

Anterior Cingulate Cortex Instigates Adaptive Switches in Choice by Integrating Immediate and Delayed Components of Value in Ventromedial Prefrontal Cortex

Marcos Economides,^{1*} Marc Guitart-Masip,^{1,2*} Zeb Kurth-Nelson,¹ and Raymond J. Dolan¹

¹Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, London WC1N 3BG, United Kingdom, and ²Ageing Research Centre, Karolinska Institute, SE-11330 Stockholm, Sweden

Actions can lead to an immediate reward or punishment and a complex set of delayed outcomes. Adaptive choice necessitates the brain track and integrate both of these potential consequences. Here, we designed a sequential task whereby the decision to exploit or forego an available offer was contingent on comparing immediate value and a state-dependent future cost of expending a limited resource. Crucially, the dynamics of the task demanded frequent switches in policy based on an online computation of changing delayed consequences. We found that human subjects choose on the basis of a near-optimal integration of immediate reward and delayed consequences, with the latter computed in a prefrontal network. Within this network, anterior cingulate cortex (ACC) was dynamically coupled to ventromedial prefrontal cortex (vmPFC) when adaptive switches in choice were required. Our results suggest a choice architecture whereby interactions between ACC and vmPFC underpin an integration of immediate and delayed components of value to support flexible policy switching that accommodates the potential delayed consequences of an action.

Key words: fMRI; decision-making; value; control; computational modeling; prefrontal cortex

Introduction

Humans and other animals make decisions by assigning values to candidate options that compete for action selection (Rangel et al., 2008). To ensure an outcome is optimal, an agent needs to infer long-term expected value by integrating over several components, including the current goal and the downstream consequence of acting. A growing understanding of how hierarchical goals influence value comparison (Hare et al., 2009, 2011) contrasts with a dearth of knowledge regarding how the brain infers and integrates downstream consequence when evaluating options in a changing environment.

Paradigms requiring calculations of long-term value recruit the prefrontal cortex (PFC; Balleine and Dickinson, 1998; Wallis and Miller, 2003; Basten et al., 2010; Gläscher et al., 2010; Rangel and Hare, 2010). In particular, the dorsolateral PFC (DLPFC) has been linked to task planning (van den Heuvel et al., 2003; Wunderlich et al., 2012), the representation of abstract task rules

(Buschman et al., 2012; Stokes et al., 2013), as well as discounted or goal values (McClure et al., 2004; Plassmann et al., 2010). However, these studies do not address how the brain infers long-term value when decisions are sequential and integrative. It is of interest that several tasks requiring cognitive control implicitly evoke representations of downstream consequence, and as such it seems plausible that these processes could be subserved by a common neural mechanism. In a typical example, an external cue signals a categorical contingency switch that instantiates a change in action or the inhibition of a prepotent response (Kerns et al., 2004). Although such tasks highlight a frontoparietal network as being central to control (Botvinick et al., 2001; Badre, 2008), they are seldom deployed in the value domain, and a focus on isolated choice neglects downstream consequences of decisions. Recent studies have touched on these issues implicating parietal regions and PFC in representing the state transitions necessary for building a model of the world (Gläscher et al., 2010; Wunderlich et al., 2012). It remains unclear what computational role these regions play when action control is reliant on a subjective inference about a change in expected value.

Here, we tested whether a context-specific evaluation of action could explain choice in a value-guided sequential go/no-go paradigm, whereby an agent tracks time-varying contingencies of a dynamic environment to adapt behavior in anticipation of future value. Building on previous studies, our paradigm allowed comparisons between policy switches arising from either inference or an external cue that the environment had changed. Thus, this task enabled us to characterize the computations tracked by the brain in a dynamic world. We

Received Oct. 8, 2013; revised Jan. 7, 2014; accepted Jan. 25, 2014.

Author contributions: M.E., M.G.-M., and R.J.D. designed research; M.E. and M.G.-M. performed research; M.E., M.G.-M., and Z.K.-N. analyzed data; M.E., M.G.-M., Z.K.-N., and R.J.D. wrote the paper.

This work was supported by Wellcome Trust Ray Dolan Senior Investigator Award 098362/Z/12/Z and Medical Research Council Grant G1000411. The Wellcome Trust Centre for Neuroimaging is supported by core funding from the Wellcome Trust (Grant 091593/Z/10/Z). We thank Peter Dayan for valuable comments on a previous version of this manuscript.

The authors declare no competing financial interests.

*M.E. and M.G.-M. contributed equally to this work.

Correspondence should be addressed to Marcos Economides at the above address. E-mail: m.economides@ucl.ac.uk.

