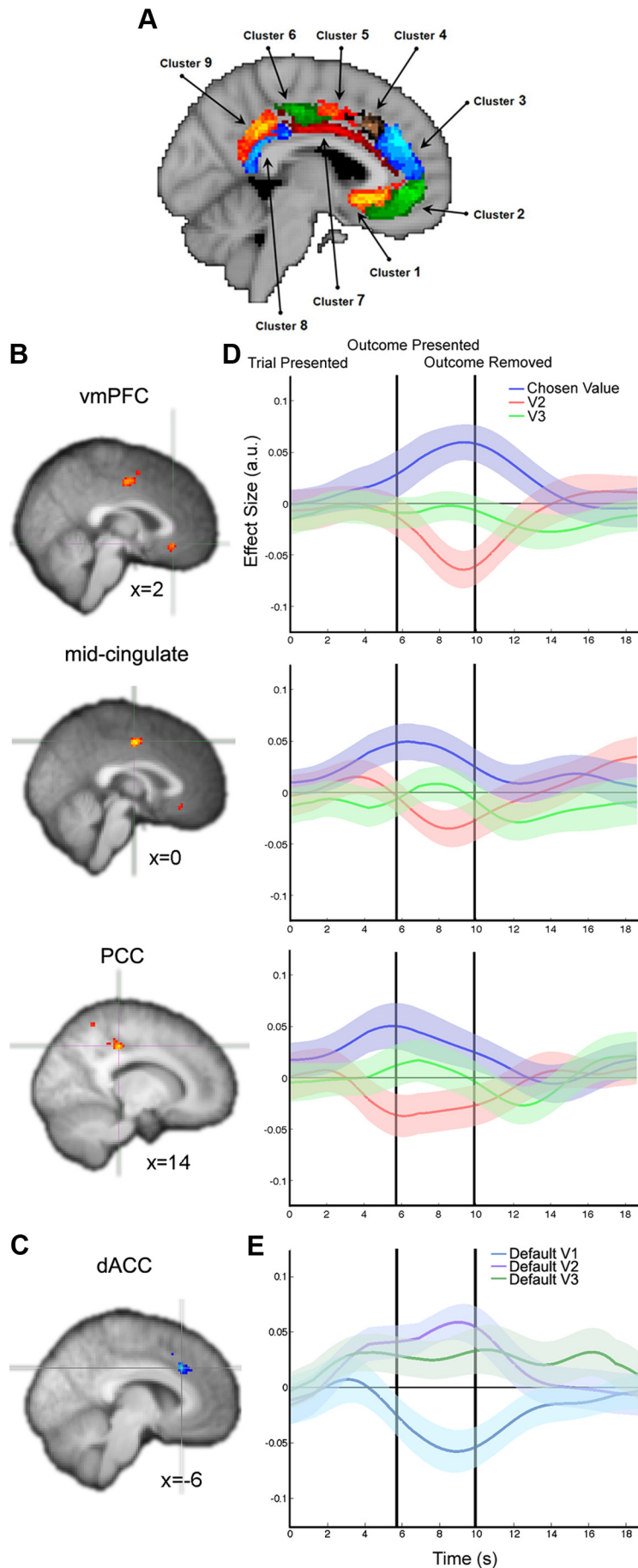
but significant positive effect of default V3 ($t_{(18)} = 1.93, p < 0.05$) in dACC (Fig. 3E). dACC activity is therefore best described by the present evidence [(default V2 + default V3) − default V1] favoring adaptation away from the default option in the environment. Finally, we formally compared the evidence favoring either coding scheme (relative chosen or relative default value) in vmPFC and dACC. To do so we inverted the sign of the regression coefficients in dACC so that both regions would encode relative chosen and relative default value positively (i.e., in the same direction). A 2 × 2 ANOVA of these regression coefficients revealed a significant interaction between region and coding scheme ($F_{(2,18)} = 7.57, p < 0.01$; Fig. 4). Together, these analyses reveal relative value effects in vmPFC and dACC with dissociable reference frames during trinary choice.

