Neuron
Article

How Green Is the Grass on the Other Side? Frontopolar Cortex and the Evidence in Favor of Alternative Courses of Action

Erie D. Boorman,[1,2,3,*] Timothy E.J. Behrens,[1,2,3] Mark W. Woolrich,[1] and Matthew F.S. Rushworth[1,2]
[1]Centre for Functional MRI of the Brain, University of Oxford, Oxford 0X3 9DU, UK
[2]Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, UK
[3]These authors contributed equally to this work
*Correspondence: erie.boorman@psy.ox.ac.uk
DOI 10.1016/j.neuron.2009.05.014

SUMMARY

Behavioral flexibility is the hallmark of goal-directed behavior. Whereas a great deal is known about the neural substrates of behavioral adjustment when it is explicitly cued by features of the external environment, little is known about how we adapt our behavior when such changes are made on the basis of uncertain evidence. Using a Bayesian reinforcement-learning model and fMRI, we show that frontopolar cortex (FPC) tracks the relative advantage in favor of switching to a foregone alternative when choices are made voluntarily. Changes in FPC functional connectivity occur when subjects finally decide to switch to the alternative behavior. Moreover, interindividual variation in the FPC signal predicts interindividual differences in effectively adapting behavior. By contrast, ventromedial prefrontal cortex (vmPFC) encodes the relative value of the current decision. Collectively, these findings reveal complementary prefrontal computations essential for promoting short- and long-term behavioral flexibility.

INTRODUCTION

The ability to adapt behavior to suit the current environment is central to survival in a changing world. Much experimental work has been devoted to unraveling the neural mechanisms that underpin adaptive behavior when such changes are explicitly instructed either by external stimuli (Grol et al., 2006; Passingham, 2008; Schluter et al., 1998; Toni et al., 2001; Wise et al., 1997) or previously learned response strategies (Genovesio et al., 2005; Genovesio and Wise, 2008; Miller and Buschman, 2008; Wise, 2008). However, in everyday life, organisms must often determine what to do in the absence of such explicit cues. In such contexts, adaptive behavior might depend upon tracking the evidence in favor of current and alternative courses of action.

It is now well established that frontopolar cortex (FPC) is active when human subjects switch between one complex behavioral task and another, particularly when it is necessary to hold information about one task in working memory when switching to the alternative (Braver et al., 2003; Koechlin et al., 1999; Koechlin and Hyafil, 2007; Koechlin and Summerfield, 2007; Ramnani and Owen, 2004). Typically, such experiments provide participants with explicit instructions about when to switch from one task to another. Yet whether or how the FPC contributes to the control of behavior when participants freely select between tasks or even between simple choices remains unclear.

Furthermore, it is established that other brain areas, particularly in the parietal cortex, are active during cued behavioral switching (Braver et al., 2003; Glascher et al., 2009; Jubault et al., 2007; Rushworth et al., 2001). However, the distinct contributions of, and interplay between, parietal and frontopolar cortex during behavioral switching have proven elusive (Daw et al., 2006; Glascher et al., 2009).

We addressed these issues by examining human FPC activity while subjects freely selected between two actions during a simple decision-making task. Crucially, no explicit instructions signaling behavioral change were provided. Because our focus was on how subjects adapt behavior based on uncertain evidence, rather than on other aspects of task control, subjects were not required to switch between complex behavioral tasks but simply between one of two possible actions on the basis of the expected values of reward associated with the actions. The expected value of each action was in turn a function of the probability that it would yield rewards if chosen (which the subject must estimate from recent outcomes) and the magnitude of reward associated with the action (which was indicated on the task screen on each trial and changed unpredictably from trial to trial). A Bayesian model was used to infer subject estimates of the outcome probabilities. By carefully decorrelating the critical parameters (probabilities and expected values of chosen and unchosen options) across trials, we were able to regress these parameters against the blood oxygenation level-dependent (BOLD) signal acquired during functional magnetic resonance imaging (fMRI). By further dissociating different stages of the decision-making process within a trial in time, we could assess both the short-term representation of variables crucial to the current decision (option *expected values*) and the extended representation of variables important for future decisions (option reward *probabilities*).

Here, we show that the FPC continually tracked the evidence accumulated in favor of switching to the alternative course of action. Moreover, immediately before a behavioral switch actually occurred, the FPC entered into a distinct pattern of functional connectivity with the parietal cortex, suggesting that when the

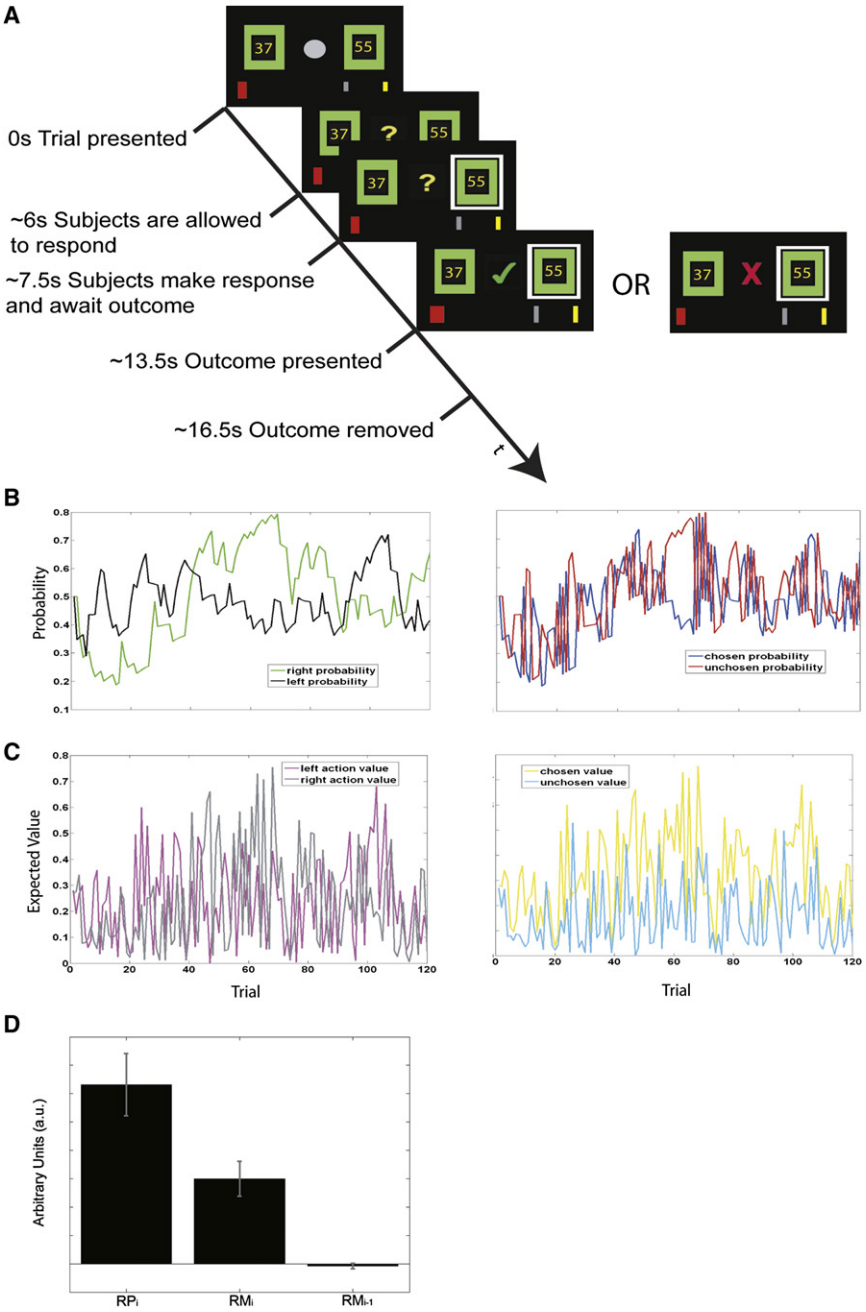

Figure 1. Two-Armed Bandit Task and Representative Probability and Value Estimates

(A) Representative trial and timeline. Subjects performed a two-armed bandit task in which they chose between left and right options on the basis of past outcomes and the reward magnitudes (yellow numbers). If the subject's choice was rewarded, the red prize bar at the bottom of the screen moved toward the silver and gold targets at the bottom right of the screen in proportion to the chosen reward magnitude. If the red prize bar reached the silver target, subjects earned £20; if it reached the yellow target, they earned £30. In this example, the subject's choice was rewarded, so the red prize bar moved toward the targets by a distance in proportion to 55 points. If the subject's choice was not rewarded, the red prize bar remained the same (also shown). Importantly, no feedback was given on the unselected option. Times displayed on the timeline refer to mean onset times for each event across trials and subjects.