DOI:10.1523/JNEUROSCI.4313-13.2014

Copyright © 2014 the authors 0270-6474/14/343340-10\$15.00/0

predicted that PFC would compute the downstream consequence of acting by tracking changing aspects of the environment and interact with regions such as ventromedial PFC (vmPFC) and striatum, both strongly implicated in reward (Kable and Glimcher, 2007), to compute an integrated signal of long-term value for guiding choice policy.

Materials and Methods

Subjects

Twenty-one adults participated in the experiment (nine males and 12 females; age range, 19–28 years; mean \pm SD, 23.2 \pm 2.3 years). All were healthy, reporting no history of neurological, psychiatric, or other current medical problems. Subjects provided written informed consent to partake in the study, which was approved by the local ethics board (University College London, London, UK).

Training paradigm

In a conditioning phase, performed outside of the scanner, subjects learned stimulus–reward associations between a set of four differently colored rectangular cues and their respective monetary values. Each colored rectangle corresponded to one of four possible value outcomes—1, 2, 3, or 4 tokens—randomized across individuals. Subjects were instructed that each token would translate into a fixed sum of money at the end of the experiment.

Each trial began with a central fixation cross presented for 1000 ms, followed by presentation of a random pair of colored boxes, one appearing to the left of the screen and one to the right. Subjects had a 2000 ms time window to choose between these two boxes via a left or right button press, followed by presentation of the outcome of their choice for 1000 ms. The outcome was revealed as a written message indicating the total number of tokens won. Subjects were instructed to explore all options until they were confident they had learned all four associations, after which they should choose the box from the pair with the higher value. Each trial was defined as correct if the subject chose the more valuable of the two options and incorrect if they failed to do so. To ensure adequate learning, performance was calculated over six bins of 20 trials, with all subjects reaching a performance criterion of $\geq 90\%$ by trial 60 onward. For absolute verification, subjects were asked to verbally communicate the nature of the learned associations.

Task paradigm

On every trial, subjects were presented with a random sequence of trained stimuli (see training paradigm), appearing individually and sequentially, with a variable interstimulus interval (750–1250 ms). The sequence order was pseudorandom and thus unpredictable, with each stimulus having an equal probability of being one of the four possible colors. In addition, the precise number of stimuli to be offered on any trial was uncertain, fluctuating under a uniform distribution between 3 and 7.

Each stimulus constituted an offer with a worth equivalent to its respective token value, for which subjects had 1500 ms to accept or reject via a go or no-go response, respectively. A restriction was placed on the number of offers that could be exploited. In high-constraint (HC) trials, the acceptance budget was between 1 and 3 offers, whereas in low-constraint (LC) trials, it ranged between 4 and 6 offers, both varying under a uniform distribution independent of the total number of offers made on the current trial. Subjects were not explicitly told the bounds of the distributions from which the number of offers and total budget were drawn, only that they were uniform. All subjects received 30 training trials (15 per condition) to infer these distributions and familiarize themselves with the task attributes.

HC and LC trials were pseudorandomly interleaved. The trial type was indicated via a small or large green circle, in the top central portion of the screen, for HC and LC, respectively. This appeared at trial onset and turned red after exhaustion of the budget, indicating that no-go responses were obligatory for the remainder of the trial. After the final offer, an outcome incorporating the total number of tokens won and the corresponding cue-token credit breakdown was revealed for 2500 ms.

One hundred twenty trials (60 per condition) were completed in the

scanner across four sessions. The number of tokens won across sessions was summed and converted to a cash prize.

Behavioral data analyses

Global behavior. Our analysis focused exclusively on choices pertaining to within-budget offers. Accepts (go responses) were obtained as a percentage of the total offer number at each offer value, conditional on HC and LC trials. These measures were entered into a two-way repeated-measures ANOVA with factors control (HC/LC) and offer value (1, 2, 3, or 4). The data were analyzed in the statistical software package SPSS, version 20.0.

Within-trial modulation of choice. Within a trial, a player transitioned through a number of discrete states dependent on two fluctuating variables, the number of offers already seen, and the number of accepts already used. To assess whether the probability of accepting a given offer was flat across the entire length of a given trial or fluctuated as a function of these variables, we split trials by offer index (i.e., 1–7) and number of offers already rejected (i.e., 0–6), recalculating the probability of accepting at every possible permutation (see Fig. 2). For each participant, we summed the number of offers with a given value presented at each possible state within a trial and then summed the number of accepts at each of those states. Dividing these measures gave us a probability of acceptance at every choice point. Thus, for both HC and LC trials and each offer value, we generated a separate probability accept matrix, with offer number increasing along the x -dimension and number of rejects increasing along the y -dimension. These matrices were averaged across all participants. For display purposes, we discarded cells with less than a total of 10 data points.