It has been suggested that dACC activity may reflect task difficulty. To control for this alternative explanation, we allowed relative default value (default V1 − default V2) and choice difficulty ($-abs(Vch − V2)$) to compete in the same GLM (*GLM6*), which revealed a very strong effect of relative default value ($t_{(18)} = -4.29, p < 0.0005$), but no significant effect of choice difficulty ($t_{(18)} = 0.56, p = 0.29$). If, alternatively, choice difficulty is defined based on the difference between the chosen and average value of the other options ($-abs(Vch − mean(V2,V3))$), then there is still a robust effect of relative default value ($t_{(18)} = -4.31, p < 0.0005$), but the effect of choice difficulty does not approach significance ($t_{(18)} = 0.77, p = 0.28$). Yet another possibility is to assume that task difficulty is defined based on default (long-term best) options ($-abs(default V1 − default$

←

Figure 3. Value coding during multi-alternative choice. ***A***, Reference image for comparison with ***B*** and ***C*** showing diffusion-weighted imaging-based parcellation of the cingulate cortex based on clustering of probabilistic connectivity profiles (adapted from Beckmann et al., 2009). ***B***, ***C***, Sagittal slices through *z*-statistic maps relating to subjective *EV* of the chosen option during decisions. Maps resulted from a random effects group analysis that included all 19 subjects. Positive effects are shown in red–yellow (***B***) and negative effects in blue–light blue (***C***). Maps are thresholded at $Z > 2.8$, $p < 0.003$ for display purposes, and are displayed according to radiological convention. ***D***, Time course of the effect size of the chosen *EV*, V2, and V3, generated using a leave-one-out procedure (see Materials and Methods), is plotted across the first decision-making and feedback phases in vmPFC, mid-cingulate, and PCC. ***E***, The same is shown for the default V1, default V2, and default V3 in dACC. Thick lines, mean; shadows ± SEM.

Table 2. Activation details

Comparison	Anatomical region	Cluster extent (voxels)	Hemisphere	Peak coordinates (mm) (x, y, z)	Maximum Z score
Chosen subjective value (subjective expected value of the chosen option during the decision phase)	vmPFC	11	R	$2, 32, -10$	3.23
	PCC	18	R	$14, -30, 42$	3.66
	Mid-cingulate	57	N/A	$0, -10, 46$	3.73
Inverse chosen subjective value	dACC	14	R	$-6, 24, 34$	3.43

Details of regions identified from a whole-brain group-level analysis on all 19 subjects, for positive and negative effects of the contrast of chosen value at decision. Regions are listed that survived voxel-based thresholding of $Z > 3.1, p < 0.001$ uncorrected, extent > 10 voxels. It should be noted that this procedure was only used to identify ROIs for further tests using the leave-one-out procedure described (see Materials and Methods). dACC, dorsal anterior cingulate cortex; PCC, posterior cingulate cortex; vmPFC, ventromedial prefrontal cortex.
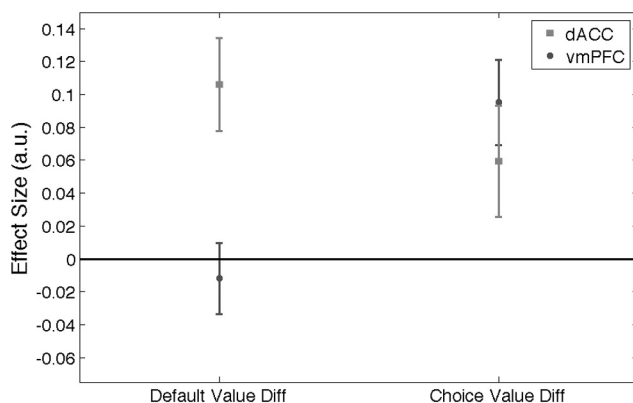

Figure 4. Comparison of chosen and default reference frames. Effect sizes of the relative chosen value (chosen EV − next best option's EV) and inverse relative default value (default V2 − default V1) at the first decision are plotted for vmPFC (dark gray) and dACC (light gray). Both regressors were entered in the same multiple regression without orthogonalization, since there was a substantial orthogonal component.