(B) Left panel: probabilities associated with the left and right options being rewarded are shown in black and green, respectively, for one subject. Right panel: probabilities associated with the chosen and unchosen options being rewarded are shown for the same subject in blue and red, respectively. These reward probabilities were estimated from the observed outcomes of each individual subject's choices by a Bayesian reinforcement learning algorithm (see Supplemental Data).

(C) Left panel: expected values associated with the left and right options are shown in purple and silver, respectively. Right panel: expected values associated with the chosen and unchosen options are shown in yellow and cyan, respectively, for the same subject.

(D) Regression coefficients relating to the difference between model-estimated reward probabilities from the current trial (RPi), the difference between random reward magnitudes from the current trial (RMi), and the difference between random reward magnitudes from the previous trial (RMi-1). Error bars represent ±SEM.

FPC has recruited sufficient evidence to support a behavioral switch, it engages the parietal cortex to implement the switch. These data isolate a distinct computation performed by the FPC in the service of behavioral flexibility, even in situations in which choices are made voluntarily, and suggest a new account of frontal lobe activity in decision making.

RESULTS

Experimental Design

Participants performed a decision-making task in which they repeatedly chose between left and right options by making left- or right-hand button presses (Figure 1A). The task was modeled on one previously used by Behrens and colleagues (Behrens et al., 2007). By changing a key task parameter—the statistical independence between options—it was possible to pursue a completely distinct experimental goal. Instead of focusing on the rate of learning (as in the previous study), this feature of the experimental design enabled us to examine the neural correlates of alternative courses of action. On each trial, random integers between 1 and 100 were displayed, one in the center of the rectangle denoting each option, that indicated the size of reward available for taking that option. Subjects were informed that because these reward *magnitudes* were

determined randomly on each trial, it was neither beneficial nor possible to track them across trials. However, participants were not explicitly cued about the *probability* that each option would be rewarded if chosen. Instead, subjects were told that these reward probabilities depended only on the recent outcome history and could therefore be tracked across trials. To simulate a changing environment, these underlying reward probabilities drifted from trial to trial during the course of the experiment (Figure 1B).

We implemented an optimal Bayesian learner (Behrens et al., 2007) to model subject estimates of the probabilities of reward associated with the chosen and unchosen options given the history of recent choice outcomes (i.e., rewarded or unrewarded) (see Supplemental Data). To confirm that the Bayesian model captured human behavior in our task, we used logistic regression to determine the extent to which subject choices were influenced by optimally tracked outcome probabilities and by the reward magnitudes that changed randomly on each trial. This analysis confirmed that subject choices were significantly influenced by the difference between the model-derived reward probabilities on the current trial [$t(17) = 5.73$, $p < 0.0001$] and the difference in reward magnitudes on the *current* trial [$t(17) = 4.9$, $p < 0.0002$] (Figure 1D). The relative reward magnitude on the current trial should influence choice, but the reward magnitudes on previous trials should theoretically not have any effect because they changed randomly from trial to trial. Consistent with the model, there was no significant effect of the difference in reward magnitudes from the previous trial on subject choices [$t(17) - 0.85$, $p > 0.20$] (Figure 1D), indicating that subject behavior did not reflect any attempt to track the random reward magnitudes.

Crucially, the Bayesian learner enabled us to select a reward schedule that decorrelated the reward probabilities associated with left and right options (subsequently referred to as *action probabilities*). Because we could not know what choices our participants would ultimately make, this step was critical, as it increased the likelihood that chosen and unchosen action probabilities would also be decorrelated (Figure 1B). As predicted, there was indeed little correlation between the model-estimated left and right action probabilities (mean signed r^2 across subjects = −0.08, SD = 0.09) or between chosen and unchosen action probabilities (mean signed $r^2 = -0.10$, SD = 0.09). Similarly, the expected value (reward probability × reward magnitude) associated with left and right options (subsequently referred to as *action values*) were decorrelated, which also increased the likelihood that chosen and unchosen action values would be decorrelated (Figure 1C). Accordingly, there was limited correlation between the model-estimated left and right action values (mean signed $r^2 = 0.02$, SD = 0.01) and between chosen and unchosen action values (mean signed $r^2 = 0.11$, SD = 0.09) (Figure S1; see Table S1 for individual subject correlations). It is important to emphasize that while *action values* (which are a function of both reward probability and magnitude) were relevant for making a decision on any given trial, the random trial-by-trial fluctuations in reward magnitude meant that longer-term behavioral strategies depended only on *action probabilities*.

The FPC Tracks the Relative Advantage Associated with the Alternative Course of Action

If it is true that a brain region encodes the long-term evidence accumulated in support of a switch in behavior, then it should encode the *relative* evidence in favor of the alternative action. To search for such a neural signal, we tested for regions throughout the whole brain where activity correlated with the log-ratio between the model-estimated unchosen and chosen action probabilities or the *relative unchosen probability*—the relevant parameter to track through and across trials to inform switches to the alternative course of action in our task. This analysis, which used standard corrections for multiple comparisons across the whole brain (Smith et al., 2004), revealed bilateral regions of lateral FPC whose activity tracked the relative unchosen probability ($Z = 3.89$, MNI $x = -34$, $y = 56$, $z = -8$, $Z = 3.34$, $x = 36$, $y = 54$, $z = 0$) (Figure 2A; Table S2). Region of interest (ROI) analyses demonstrated that the FPC signal correlates with the relative unchosen probability throughout the length of the trial, with peaks occurring at two time points: during the decision-making period and in the intertrial interval (ITI) (Figure 2B). Notably, taking the difference between probabilities, rather than the log-ratio, revealed virtually identical activations and time series of correlation (Figure S2). We therefore do not make a strong claim about exactly what function of probability is coded in FPC. However, we do have strong evidence that this function increases monotonically with relative unchosen probability (Figure 2C).

The correlation between FPC signal and relative unchosen probability is not driven by the current behavioral choice; the effect survived the inclusion of binary indicator confound regressors that coded for exploratory versus exploitative decisions, switching versus staying, and left versus right choices (Figure S3). Furthermore, separating the relative unchosen probability into its component parts revealed a positive correlation with the unchosen action probability and a negative correlation with the chosen action probability (Figure 2B), demonstrating that the FPC does indeed encode a *relative* signal. Although negative fMRI findings must be interpreted with caution, we did not observe any region whose activity correlated significantly with either the chosen or unchosen action probability in isolation. In addition, the FPC was not sensitive to the relative difference between unchosen and chosen reward magnitudes [$t(17) < 0.4$, $p > 0.4$]. Intriguingly, the FPC signal scales with the relative unchosen probability on both stay [$t(17) = 2.32$, $p < 0.05$] and switch trials [$t(17) = 2.65$, $p < 0.01$] (Figure 2C), indicating that even before subjects have responded, the FPC already encodes the evidence in favor of the option that will soon become unchosen after subjects switch. In other words, the FPC changes its frame of reference in terms of which action is chosen even before subjects adapt their behavior. The second peak in the FPC signal, which occurs after the completion of a decision on one trial, contains the evidence in favor of a behavioral switch on the following trial.