Computational modeling

Because we were interested in assaying subjects' strategy for maximizing reward, we evaluated evidence for four competing choice models. Broadly, we conjectured that subjects might approach trials with a predetermined decision rule, in effect applying a heuristic uniformly throughout a trial. Alternatively, because of uncertainty surrounding the number of expected offers and the go budget (the number of offers they can exploit for reward in a trial), subjects might continually adapt their threshold for accepting offers across a trial. We outline the distinct models below, ordered by increasing complexity, in which each model calculated the expected value of accepting an offer (V_A), which was then passed through a sigmoid function to determine action probabilities as follows:

$$P_A = \frac{1}{1 + \exp(-\tau \times V_A)},$$

where τ is a temperature parameter that governs the stochasticity of choices.

Baseline heuristic model

We specified a baseline heuristic model that calculates V_A by comparing the (face) value of every offer to a stationary decision threshold:

$$V_A = R - c_1,$$

where R is the (face) value of the current offer, and c_1 is a value threshold.

Thus, this model makes choices based solely on the immediate (face) value of an offer with the probability of acceptance fixed throughout a trial.

The model has three free parameters: the associated decision threshold for both HC and LC separately, and the steepness of the sigmoid function.

Sliding offer model

We conjectured that subjects might track the number of offers seen in a trial and adjust a decision threshold such that an offer is more likely to be accepted if forthcoming offers were scarce. We added a linear slope parameter to the baseline heuristic model that governed the steepness of this decay across a trial, such that