V2)) rather than the choice. Once again, we observe a very strong effect of relative default value ($t_{(18)} = -4.53, p < 0.0005$), but no significant effect of long-term task difficulty ($t_{(18)} = -0.14, p = 0.44$), indicating that relative default value effects in dACC cannot be explained by task difficulty.

Relative value at each decision
A brain region that compares options for the current decision would be expected to do so whenever a choice is made. We therefore examined the time course of the chosen and unchosen value effects from the second decision across both the initial and subsequent decisions that followed on two-thirds of trials (conditions 2 and 3). The coding of the best unchosen value in the vmPFC, mid-cingulate, and PCC flips from negative at the first decision to positive at the second decision when it becomes chosen (GLM7; negative effect of V2 at decision 1 in vmPFC: $t_{(18)} = -3.03, p < 0.005$; mid-cingulate $t_{(18)} = -1.93, p < 0.05$; PCC: $t_{(18)} = -2.22, p < 0.05$; positive effect of chosen value at decision 2 in vmPFC: $t_{(18)} = 1.83, p < 0.05$; mid-cingulate: $t_{(18)} = 4.57, p > 0.0005$; PCC: $t_{(18)} = 3.31, p < 0.005$), whereas the worst unchosen value begins to be coded negatively when it becomes the only unchosen option at the second decision (Fig. 5A; negative effect of unchosen value at decision 2 in vmPFC: $t_{(18)} = -2.25, p < 0.05$; mid-cingulate: $t_{(18)} = -2.82, p < 0.01$; PCC: $t_{(18)} = -1.83, p < 0.05$). These analyses provide strong evidence

that the vmPFC and associated regions flexibly encode the relative chosen value of the current decision, and moreover, that the relative value signal consists of the two options' values most pertinent to the comparison that is being performed (Fig. 2C).

Conversely, in the dACC there was no evidence of a flip in coding of the best unchosen option's value from the first to second decision ($t_{(18)} > 0.3, p > 0.1$) and no evidence of a relative chosen value effect at the second decision ($t_{(18)} > 0.25, p > 0.1$; Fig. 5B). Examination of the Z-stat map relating to the chosen value, unchosen value, or (unchosen value − chosen value) at the second decision failed to identify effects in any region of DMFC even at the reduced threshold of $Z > 2.3, p < 0.01$, uncorrected. Furthermore, the absolute relative chosen value effect at the second decision in vmPFC was significantly stronger than the absolute relative chosen value effect in dACC ($t_{(18)} = 2.20, p < 0.05$). Notably, there was a trend toward an effect of the relative default value when the default option was still on offer at second decisions (GLM8; $t_{(18)} = -1.40, p < 0.1$), despite the largely reduced number of such trials, but not when it had already been chosen at the first decision and was therefore no longer available ($t_{(18)} > -0.35, p > 0.1$). Indeed, on these latter trials, we were unable to find any effects of relative or absolute value in either frame of reference (all $t_{(18)} > 0.25, p > 0.1$). These analyses suggest that during repetitive decision making, dACC encodes relative value only when the default option is available to choose.

Future choices
The data presented so far are suggestive of distinctive roles for the vmPFC and dACC, with vmPFC activity reflecting variables relevant to the decision at hand and dACC activity reflecting variables relevant to the long-term best option. If vmPFC is concerned with comparing the currently available choice options, then it might also determine whether it is worth continuing with that most recent choice on the subsequent trial. If dACC plays a role in determining whether to adapt behavior away from the default option in the environment, conversely, then dACC activity might predict future switches away from that option. We therefore tested whether there was an effect on trial n of repeating or switching choices on trial $n + 1$ with respect to either the most recent choice made on trial n or the default option. At the time feedback was presented for the second choice in a trial, there was elevated vmPFC activity preceding trials on which subjects went on to repeat their most recent choice, relative to trials on which subjects went on to switch from their most recent choice (Fig. 6A; GLM9 $t_{(18)} = 2.63, p < 0.01$). However, there was no effect at feedback preceding trials on which they would repeat or switch away from the default option ($t_{(18)} < 1.0, p > 0.1$). In contrast, analysis of the dACC time course revealed a significant effect of switching away from the long-term best option that also occurred immediately before the onset of the new trial (Fig. 6B; $t_{(18)} = 2.40, p = 0.01$), but not of repeating or switching away from the most recent choice ($t_{(18)} < 1.0, p > 0.1$). Formal comparison of the two coding schemes at feedback revealed a significant interaction ($F_{(2,18)} = 7.58, p < 0.01$). Importantly, the chosen and counterfactual outcomes and their EVs were included in the respective GLMs, so these perifeedback effects cannot be explained by either chosen or counterfactual outcomes or their expectations.