Within- and Between-Subject Variation in the FPC Signal Predicts Switching to the Alternative Action

The FPC codes the relative evidence in favor of the alternative behavior even in the period after a decision is made and its

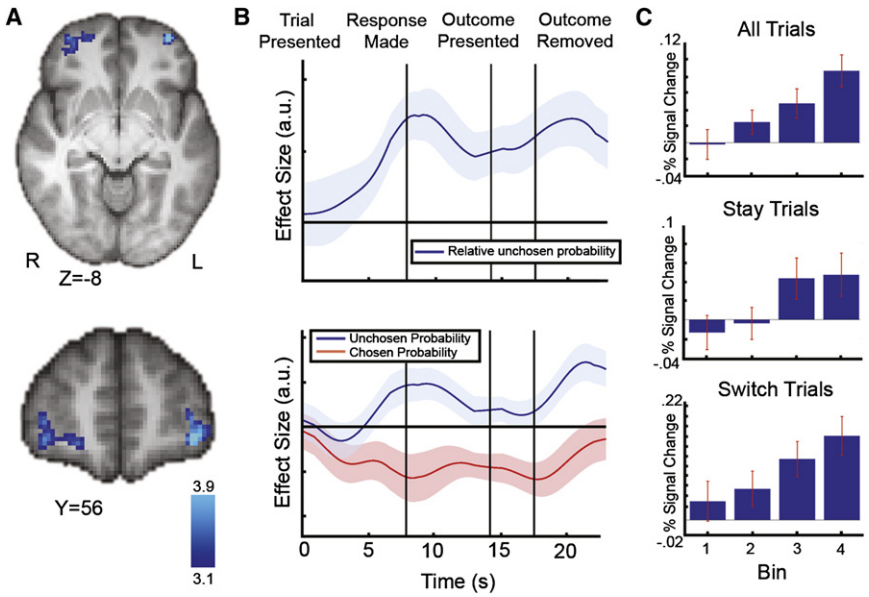

Figure 2. The Frontopolar Cortex Tracks the Relative Unchosen Probability

(A) Axial and coronal slices through z-statistic maps relating to the relative unchosen probability [log(unchosen action probability/chosen action probability)]. Z-statistic maps are corrected for multiple comparisons across the whole-brain by means of cluster-based Gaussian random-field correction at $p < 0.05$. Maps are displayed according to radiological convention. Colorbar indicates Z-score.

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

(C) The bar plots show the mean BOLD percent signal change in the FPC, with trials binned by the relative unchosen probability into quartiles such that each bin has an equivalent number of trials on all trials (top), stay trials (middle), and switch trials (bottom). Error bars represent ± SEM.

outcome has been presented. Such a signal may be viewed as a neural representation of the evidence in favor of switching behavior at the following trial. This interpretation makes strong and testable predictions about the effect of this signal on behavior. First, within subjects, the FPC signal during the ITI might predict switching on the following trial. Second, subjects in whom this evidence is poorly represented should not switch to the alternative behavior at appropriate times.

We tested these predictions in two complementary ways. First, we assessed whether greater FPC activity at any point of the current trial predicted switching on the following trial. Remarkably, even though reward magnitudes were not yet revealed to subjects, the FPC signal between trials correlated with switching on the following trial [Figure 3A; $t(17) = 2.1$, $p < 0.05$]. It is likely that this correlation would have been stronger had reward probabilities been the only source of information guiding subject decisions on the upcoming trial. Second, we assessed whether variation across subjects in the FPC signal during the ITI predicted switching to the alternative when it was the more valuable option. Because reward magnitudes might render the more probable option less valuable on a given trial, switching to the previously unchosen option would only be advantageous when it also has the higher expected value. We reasoned that if the FPC signal between trials is guiding switches, then variation in this signal across subjects should predict switching when it is advantageous on the following trial. As predicted, subjects with a greater effect of the relative unchosen probability in the FPC during the ITI were more likely to switch to the more valuable option (Figure 3B).

We further reasoned that subjects with a stronger representation of relative unchosen probability in the FPC would be more likely to behave as if they retained this value accurately. We therefore sought to explain behavior in each subject by two competing models. The first (OPTIMAL) leaked information about the unchosen probability at an optimal rate. In our task, the subject does not receive feedback about rewards that would have been available on the unchosen option. Therefore, optimal behavior is for the estimate of the unchosen option's probability to decay to chance levels (0.5) at a rate determined by the volatility of the environment (Supplemental Data) (Behrens et al., 2007). In the second model (LEAKY), estimates of the unchosen probability decay at a faster rate (Supplemental Data). The extent to which a subject's behavior is better predicted by OPTIMAL than LEAKY is a measure of how accurately the subject retains information about the unchosen option across trials. We therefore used a logistic regression analysis in which these two models would compete to describe subject behavior. We computed the ratio of the regression coefficients from the logistic regression to give β; the extent to which a subject's behavior is better fit by OPTIMAL than by LEAKY. Individuals with a greater effect of unchosen probability in the FPC during the ITI had higher β values (Figure 3C). Taken together, these within- and between-subject correlations suggest that the FPC not only maintains the evidence in favor of switching but also might play a role in implementing switches to foregone options.

Implementation of the Behavioral Switch: Functional Connectivity between FPC and Parietal and Premotor Regions

In addition to the FPC, our whole-brain analysis of the relative unchosen probability identified bilateral regions midway along the intraparietal sulcus, bordering the inferior parietal lobule (mid-IPS) ($Z = 3.89$, MNI $x = 50$, $y = -46$, $z = 46$, $Z = 3.69$, $x = -32$, $y = -60$, $z = 52$) (Figure 4A). A similar mid-IPS region has previously been reported to be coactive with the FPC during exploratory behavior (Daw et al., 2006), although the difference in the

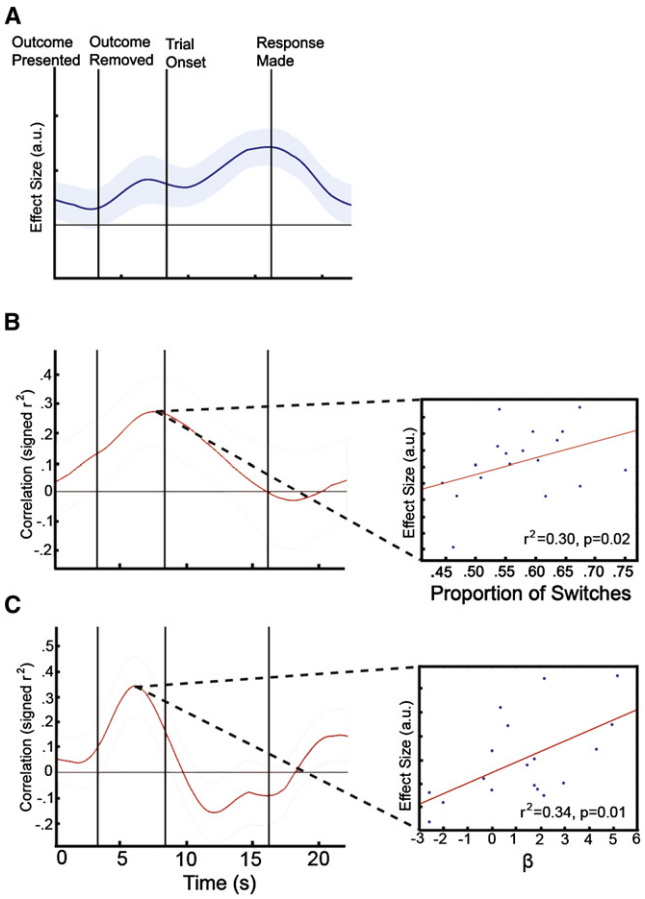

Figure 3. Frontopolor Cortex Activity during the ITI Predicts Within- and Between-Subject Variability in Behavior

(A) FPC effect of switching on the following trial plotted from the time of outcome presentation on the current trial to the time of outcome presentation on the following trial. There is a significant effect of switching on the following trial during the ITI. The time course plot is displayed according to the same conventions used in Figure 2.