$$V_A = R - (c_1 - o \times c_2),$$

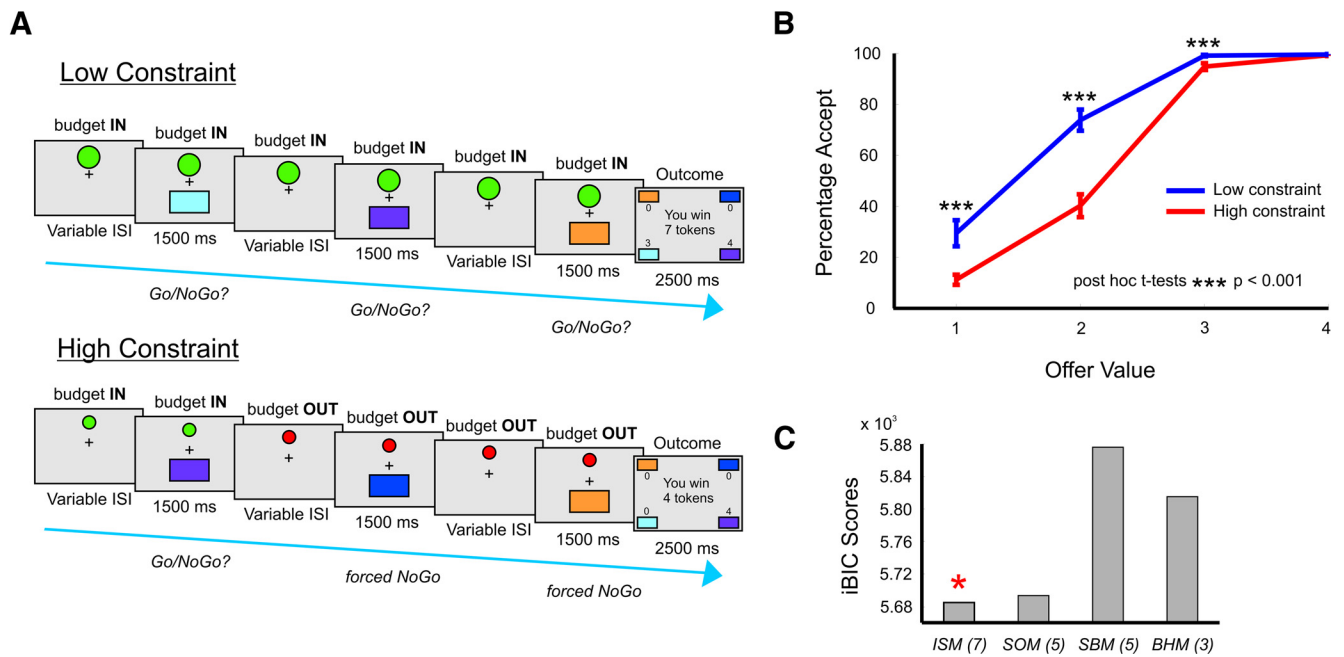


Figure 1. Behavioral paradigm and results. **A**, Subjects learned stimulus–value associations, ranging from one to four tokens, for four colored stimuli. On every trial, participants saw a random sequence of these stimuli, varying unpredictably in length between three and seven, with each stimulus representing an offer requiring either a go response to win the associated tokens or a no-go response for no token gain (for simplicity, the illustrations span 3 offers). Subjects had a predetermined go budget that placed a restriction on the number of offers that could be accepted. In an LC context, subjects could accept between four and six offers but only between one and three in an HC context, with the exact budget being uncertain. After exhausting a go budget, no-go responses were enforced for the remainder of the trial. The context or condition was cued via a large (LC) or small (HC) green circle, whereas a depleted budget was signaled via the green circle turning red. **B**, Mean percentage of offers accepted split by token value and condition (HC in red, LC in blue). Subjects were less willing to accept low-value offers when the budget was scarce. *Post hoc* paired *t* tests revealed significant decreases in percentage accept for offer values 1, 2, and 3 in HC compared to LC (all $p < 0.001$). Vertical lines represent SEM. **C**, Integrated BIC scores (for the group as a whole) show that a model in which both the number of offers already seen and number of offers already accepted/rejected are used to adjust the threshold for action fits behavior best. ISM, Integrated sliding model; SOM, sliding offers model; SBM, sliding budget model; BHM, baseline heuristic model. The number of free parameters built into each model is indicated in parentheses.

where R is the (face) value of the current offer, c_1 is a value threshold, o is the current offer index, and c_2 governed the steepness of the associated slope.

The model has five free parameters: the associated decision threshold and a slope parameter for both HC and LC separately, and the steepness of the sigmoid function.

Sliding budget model

A second variable that subjects could track to dynamically adjust their decision threshold is the number of offers already accepted in a trial. Given a limited go budget, a player may be less likely to accept an offer as this resource is exhausted, assuming ample offers. This model linearly increased the decision threshold with every additional offer accepted but did not take into account the abundance of remaining offers, such that

$$V_A = R - (c_1 + a \times c_2),$$

where R is the (face) value of the current offer, c_1 is a value threshold, a is the number of offers accepted previously, and c_2 governed the steepness of the associated slope.

The model has five free parameters: the associated decision threshold and a slope parameter for both HC and LC separately, and the steepness of the sigmoid function.

Integrated sliding model

Combining the sliding offer and sliding budget models, subjects could track both the number of offers seen and the number of offers already accepted in a trial, using each source of information to adjust the decision threshold. The threshold should drop linearly with every mounting offer and rise linearly with every mounting go response. We fit separate slope parameters that governed the linear gradient for the number of offers and number of accepts, such that

$$V_A = R - (c_1 + a \times c_2 - o \times c_3),$$

where R is the (face) value of the current offer, c_1 is a value threshold, a is the number of offers accepted previously, o is the current offer index, and c_2 and c_3 govern the steepness of the associated slopes.

Interestingly, this two-factor model predicts the optimal action with a frequency of 87% (based on group mean parameter fits).

The model has seven free parameters: the associated decision threshold, a slope parameter for the number of offers, a slope parameter for the number of accepts for both HC and LC separately, and a parameter for the steepness of the sigmoid function.

Model comparison

As described previously (Huys et al., 2011; Guitart-Masip et al., 2012), we used a hierarchical type II Bayesian (or random effects) procedure using maximum likelihood to fit simple parameterized distributions for higher-level statistics of the parameters. Because the values of parameters for each subject are “hidden,” this uses the expectation-maximization procedure. Thus, on each iteration, the posterior distribution over the group for each parameter is used to specify the prior over the individual parameter fits on the next iteration. For each parameter, we used a single distribution for all participants. Before inference, all parameters were suitably transformed to enforce constraints (log and inverse sigmoid transforms).

Models were compared using the integrated Bayesian information criterion (iBIC), in which small iBIC values indicate a model that fits the data better after penalizing for the number of parameters. Comparing iBIC values is akin to a likelihood ratio test (Kass and Raftery, 1995).