Discussion
We have identified a functional dissociation between vmPFC and dACC during sequential multi-alternative choice. vmPFC ad-
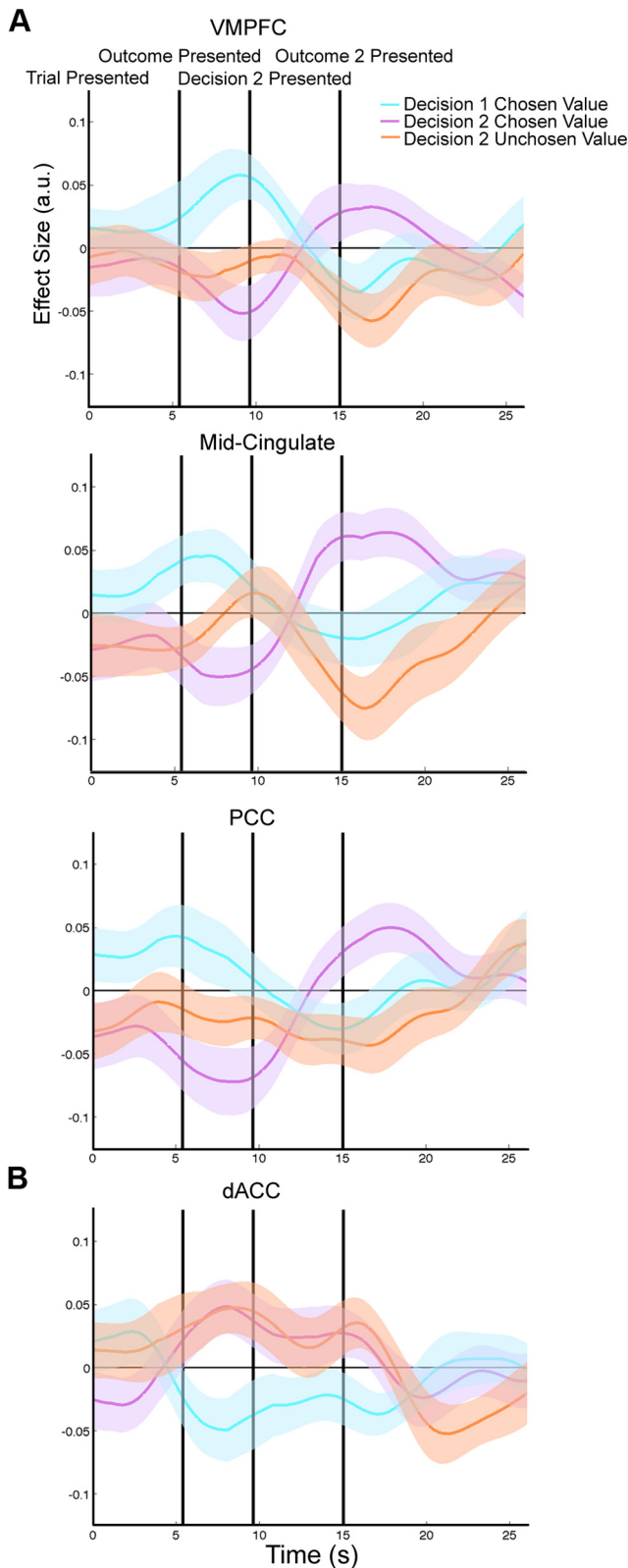
A

B

Figure 5. Value coding during sequential choices. *A*, Time course of the effect size of the chosen value of the first decision, the chosen value at the second decision, and the unchosen value at the second decision from conditions 2 and 3 are plotted across both decisions. Top row, vmPFC; second row, mid-cingulate; third row, PCC. *B*, The same is shown for dACC. Images are displayed according to the same conventions used in Figure 3.

opted the frame of the current choice, whereas dACC adopted the frame of the default option. vmPFC encoded the relative value between the chosen and next best option at each sequential choice, findings consistent with a system that flexibly compares the best two available options in the environment for the current choice. In contrast, we found that dACC encoded the current advantage of switching away from the long-term best option in the environment, only when it was available to choose, consistent with a mechanism for determining when to adapt behavior away from the default position.