(B and C) Left column: time series of between-subject correlation (signed r^2) between the effect size in the FPC and the proportion of trials on which subjects switched to the better option (B) and the ratio of the fit between models incorporating optimally retained or quickly forgotten estimates of an unchosen action's probability, β (C). The correlation is shown at every time point in the trial, plotted from the time of outcome presentation on the current trial to the time of outcome presentation on the following trial. Right column: scatter plots were produced by fitting a hemodynamic model in each subject to the time course of the effect in the FPC. For this model, a canonical hemodynamic response function was time locked to the start of the intertrial interval. The corresponding effect size (i.e., regression coefficient) for the relative unchosen probability in the FPC during the ITI for each subject was then plotted against the proportion of trials when a subject switched as a fraction of the trials when it would have been advantageous to switch (A) and β (B).

functions of the two regions has been difficult to establish. Dividing trials into switch and stay trials, however, revealed that while the FPC encoded the relative unchosen probability monotonically on all trials (Figure 2C), the mid-IPS correlated with the relative unchosen probability on switch trials only (Figure 4B) (repeated-measures ANOVA for interaction between regression coefficients relating to the relative unchosen probability on switch and stay trials: $F_{1,17} = 4.6$, $p < 0.05$; relative unchosen probability on switch trials only: $t(17) = 3.2$, $p < 0.005$; relative unchosen probability on stay trials only: $t(17) < .5$, $p > 0.4$). Given this region's proposed role in action updating and behavioral switching (Glascher et al., 2009; Jubault et al., 2007; Rushworth et al., 2001), this observation raised the possibility that the FPC might convey the amount of evidence favoring switches to the mid-IPS and related regions responsible for implementing a switch in behavior.

We reasoned that since the FPC accumulates the evidence in favor of switching across all trial types and predicts switches within and between participants, it might transmit this evidence to other brain regions when there is sufficient evidence to support a switch in behavior before switches actually occur. To test this hypothesis, we performed a functional connectivity analysis designed to ask the following question: where in the brain is there a change in interaction with the FPC and the relative unchosen probability, specifically before switches are made? Importantly, this is the specific temporal window during which we would expect brain regions to prepare to implement switches. The right parietal cortex and interconnected right ventral premotor cortex (PMv) (Matelli et al., 1986; Rushworth et al., 2006) showed enhanced functional connectivity with the right FPC that was dependent on the estimated size of the relative unchosen probability, specifically before switches in behavior took place ($Z = 3.41$, MNI $x = 48$, $y = -42$, $z = 44$) (Figure 4C). Although no regions survived thresholding when we investigated functional connectivity with the left FPC, the left parietal cortex and PMv survived the reduced threshold of $Z > 2.8$, $p < 0.005$, uncorrected (Figure S4). Importantly, because the psychological regressor (the relative unchosen probability preceding switches) and the physiological regressor (the FPC timeseries) were included in the general linear model for the functional connectivity analysis, these activations were not due to the psychological or physiological regressors alone but to their interaction. Notably, the parietal region exhibiting enhanced functional connectivity overlapped with the mid-IPS region found to code for the relative unchosen probability on switch trials (Figures 4A and 4C). Consistent with the proposal that this region plays a role in behavioral switching (Glascher et al., 2009; Jubault et al., 2007; Rushworth et al., 2001), ROI analyses also revealed a main effect of switching during the decision-making phase in the mid-IPS in the current study [$t(17) = 4.0$, $p < 0.001$]. Although it is not possible to determine whether the FPC or IPS is driving the interaction from the functional connectivity analysis alone, these analyses, in conjunction with the findings that the FPC predicts switching both within and between subjects and encodes the evidence advocating a switch on both stay and switch trials, while the IPS only encodes this information before switches, strongly suggest that the FPC both represents the evidence supporting selection of the unchosen option and transmits the information to the mid-IPS and PMv in order to implement switches.

The Evidence in Favor of the Current Decision

Because of their relatively slow rate of change, reward probabilities were relevant for tracking the long-term evidence accumulated in favor of different choice strategies. However, on any

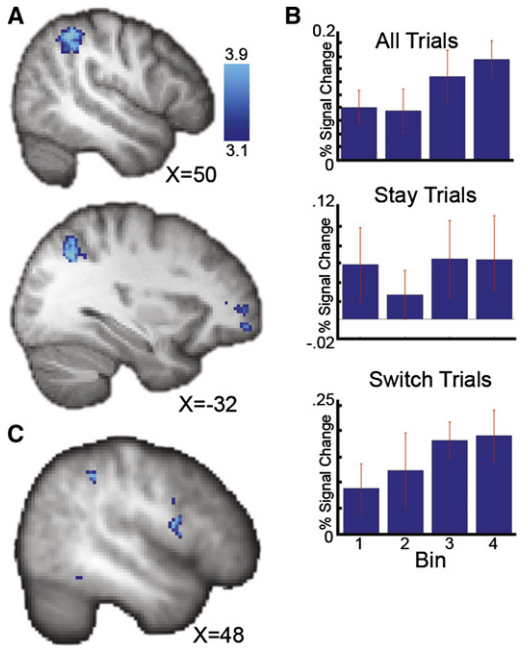

Figure 4. Functional Connectivity between FPC and Mid-IPS and Premotor Regions

(A) Sagittal slices through z-statistic maps relating to the relative unchosen probability.

(B) Bar plots showing mean mid-IPS BOLD responses, binned by the relative unchosen probability across all trials (top), stay trials (middle), and switch trials (bottom). The mid-IPS only correlates linearly with the relative unchosen probability on switch trials.

Images in (A) and (B) are displayed according to the same conventions used in Figure 2.

(C) Functional connectivity analysis. Sagittal slice through z-statistic map showing regions that interact with the right FPC signal and the relative unchosen probability before switches are made. Z-statistic map in (C) is displayed with a voxel-wise threshold of $Z > 3.1$, $p < 0.001$ uncorrected.

given trial, decisions should be determined not only by the reward probability but also by the reward magnitude—that is, choices should be a function of the (subjective) action values (Supplemental Data). Previous fMRI data have implicated a region of the vmPFC in coding for the expected value of the chosen action or for related variables (Behrens et al., 2008; Daw et al., 2006; Hampton et al., 2006; Hare et al., 2008; Kable and Glimcher, 2007; Knutson et al., 2005; Plassmann et al., 2007; Tanaka et al., 2004). These studies have not, however, investigated how this signal might depend upon alternative, competing actions. If it is true that this signal encodes a decision between maintaining and adapting behavior, then it should encode the evidence in favor of the chosen action. A second general linear model (GLM) was therefore constructed to search for brain regions whose activity correlated with the difference between the chosen and unchosen subjective expected values, or the *relative chosen value*. The only region across the entire brain to survive thresholding was located in the vmPFC (max $Z = 3.75$, at MNI $x = -6$, $y = 48$, $z = -8$) (Figure 5A) (Table S2). ROI analysis revealed that the vmPFC encoded the relative chosen value during the decision period and not at any other time point in the trial (Figure 5B). Furthermore, separating the relative chosen value into its component parts revealed a positive correlation with chosen action value and a negative correlation with unchosen action value, demonstrating that, like the FPC, the vmPFC encodes a *relative* signal (Figure 5B). However, while the FPC tracked the relative unchosen *probability*, the vmPFC signal correlated best with the relative chosen *value*. As with the probability-based analysis, no region's activity correlated significantly with either the chosen or unchosen action value in isolation. Finally, by including decorrelated measures of both reward probability and reward magnitude in our experimental design (Figure S1), we were able to demonstrate that the vmPFC encodes both metrics monotonically during the time of the decision [effect of reward probability: $t(17) = 2.23$, $p < 0.05$; effect of reward magnitude: $t(17) = 2.48$, $p < 0.05$] (Figure 5C).

DISCUSSION

This study reveals the specific computations performed by human FPC in service of behavioral flexibility during voluntary choice. We have shown that the FPC is not just active when a change in behavior occurs but that it continually and monotonically tracks the long-term evidence accrued to support a switch in behavior both during decisions and in the intervals between trials (Figure 2). Moreover, rather than passively monitoring these parameters without consequence for behavior, ROI analyses indicated that the greater the effect size in the FPC between trials, the more likely subjects were to appropriately switch their behavior on the next trial (Figure 3). Finally, changes in functional connectivity between the FPC and parietal and premotor regions occur immediately before switches to the alternative actually take place (Figures 4 and S4). These findings demonstrate a pivotal role for the FPC in promoting behavioral flexibility, even when decisions are made voluntarily. Furthermore, that vmPFC encodes the relative value of the current decision suggests that vmPFC and FPC may perform key complementary roles during decision making.

There is an emerging view that complex behavior can be understood with respect to hierarchical organized systems for action selection (Bunge et al., 2005; Koechlin and Jubault, 2006; Koechlin et al., 2003; Koechlin and Summerfield, 2007). According to this perspective, functions of lateral frontal cortex can be described in relation to a rostral-caudal functional axis with increasing complexity in increasingly rostral brain regions. Along this axis caudal regions select actions on the basis of learned associations with simple cues, intermediate areas are recruited when the significance of a cue is contextually dependent, while more rostral regions subserve additional control related to cues provided from a previous temporal episode. In other words, these more rostral regions are employed when the rules for action selection change with task context or episode. By contrast, it has been argued that FPC maintains representations based on past experience but relating to potential future or "pending" behavioral states (Koechlin, 2008; Koechlin and Hyafil, 2007). Although FPC activity has previously been reported when cue, context, and episodic contingencies are *insufficient* to guide behavior, the present study demonstrates a role for FPC even when these contingencies are *absent* and the behavioral

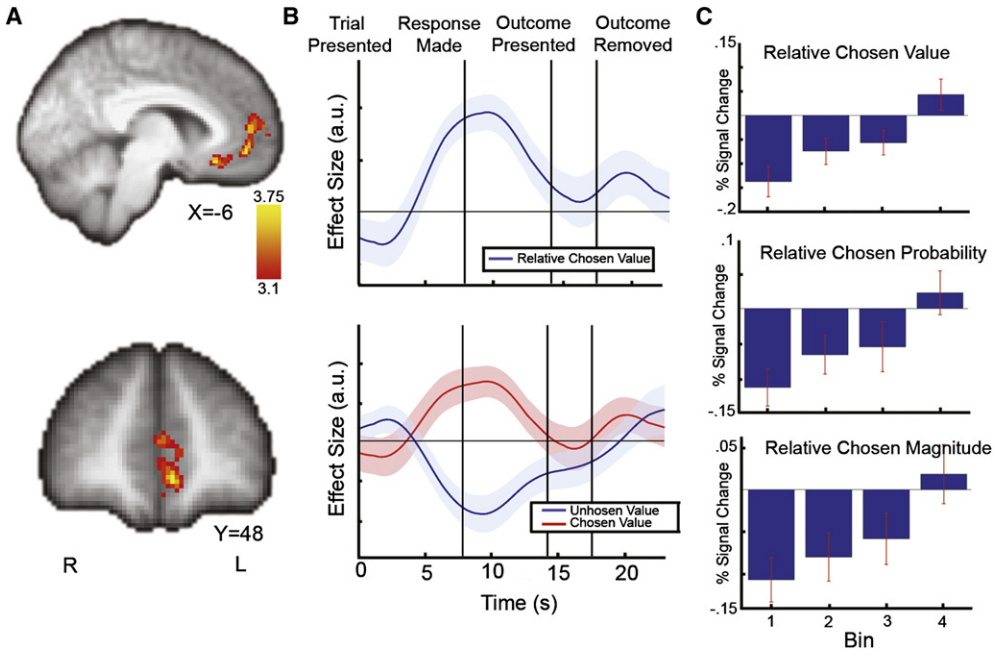

Figure 5. The Ventromedial PFC Encodes the Relative Chosen Value during the Decision-Making Phase

(A) Sagittal and coronal slices through z-statistic maps relating to the relative chosen value (chosen – unchosen expected value).

(B) Top panel: time course for the effect size of the relative chosen value in the vmPFC is shown throughout the duration of the trial. Bottom panel: the same time course is shown with the signal decomposed into chosen and unchosen action values. There is a positive correlation with chosen value and a negative correlation with unchosen value during the decision-making phase.

(C) Bar plots showing mean vmPFC BOLD responses, binned by relative chosen value (chosen – unchosen expected value) (top), relative chosen probability (chosen – unchosen probability) (middle), and relative chosen magnitude (chosen – unchosen magnitude) (bottom). Images are displayed according to the same conventions used in Figure 2.

task involves simple responses but where response selection is based on uncertain evidence. Despite the absence of these contingencies, the evidence in favor of a pending behavioral state is estimated on the basis of past experience by the FPC in the current task. The FPC thus encodes information pertaining to potential but still temporally distant episodes of behavior (Koechlin, 2008; Koechlin and Hyafil, 2007).

It should be noted that, in our experimental task, there was no advantage to tracking reward magnitudes from one trial to the next. Consistent with this feature of the task design, subject decisions were not influenced by reward magnitudes from the past trial (Figure 1D). The only information guiding expectations about *future* behavior were therefore the reward probabilities associated with the options. The FPC signal uncovered in the current study may thus be conceived of as encoding the relative expectation of reward associated with the unchosen compared to the chosen option at the upcoming trial. Such an evidence-based signal should theoretically guide switching on the forthcoming trial, a hypothesis supported by correlations between both intra- and intersubject variability in FPC activity and switching behavior (Figures 3A and 3B).

The FPC has been implicated in multitasking (Braver et al., 2003; Burgess, 2000; Burgess et al., 2000; Koechlin et al., 1999, 2000), prospective memory (Burgess et al., 2003; Sakai and Passingham, 2006; Sakai et al., 2002), and relational reasoning (Bunge et al., 2005; Prabhakaran et al., 2000; Wendelken et al., 2008). These functions all require the maintenance of a ''pending'' task set or of information while another behavior is ongoing (Koechlin et al., 1999; Koechlin and Hyafil, 2007; Ramnani and Owen, 2004). However, these previous studies all provided explicit cues or rules instructing subjects when to switch to the pending representation. The current findings show that the FPC plays a general role in representing a pending or alternative potential course of action in order to determine when to switch to that behavior, independently of whether such switches are externally instructed or voluntary, and independently of the complexity of the pending behavior. Moreover, they demonstrate that FPC is not merely more active at the point of switching but, first, that it continually tracks the evidence favoring a switch in behavior and, second, that it enters into a distinct pattern of functional connectivity with parietal and premotor regions as the switch in behavior actually occurs. Previous studies have reported changes in functional interactions between FPC and other brain regions when specific instructions to change the way tasks are performed are given to subjects (Sakai and Passingham, 2006; Sakai et al., 2002), but the present results suggest that an equally important function of the FPC is to accumulate evidence for when a switch will be needed and then to functionally interact with other brain regions even when explicit instructions to switch are not provided.

It is noteworthy that medial area 10 in the macaque monkey does not possess efferent projections to parietal cortex (Petrides

and Pandya, 2007). However, after injections were made in area 10, terminal label was identified just posterior to the arcuate sulcus in PMv, a region which possesses dense projections to the parietal region identified by our functional connectivity analysis (Matelli et al., 1986; Rushworth et al., 2006). It is therefore plausible that PMv, which was also activated in our functional connectivity analysis, served as a relay region to transmit the evidence favoring switches from FPC to parietal cortex. Another route that could potentially mediate the influence of the FPC on the parietal cortex might run via the dorsolateral prefrontal cortex (Petrides and Pandya, 2007; Cavada and Goldman-Rakic, 1989). We note that some changes in dorsolateral prefrontal activity were identified by the PPI analysis, but they remained below the threshold for statistical significance. Given that there is some evidence that the FPC and mid-IPS region are disproportionally expanded in the human brain (Husain and Nachev, 2007; Semendeferi et al., 2001), it is also possible that other routes between these regions may be found to exist in the human.