A recently developed recurrent cortical circuit model (Wang, 2008; Hunt et al., 2012) predicts that the integrated activity of neuronal pools in a decision-making network will correlate with the chosen relative to the unchosen option's value during binary choice. We found that vmPFC, mid-cingulate, and PCC encoded the chosen relative to next best option value, a signal that may correspond to predictions from such a biophysically inspired decision network, pointing to the possibility that vmPFC and associated regions may constitute at least part of a decision network during goal-based choice. Notably, a comparison between the best two options would also be predicted by a recent computational account of trinary choice (Krajbich and Rangel, 2011). This interpretation is also consistent with impairments in decision-making seen following selective lesions of medial OFC, a subdivision of vmPFC, but not lateral OFC, in monkeys, particularly when the proximity between options' values makes comparison difficult (Noonan et al., 2010; Rudebeck and Murray, 2011a, b), and following damage to vmPFC in humans (Fellows, 2007).

One alternative possibility is that vmPFC, mid-cingulate, and PCC encode the net subjective value of the decision at hand (Tanaka et al., 2004; Daw et al., 2006; Hampton et al., 2006; Kable and Glimcher, 2007; Tom et al., 2007). In the present experiment, the net value signal would constitute the value of the preferred option minus the next best option. Such an effect is consistent with a value signal that incorporates the opportunity cost—the value of the next best option on offer in the environment—and discards inferior options. This finding suggests the possibility that neural activity in these regions may underlie the sensitivity of behavior to opportunity costs emphasized in economics (McConnell and Brue, 2004). That vmPFC activity in the current study decreases with the value of the next best option dovetails with demonstrations that vmPFC activity decreases as other costs associated with the outcome, such as money (Tom et al., 2007; Basten et al., 2010) or delay (Kable and Glimcher, 2007; Prévost et al., 2010) increase (Rangel and Hare, 2010; Rushworth et al., 2011).

An interesting question concerns whether vmPFC activity reflects differences between the chosen and next best alternative or the best two options. In our task, as in most situations, these are not possible to dissociate due to their high correlation and our reliance on a behavioral model to infer subjects' valuation of options. Distinguishing between these possibilities would be an interesting avenue for future research.

We also found that vmPFC activity just before a new trial predicted choice repetition over and above any effect of relative value, chosen or counterfactual outcomes. Notably, such a choice repetition effect is inconsistent with the influential notion that vmPFC damage causes perseveration—the inability to inhibit a previously rewarded choice—and that vmPFC thus promotes behavioral flexibility. It instead suggests that vmPFC is important for promoting previously rewarded choices. It is notable, however, that almost all cases of vmPFC damage also exhibit damage to many neighboring frontal re-
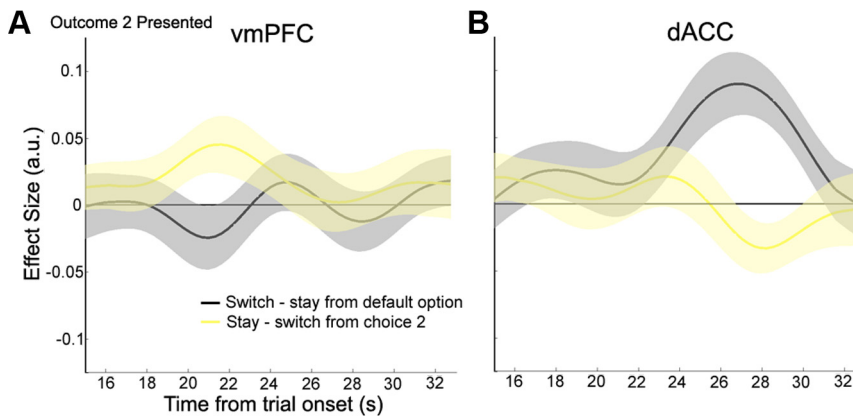
Figure 6. Future choice effects. Time course of the effect of future stays compared with future switches with reference to the second choice in a trial is plotted alongside the effect of future switches compared with future stays with reference to the default option in vmPFC (*A*) and dACC (*B*). Regressors shown were included in the same GLMs, in addition to the chosen and counterfactual outcomes and their associated *EVs* (data not shown for clarity).

showed that the population of dACC neurons encodes the relative evidence for a patch-leaving decision during sequential foraging decisions. The BOLD findings from dACC we have presented are thus consistent with population coding of dACC neurons in monkeys, as well as previous demonstrations that the dACC plays a general role in voluntarily selecting or adapting choices on the basis of reinforcement history (Shima and Tanji, 1998; Walton et al., 2004; Kennerley et al., 2006; Johnston et al., 2007). They also extend a very recent demonstration (Kolling et al., 2012) that dACC activity recorded in humans incorporates key variables of foraging decisions to sequential and multi-alternative choices—two cardinal features of ecological choice. In the present study, we additionally reconcile choice and feedback-related dACC activity and suggest that dACC activity reflects a more general decision variable—the evidence favoring adaptation away from any default position. Finally, our findings complement a recent demonstration that switches from a default option during nonsequential choices between mixed gain/loss gambles may depend on the balance of activity in caudate and anterior insula (Yu et al., 2010).

The interpretation of the choice-related dACC signal as promoting behavioral adaptation brings choice and outcome signals in the dACC into line. During outcome monitoring, single-unit and BOLD activity reflect the change in beliefs or behavior that will ensue from observing an outcome (Rushworth and Behrens, 2008). That the choice-related dACC signal might also reflect the urge for behavioral change opens up the possibility for a combined theory that will explain these two previously incongruous forms of activity.

It is important to note that our characterization of dACC is not necessarily inconsistent with recent demonstrations that activity in DMFC scales with relative unchosen value even in tasks where decisions are not repetitive (Hare et al., 2011). In such situations, where there is no identifiable default position, we hypothesize that dACC and vmPFC reference frames would align with the current choice.

A natural question concerns the precise relationship between vmPFC and dACC-centered systems. We hypothesize that there will be situations in which vmPFC and dACC-centered systems compete, such as when the default position conflicts with the current choice because another option is temporarily more attractive, and situations in which they cooperate, such as when the default position aligns with the current choice. Understanding the contextual dependence of vmPFC–dACC dynamics represents an important step for future research.

In our previous report (Boorman et al., 2011), slightly more dorsal regions of DMFC and posteromedial cortex (PMC) than the dACC and PCC regions identified here also encoded the relative unchosen probability and counterfactual prediction errors. Further inspection of the more dorsal DMFC region revealed that it exhibited rather similar coding to the dACC region reported here, although it was more sensitive to probabilities compared with magnitudes than was

gions and the consequences of damage are tested in a limited number of behavioral situations. When a focused lesion is made to the medial OFC in macaque monkeys, switch rates are indeed elevated, particularly when the proximity between options' values makes comparison difficult (Noonan et al., 2012). Although these monkeys do not exhibit behavioral inflexibility, they are unable to focus their decision space on the two most relevant or valuable options (Noonan et al., 2010) and make choices that are less internally consistent (Rudebeck and Murray, 2011a, b), consistent with the finding that vmPFC activity reflects a comparison between the chosen and next best option's values reported here. Extending previous reports (Gläscher et al., 2009), we were able to show that a choice repetition effect in vmPFC also occurs during the period directly preceding a new trial and is dissociable from relative chosen value effects that occur during decisions.