In addition, the FPC has recently been implicated in exploration (Burgess et al., 2007; Daw et al., 2006; Yoshida and Ishii, 2006). In particular, one recent study identified a network of regions including similar FPC and mid-IPS regions when subjects forewent the most valuable option in order to "explore" alternatives in a changing environment (Daw et al., 2006). While such activation patterns have been interpreted as indicative of a neural mechanism for efficient sampling of a changing environment, it is nevertheless the case that it is on these trials that foregone options have particularly high value. The present results begin to explain the nature of FPC's contribution to exploration and how it is distinguished from that of mid-IPS. Specifically, the present study demonstrated that FPC monotonically encodes the evidence in favor of the alternative behavior across *all* trial types, whether the decision is to switch or stay, to explore or exploit. Moreover, the FPC enters into an altered pattern of interaction with the mid-IPS that is dependent on the estimated size of the relative unchosen probability only on switch trials. Unlike the FPC, the mid-IPS only encodes the evidence in favor of switching immediately before switches actually occur. Our data therefore suggest that FPC's and mid-IPS's proposed roles in exploration might be best understood in terms of continually tracking the evidence in favor of switching to an alternative behavior and intermittently implementing switches to that alternative, respectively.

Our experimental design enabled us to dissociate long-term signals accumulated across trials from short-term signals that should reflect decisions on the current trial. Whereas the FPC tracked the strategic advantage of switching to a foregone alternative across trials, the vmPFC encoded the relative advantage of the current decision on a given trial. Activity in vmPFC has previously been shown to correlate with parameters associated with the value of a chosen option (Daw et al., 2006; Hampton et al., 2006; Kable and Glimcher, 2007). However, it has not been possible to disentangle precisely which features of decisions are encoded by vmPFC. By minimizing the correlation between candidate parameters, we were able to demonstrate that the vmPFC encodes the *relative* chosen value, or the value-based evidence in favor of selecting the chosen option, during repetitive decision making. Notably, vmPFC sensitivity to competing but unchosen options suggests that neural activity in this region may underlie the sensitivity of behavior to opportunity costs that has been emphasized in economics (McConnell and Brue, 2004). This perspective accords well with recent demonstrations that vmPFC and other prefrontal regions are sensitive to both potential gains and costs during decision making (Hare et al., 2008; Kable and Glimcher, 2007; Kim et al., 2008; Rushworth et al., 2007; Tanaka et al., 2004; Tom et al., 2007). Our findings are therefore broadly consistent with the notion that vmPFC encodes the *subjective value* of the chosen option during the decision (Kable and Glimcher, 2007). However, in the present study, the overall value parameter incorporates the opportunity cost associated with not selecting alternative actions.

Striatal and orbitofrontal neurons in macaque monkeys have been shown to be sensitive to the chosen value when choices are made between actions and goods, respectively (Lau and Glimcher, 2008; Padoa-Schioppa and Assad, 2006, 2008). It has been suggested that such signals might provide the predicted value component of prediction error signals in dopaminergic neurons crucial for learning (Lau and Glimcher, 2008). It should be noted, however, that striatal coding of chosen value predominately occurs at the time of reward delivery and has not been shown to depend on the value of the alternative action (Lau and Glimcher, 2008). By contrast, the vmPFC signal revealed here occurs at the time of the decision and encodes the value of the chosen *relative* to the unchosen action. Although the positive component of this signal may contribute to prediction error computation, it is also possible that vmPFC encodes the subjective value of a choice—here, a decision reflecting the relative merits of maintaining or switching behavior.

Although relatively little is known of the activity of neurons in the vmPFC of monkeys, a relative value signal has been identified in an adjacent and anatomically interconnected region, the central orbitofrontal cortex (OFC), during a delayed response task (in the absence of a decision) when the same two reinforcements were presented repeatedly within a block of trials (Tremblay and Schultz, 1999). Although the OFC signal may change when the same options are not repeatedly paired with one another (Padoa-Schioppa and Assad, 2008), it is conceivable that the relative chosen value signal uncovered in the vmPFC in the present study reflects afferent input from OFC among other areas. Another distinct possibility is that chosen and unchosen action and offer value signals are integrated in the vmPFC, a region anatomically positioned to send and receive information to and from the striatum and the OFC (Carmichael and Price, 1996; Ferry et al., 2000; Haber, 2003; Parent, 1986).

It has been proposed that organisms are able to change their behavioral patterns flexibly simply by choosing actions according to their expected returns (Bogacz, 2007; Dayan and Abbott, 2001; Sutton and Barto, 1998). While it is easy to see how this strategy might generate adaptive behavior, it requires evaluation and comparison of every possible action at each decision point. Such continual evaluation and comparison would come at the price of a heavy computational load when there are multiple alternatives to choose between. An alternative strategy capable of generating similar behavior is for the animal to continue along a behavioral *policy* until evidence has accumulated in favor

of a behavioral switch. Although the details of such a decision-making mechanism are far from clear, it is possible that only at times of potential switches must the animal perform a comparison between available actions. Because the present study employed a two-alternative choice task, it cannot arbitrate between these two theories of neural coding during repetitive choice. Nevertheless, the FPC signal identified in the current study could potentially accumulate the evidence in favor of a future switch in such a decision-making model.

FPC and vmPFC perform complementary computations during voluntary decision making. Whereas FPC encodes the long-term evidence collected in favor of adapting future behavior, vmPFC encodes a short-term signal reflecting a comparison between the currently chosen and alternative actions. The existence of such signals suggests a novel mechanism for deciding whether or not it is worth adapting or maintaining decisions.

EXPERIMENTAL PROCEDURES

Subjects

Twenty-one healthy volunteers participated in the fMRI experiment. Two volunteers reported discovering and using a nonexistent relationship between the cued reward magnitudes and the rewarded choice, so their data were discarded. An additional participant's data was discarded due to poor echoplanar imaging data quality resulting from gradient coil malfunction. The resulting 18 subjects (8 women, mean age = 27.3 years, standard deviation 3.7 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 task, subjects decided repeatedly between left and right options based on their expectation of reward and the number of points associated with each option (Figure 1A). When the yellow question mark appeared in the center of the screen, subjects indicated their choices with left- and right-hand index finger responses for left and right options, respectively. After subjects indicated their decision, the chosen option was highlighted by a gray rectangle that framed the chosen green rectangle, and subjects awaited the outcome. If the subject's choice was rewarded, a green tick appeared in the center of the screen and the red prize bar also updated toward the silver and gold rectangular targets in proportion to the amount of points won on that trial during the outcome phase. Subjects were rewarded with £20 if the prize bar reached the silver target and £30 if it reached the gold target. No subject failed to reach the first target. If the subject's choice was not rewarded, a red X appeared in the center of the screen. Importantly, there was no feedback given for the unselected option. Each trial consisted of a decide phase (4.5–7.5 s jittered + reaction time) during which subjects made the decision and response; an interval phase (4.5–7.5 s jittered) during which subjects awaited the outcome of their decision; and an outcome phase (3 s) during which the outcome (rewarded or unrewarded) was presented in the center of the screen. The outcome phase was followed by an ITI (3–7 s jittered). There were 120 trials with an average trial time of 21.5 s, resulting in a scanning time of ~43 min.

fMRI Data Acquisition and Analysis

fMRI data were acquired using standard procedures reported in full in the Supplemental Data. The repetition time for acquiring fMRI volumes was 3 s.

fMRI analysis was carried out using FMRIB's Software Library (FSL) (Smith et al., 2004). Conservative independent component analysis was performed using MELODIC to identify and remove obvious motion artifacts (Damoiseaux et al., 2006). Data were then preprocessed using the default options in 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); high-pass temporal filtering was applied using a Gaussian-weighted running lines filter, with a 3 dB cutoff of 100 s.