We found that behavior was biased away from the best option by the allure of the default option and that subjects exhibited a general tendency to prefer the default option (Fig. 2C,D). Our fMRI results show that in the context of sequential multi-alternative choice, dACC adopts a default reference frame. We propose that the dACC encodes a decision variable whose magnitude corresponds to the relative value of adapting behavior from the default position. This interpretation is based on several features of dACC effects in the current study. First, dACC activity that could not be explained by the relative chosen value could be explained by the relative default value. Second, the sign of the relative value signal in dACC [(default V2 + default V3) − default V1] is consistent with a region encoding the evidence supporting adaptation away from, rather than choices of, the default option. Third, dACC only encodes the relative value when the default option is on offer. Fourth, dACC activity just before the onset of a new trial predicts switching away from the default option, but not the recently chosen option, on that new trial, again pointing to a role tied to the default position. As in vmPFC this future choice effect survived the inclusion of relative value, chosen and counterfactual outcomes in the same GLM. Importantly, including these nuisance regressors ensures that the dACC effect at feedback cannot simply be explained by value difference or error-related responses (Fig. 6).

Our interpretation accords well with a recent electrophysiological study in monkeys (Hayden et al., 2011a), which

dACC, suggesting a possible gradient. In contrast, coding in the more dorsal region of PMC and the more ventral PCC were strikingly different, with dorsal PMC reflecting relative unchosen probability and PCC reflecting relative chosen value. This distinction parallels that seen between anterior PFC (aPFC) and vmPFC (Boorman et al., 2009, 2011) and suggests that aPFC, DMFC, and PMC may form one functional network and vmPFC, mid-cingulate, and PCC, a neighboring functional network, which are relatively more concerned with behavioral adaptation and flexible comparative evaluation during explicit choices, respectively.

We have characterized value comparison signals in vmPFC, mid-cingulate, PCC, and dACC during trinary and sequential choice. Our findings suggest that the best and next best options have privileged status during value comparison in vmPFC, mid-cingulate, and PCC. We propose that the effects we observe in vmPFC reflect the comparison of the best two options from a competitive decision-making mechanism during goal-based choice. The findings in dACC, on the other hand, concur with the perspective that dACC encodes a decision variable whose magnitude amounts to the current evidence favoring adaptation away from a default position. Collectively, these findings endorse the view that behavior is guided by multiple controllers and should serve to inform computational models of value-based decision making.

References

Basten U, Biele G, Heekeren HR, Fiebach CJ (2010) How the brain integrates costs and benefits during decision making. Proc Natl Acad Sci U S A 107:21767–21772. CrossRef Medline

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. CrossRef Medline

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

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. CrossRef Medline

Boorman ED, Behrens TE, Rushworth MF (2011) Counterfactual choice and learning in a neural network centered on human lateral frontopolar cortex. PLoS Biol 9:e1001093. CrossRef Medline

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

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

Fellows LK (2007) The role of orbitofrontal cortex in decision making: a component process account. Ann N Y Acad Sci 1121:421–430. CrossRef Medline

Fellows LK (2011) Orbitofrontal contributions to value-based decision making: evidence from humans with frontal lobe damage. Ann NY Acad Sci 1239:51–58. CrossRef

FitzGerald TH, Seymour B, Dolan RJ (2009) The role of human orbitofrontal cortex in value comparison for incommensurable objects. J Neurosci 29:8388–8395. CrossRef Medline

Freidin E, Kacelnik A (2011) Rational choice, context dependence, and the value of information in European starlings (Sturnus vulgaris). Science 334:1000–1002. CrossRef Medline

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

Hampton AN, Bossaerts P, O'Doherty JP (2006) The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. J Neurosci 26:8360–8367. CrossRef Medline

Hare TA, Schultz W, Camerer CF, O'Doherty JP, Rangel A (2011) Transformation of stimulus value signals into motor commands during simple choice. Proc Natl Acad Sci U S A 108:18120–18125. CrossRef Medline

Hayden BY, Pearson JM, Platt ML (2009) Fictive reward signals in the anterior cingulate cortex. Science 324:948–950. CrossRef Medline

Hayden BY, Pearson JM, Platt ML (2011a) Neuronal basis of sequential foraging decisions in a patchy environment. Nat Neurosci 14:933–939. CrossRef Medline

Hayden BY, Heilbronner SR, Pearson JM, Platt ML (2011b) Surprise signals in anterior cingulate cortex: neuronal encoding of unsigned reward prediction errors driving adjustment in behavior. J Neurosci 31:4178–4187. CrossRef Medline