Separate general linear models (GLMs) were fit in prewhitened data space for probability and expected value metrics (Woolrich et al., 2001). Because probability should theoretically be tracked throughout and across trials, this metric was modeled as a prolonged response. In the GLM for unchosen probability, 11 regressors were included: the main effect of the trial, lasting the duration of the trial, the interaction of demeaned log chosen probability (inferred by the model) and the trial, the interaction of demeaned log unchosen probability (inferred by the model) and the trial, left and right response regressors, and six motion regressors produced during realignment. The means of each regressor listed above, as well as an additional contrast of parameter estimates (COPE) which defined the relative unchosen probability as the log ratio between unchosen and chosen action probabilities, were included in the analysis. Because reward magnitudes were generated randomly and could not be tracked across trials, the expected value should theoretically only be relevant to decisions between trial onset and subject response. This metric was therefore modeled during the decision phase. This second GLM, constructed to identify brain activity associated with the expected value, contained 13 regressors: the main effect of the decide phase, the interval phase, and the monitor phase, the interaction between demeaned subjective chosen expected value (inferred by the model) and the decide phase, the interaction between demeaned subjective unchosen expected value (inferred by the model) and the decide phase, left and right responses, and six motion regressors. Again, COPEs were included for the mean of each above regressor, and an additional COPE defined the relative chosen value as the difference between chosen and unchosen value regressors. Aside from the motion regressors, all regressors were convolved with the FSL default hemodynamic response function (Gamma function, delay = 6 s, standard deviation = 3 s), and filtered by the same high-pass filter as the data.

For group analyses, EPI images were first registered to the high-resolution structural image using 7° of freedom and then to the standard (Montreal Neurological Institute [MNI]) space MNI152 template using affine registration with 12° of freedom (Jenkinson and Smith, 2001). We then fit a GLM to estimate the group mean effects for the regressors described above. FMRIB's Local Analysis of Mixed Effects (FLAME) was used to perform a mixed-effects group analysis that models both "fixed effects" (i.e., within-subject) variance and "random effects" (i.e., between subject) variance (Beckmann et al., 2003; Woolrich et al., 2004). All reported fMRI Z-statistics and p values arose from these mixed effects analyses on all 18 subjects. Inference was carried out using Gaussian random-field theory and cluster-based thresholding, using a cluster-based threshold of $Z > 2.8$ and a whole-brain corrected cluster significance threshold of $p < 0.05$ (Worsley et al., 1992).

Region of Interest Analyses

ROI analyses were conducted on significant clusters of activation identified from the whole-brain voxelwise analysis in order to determine the nature of BOLD fluctuations in these regions. In each subject, we averaged BOLD data from a 27 voxel cube centered on the local maximum back-projected from the FPC, mid-IPS, and vmPFC activations identified by the regressions of interest. Repeated-measures ANOVAs were carried out to test for differences between left- and right-lateralized ROIs in the FPC and mid-IPS. These revealed no difference in effect sizes between left and right frontopolar regions at any phase of the trial (region × trial phase interaction: $F_{2,34} = 0.31$; $p = 0.695$; main effect of region: $F_{1,17} = 0.958$, $p = 0.31$) or between left and right mid-IPS regions at any phase of the trial (region × trial phase interaction: $F_{2,34} = 1.653$, $p = 0.206$; main effect of region: $F_{1,17} = 0.123$ $p = 0.73$). We therefore averaged frontopolar and mid-IPS ROIs across hemispheres to produce a single FPC ROI and a single mid-IPS ROI.

Each subject's BOLD time series was then divided into trials, which were resampled to a duration of 21.5 s, such that the decision was presented at 0 s, the response was made at 7.5 s, and the outcome was presented from 13.5 to 16.5 s, as these were the mean timings for each event across trials and subjects. The resampling resolution was 100 ms. A general linear model was then fit across trials at every time point in each subject independently

(see Figure S6 for an illustration). We then calculated group average effect sizes at each time point, and their standard errors.

To produce the time series of correlation in Figure 3, we constructed a general linear model in which the dependent variable was a matrix comprising the FPC effect size for the relative unchosen probability at every time point in the trial in each subject and the predictor was a vector comprising the behavioral index of interest for each subject. The correlation between the FPC effect size and behavioral index was then plotted at each time point in the trial. To produce scatter plots in Figure 3, we fit a canonical hemodynamic response function (Gamma function, delay = 6 s, standard deviation = 3 s) at the onset of the ITI to the time series of the parameter estimate in the FPC in each subject (Figure S5). The corresponding effect size for the relative unchosen probability in the FPC during the ITI was then plotted against the behavioral index of interest.

Details of the functional connectivity analysis can be found in the Supplemental Data.

Behavioral Model

We used a Bayesian reinforcement-learning algorithm (Behrens et al., 2007) to model subject estimates of the reward probabilities and their eventual choices. This model has two components: a "predictor" that estimates the reward probability associated with each option and other environmental statistics given only the observed data and a "selector" that chooses actions on the basis of these estimates. Our approach assumes that the predictor estimates the reward probabilities and other statistics in a Markovian fashion. On any given trial, the predictor only requires a current estimate of (or belief in) the reward probabilities, the volatility of the environment, and the distrust in the constancy of the volatility of the environment to represent the entire history of previous rewards (see Supplemental Data for an algebraic description). These estimates of the reward probabilities were then combined with reward magnitude according to subject-specific free parameters that can differentially weigh probability and magnitude, corresponding to risk-averse and risk-prone behavior, respectively, to derive estimates of the subjective expected action values (Supplemental Data). Finally, the selector assumed that subjects generated actions according to the following sigmoidal probability distribution:

$$P(C = \text{Left}) = 1/(1 + \exp(-g_{\text{left}})), \quad (1)$$

where g_{left} is the subjective expected value of the left option.

SUPPLEMENTAL DATA

Supplemental Data include Supplemental Experimental Procedures, six figures, and two tables and can be found with this article online at http://www.cell.com/neuron/supplemental/S0896-6273(09)00389-4.

ACKNOWLEDGMENTS

Funded by the Wellcome Trust (E.D.B.), the EPSRC (M.W.W.), and the MRC (T.E.J.B., M.F.S.R.). We would like to thank Dr Mark Walton and Professor Richard Passingham for helpful discussions and Mr Steven Knight for help with scanning.

Accepted: May 6, 2009
Published: June 10, 2009

REFERENCES

Beckmann, C.F., Jenkinson, M., and Smith, S.M. (2003). General multilevel linear modeling for group analysis in FMRI. Neuroimage *20*, 1052–1063.

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

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

Bogacz, R. (2007). Optimal decision-making theories: linking neurobiology with behaviour. Trends Cogn. Sci. *11*, 118–125.

Braver, T.S., Reynolds, J.R., and Donaldson, D.I. (2003). Neural mechanisms of transient and sustained cognitive control during task switching. Neuron *39*, 713–726.

Bunge, S.A., Wendelken, C., Badre, D., and Wagner, A.D. (2005). Analogical reasoning and prefrontal cortex: evidence for separable retrieval and integration mechanisms. Cereb. Cortex *15*, 239–249.

Burgess, P.W. (2000). Strategy application disorder: the role of the frontal lobes in human multitasking. Psychol. Res. *63*, 279–288.

Burgess, P.W., Veitch, E., de Lacy Costello, A., and Shallice, T. (2000). The cognitive and neuroanatomical correlates of multitasking. Neuropsychologia *38*, 848–863.

Burgess, P.W., Scott, S.K., and Frith, C.D. (2003). The role of the rostral frontal cortex (area 10) in prospective memory: a lateral versus medial dissociation. Neuropsychologia *41*, 906–918.

Burgess, P.W., Dumontheil, I., and Gilbert, S.J. (2007). The gateway hypothesis of rostral prefrontal cortex (area 10) function. Trends Cogn. Sci. *11*, 290–298.

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

Cavada, C., and Goldman-Rakic, P.S. (1989). Posterior parietal cortex in rhesus monkey: I. Parcellation of areas based on distinctive limbic and sensory corticocortical connections. J. Comp. Neurol. *287*, 393–421.

Damoiseaux, J.S., Rombouts, S.A., Barkhof, F., Scheltens, P., Stam, C.J., Smith, S.M., and Beckmann, C.F. (2006). Consistent resting-state networks across healthy subjects. Proc. Natl. Acad. Sci. USA *103*, 13848–13853.

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

Dayan, P., and Abbott, L.F. (2001). Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems (Cambridge, MA: MIT Press).

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

Genovesio, A., and Wise, S.P. (2008). The Neurophysiology of Abstract Response Strategies. In Neuroscience of Rule-Guided Behavior, S.A. Bunge and J.D. Wallis, eds. (New York: Oxford University Press), pp. 81–106.

Genovesio, A., Brasted, P.J., Mitz, A.R., and Wise, S.P. (2005). Prefrontal cortex activity related to abstract response strategies. Neuron *47*, 307–320.

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

Grol, M.J., de Lange, F.P., Verstraten, F.A., Passingham, R.E., and Toni, I. (2006). Cerebral changes during performance of overlearned arbitrary visuomotor associations. J. Neurosci. *26*, 117–125.

Haber, S.N. (2003). The primate basal ganglia: parallel and integrative networks. J. Chem. Neuroanat. *26*, 317–330.

Hampton, A.N., Bossaerts, P., and O'Doherty, J.P. (2006). The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. J. Neurosci. *26*, 8360–8367.

Hare, T.A., O'Doherty, J., Camerer, C.F., Schultz, W., and 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.

Husain, M., and Nachev, P. (2007). Space and the parietal cortex. Trends Cogn. Sci. *11*, 30–36.

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

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

Jubault, T., Ody, C., and Koechlin, E. (2007). Serial organization of human behavior in the inferior parietal cortex. J. Neurosci. *27*, 11028–11036.

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

Kim, S., Hwang, J., and Lee, D. (2008). prefrontal coding of temporally discounted values during intertemporal choice. Neuron *59*, 161–172.

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

Koechlin, E. (2008). The cognitive architecture of human lateral prefrontal cortex. In Sensorimotor Foundations of Higher Cognition, P. Haggard, Y. Rossetti, and M. Kawato, eds. (Oxford: Oxford University Press), pp. 483–509.

Koechlin, E., Ody, C., and Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. Science *302*, 1181–1185.

Koechlin, E., and Jubault, T. (2006). Broca's area and the hierarchical organization of human behavior. Neuron *50*, 963–974.

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

Koechlin, E., and Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. Trends Cogn. Sci. *11*, 229–235.

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

Koechlin, E., Corrado, G., Pietrini, P., and Grafman, J. (2000). Dissociating the role of the medial and lateral anterior prefrontal cortex in human planning. Proc. Natl. Acad. Sci. USA *97*, 7651–7656.

Lau, B., and Glimcher, P.W. (2008). Value representations in the primate striatum during matching behavior. Neuron *58*, 451–463.

Matelli, M., Camarda, R., Glickstein, M., and Rizzolatti, G. (1986). Afferent and efferent projections of the inferior area 6 in the macaque monkey. J. Comp. Neurol. *251*, 281–298.

McConnell, C.R., and Brue, S.L. (2004). Microeconomics: Principles, Problems, and Policies, 16th edition (Boston, MA: McGraw-Hill Professional).

Miller, E.K., and Buschman, T.J. (2008). Rules through Recursion: How Interactions between the Frontal Cortex and Basal Ganglia May Build Abstract, Complex Rules from Concrete, Simple Ones. In Neuroscience of Rule-Guided Behavior, S.A. Bunge and J.D. Wallis, eds. (New York: Oxford University Press), pp. 419–440.

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

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

Parent, A. (1986). Comparative Neurobiology of the Basal Ganglia (Chichester: John Wiley and Sons, Inc).

Passingham, R.E. (2008). What Is Special about the Human Brain? (New York: Oxford University Press).

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

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

Prabhakaran, V., Narayanan, K., Zhao, Z., and Gabrieli, J.D. (2000). Integration of diverse information in working memory within the frontal lobe. Nat. Neurosci. *3*, 85–90.

Ramnani, N., and Owen, A.M. (2004). Anterior prefrontal cortex: insights into function from anatomy and neuroimaging. Nat. Rev. Neurosci. *5*, 184–194.

Rushworth, M.F., Ellison, A., and Walsh, V. (2001). Complementary localization and lateralization of orienting and motor attention. Nat. Neurosci. *4*, 656–661.

Rushworth, M.F., Behrens, T.E., and Johansen-Berg, H. (2006). Connection patterns distinguish 3 regions of human parietal cortex. Cereb. Cortex *16*, 1418–1430.

Rushworth, M.F., Buckley, M.J., Behrens, T.E., Walton, M.E., and Bannerman, D.M. (2007). Functional organization of the medial frontal cortex. Curr. Opin. Neurobiol. *17*, 220–227.

Sakai, K., and Passingham, R.E. (2006). Prefrontal set activity predicts rule-specific neural processing during subsequent cognitive performance. J. Neurosci. *26*, 1211–1218.

Sakai, K., Rowe, J.B., and Passingham, R.E. (2002). Active maintenance in prefrontal area 46 creates distractor-resistant memory. Nat. Neurosci. *5*, 479–484.

Schluter, N.D., Rushworth, M.F., Passingham, R.E., and Mills, K.R. (1998). Temporary interference in human lateral premotor cortex suggests dominance for the selection of movements. A study using transcranial magnetic stimulation. Brain *121*, 785–799.

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

Smith, S.M. (2002). Fast robust automated brain extraction. Hum. Brain Mapp. *17*, 143–155.

Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage *23* (*Suppl 1*), S208–S219.

Sutton, R.S., and Barto, A.G. (1998). Reinforcement Learning: An Introduction (Cambridge, MA: MIT Press).

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

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

Toni, I., Rushworth, M.F., and Passingham, R.E. (2001). Neural correlates of visuomotor associations. Spatial rules compared with arbitrary rules. Exp. Brain Res. *141*, 359–369.

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

Wendelken, C., Nakhabenko, D., Donohue, S.E., Carter, C.S., and Bunge, S.A. (2008). "Brain is to thought as stomach is to??": investigating the role of rostrolateral prefrontal cortex in relational reasoning. J. Cogn. Neurosci. *20*, 682–693.

Wise, S.P. (2008). Forward frontal fields: phylogeny and fundamental function. Trends Neurosci. *31*, 599–608.

Wise, S.P., Boussaoud, D., Johnson, P.B., and Caminiti, R. (1997). Premotor and parietal cortex: corticocortical connectivity and combinatorial computations. Annu. Rev. Neurosci. *20*, 25–42.

Woolrich, M.W., Ripley, B.D., Brady, M., and Smith, S.M. (2001). Temporal autocorrelation in univariate linear modeling of FMRI data. Neuroimage *14*, 1370–1386.

Woolrich, M.W., Behrens, T.E., Beckmann, C.F., Jenkinson, M., and Smith, S.M. (2004). Multilevel linear modelling for FMRI group analysis using Bayesian inference. Neuroimage *21*, 1732–1747.

Worsley, K.J., Evans, A.C., Marrett, S., and Neelin, P. (1992). A three-dimensional statistical analysis for CBF activation studies in human brain. J. Cereb. Blood Flow Metab. *12*, 900–918.

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