Hunt LT, Kolling N, Soltani A, Woolrich MW, Rushworth MF, Behrens TE (2012) Mechanisms underlying cortical activity during value-guided choice. Nat Neurosci 15:470–476, S1–S3. CrossRef Medline

Jenkinson M, Bannister P, Brady M, Smith S (2002) Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17:825–841. CrossRef Medline

Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM (2012) FSL. Neuroimage 62:782–790. CrossRef Medline

Jocham G, Neumann J, Klein TA, Danielmeier C, Ullsperger M (2009) Adaptive coding of action values in the human rostral cingulate zone. J Neurosci 29:7489–7496. CrossRef Medline

Johnston K, Levin HM, Koval MJ, Everling S (2007) Top-down control-signal dynamics in anterior cingulate and prefrontal cortex neurons following task switching. Neuron 53:453–462. CrossRef Medline

Kable JW, Glimcher PW (2007) The neural correlates of subjective value during intertemporal choice. Nat Neurosci 10:1625–1633. CrossRef Medline

Kable JW, Glimcher PW (2009) The neurobiology of decision: consensus and controversy. Neuron 63:733–745. CrossRef Medline

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

Kolling N, Behrens TE, Mars RB, Rushworth MF (2012) Neural mechanisms of foraging. Science 336:95–98. CrossRef Medline

Krajbich I, Rangel A (2011) Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. Proc Natl Acad Sci U S A 108:13852–13857. CrossRef Medline

Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI (2009) Circular analysis in systems neuroscience: the dangers of double dipping. Nat Neurosci 12:535–540. CrossRef Medline

Lim SL, O'Doherty JP, Rangel A (2011) The decision value computations in the vmPFC and striatum use a relative value code that is guided by visual attention. J Neurosci 31:13214–13223. CrossRef Medline

McConnell CR, Brue SL (2004) Microeconomics: principles, problems, and policies, Ed 16. McGraw-Hill Professional.

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. CrossRef Medline

Noonan MP, Kolling N, Walton ME, Rushworth MF (2012) Re-evaluating the role of the orbitofrontal cortex in reward and reinforcement. Eur J Neurosci 35:997–1010. CrossRef Medline

Parvizi J, Van Hoesen GW, Buckwalter J, Damasio A (2006) Neural connections of the posteromedial cortex in the macaque. Proc Natl Acad Sci U S A 103:1563–1568. CrossRef Medline

Pearson JM, Hayden BY, Raghavachari S, Platt ML (2009) Neurons in posterior cingulate cortex signal exploratory decisions in a dynamic multioption choice task. Curr Biol 19:1532–1537. CrossRef Medline

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

Platt ML, Huettel SA (2008) Risky business: the neuroeconomics of decision making under uncertainty. Nat Neurosci 11:398–403. CrossRef Medline

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

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

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

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

Rudebeck PH, Murray EA (2011b) Balkanizing the primate orbitofrontal cortex: distinct subregions for comparing and contrasting values. Ann N Y Acad Sci 1239:1–13. CrossRef Medline

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

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

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

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

Sutton RS, Barto AG (1998) Reinforcement learning: an introduction. Cambridge, MA: MIT.

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. CrossRef Medline

Tom SM, Fox CR, Trepel C, Poldrack RA (2007) The neural basis of loss aversion in decision-making under risk. Science 315:515–518. CrossRef Medline

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

Wang XJ (2008) Decision making in recurrent neuronal circuits. Neuron 60:215–234. CrossRef Medline

Wessel JR, Danielmeier C, Morton JB, Ullsperger M (2012) Surprise and error: common neuronal architecture for the processing of errors and novelty. J Neurosci 32:7528–7537. CrossRef Medline

Wunderlich K, Rangel A, O'Doherty JP (2009) Neural computations underlying action-based decision making in the human brain. Proc Natl Acad Sci U S A 106:17199–17204. CrossRef Medline

Yu R, Mobbs D, Seymour B, Calder AJ (2010) Insula and striatum mediate the default bias. J Neurosci 30:14702–14707. CrossRef Medline