Reaction time analyses

We conjectured that, if subjects were evaluating choice options in light of an action threshold that fluctuated in accordance with the number of offers already seen and accepted/rejected, then reaction times should be faster when the associated threshold is low and a go response is relatively more valuable. To test this, we used multiple linear regression to model

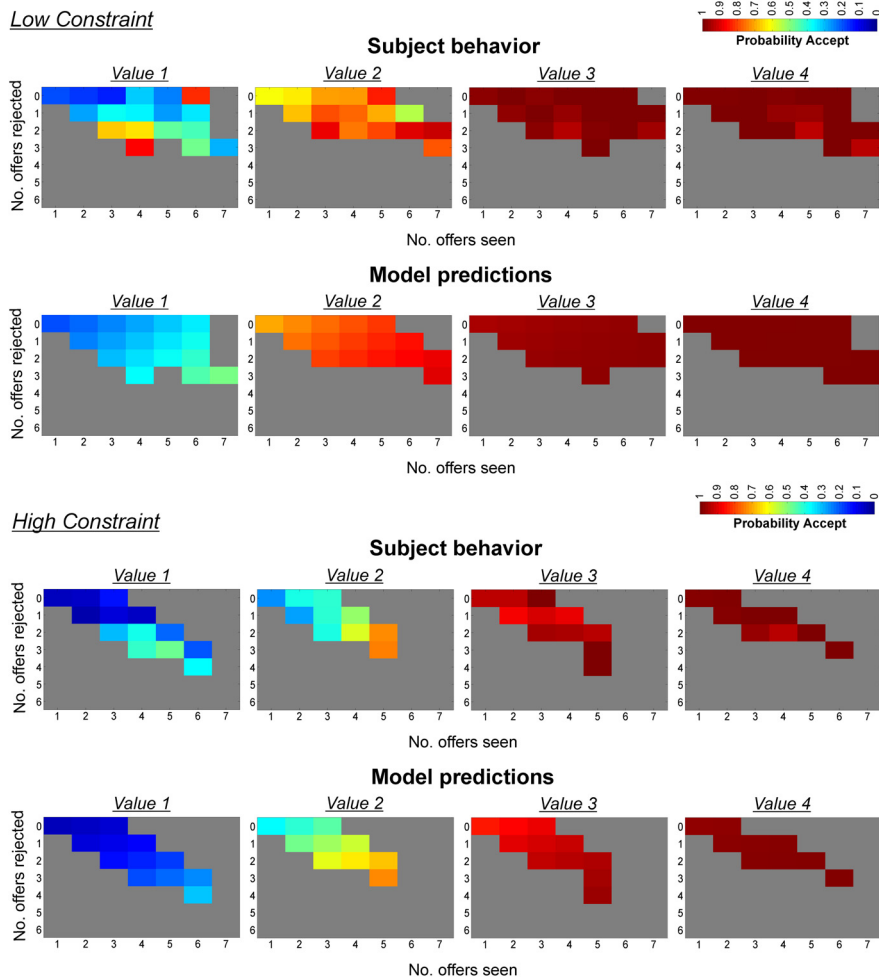


Figure 2. Subjects' control strategy. Subjects adjust the probability of accepting less desirable offers as a function of the number of offers seen (x -axis) and number of offers already rejected (y -axis). The spectrum runs from blue (probability 0) to red (probability 1).

the dependence of reaction times for all go choices on the corresponding offer values (immediate values) and model thresholds, separately for HC and LC trials. The two regressors were forced to compete for variance so as to explore dissociable contributions to the observed reaction times.

fMRI data acquisition

fMRI was performed on a 3 Tesla Siemens Quattro magnetic resonance scanner with echo planar imaging (EPI) and 32-channel head coil. Functional data were acquired over four sessions containing 166 volumes with 48 slices (664 volumes total). Acquisition parameters were as follows: matrix, 64×74 ; oblique axial slices angled at -30° in the anteroposterior axis; spatial resolution, $3 \times 3 \times 3$ mm; TR, 3360 ms; TE, 30 ms. The first five volumes were subsequently discarded to allow for steady-state magnetization. Field maps were acquired before the functional runs (matrix, 64×64 ; 64 slices; spatial resolution, $3 \times 3 \times 3$ mm; gap, 1 mm; short TE, 10 ms; long TE, 12.46 ms; TR, 1020 ms). Anatomical images of each subject's brain were collected using multi-echo 3D fast, low-angle shot sequence for mapping proton density, T1 and magnetization transfer at 1 mm^3 resolution, and by T1-weighted inversion recovery prepared EPI sequences (spatial resolution, $1 \times 1 \times 1$ mm) with B1 mapping data to correct for the effect of inhomogeneous transmit fields on the T1 maps (3D EPI transverse partition direction; matrix, 64×48 ; phase direction, right to left; 48 partitions; resolution, $4 \times 4 \times 4$ mm).

During scanning, peripheral measurements of subject pulse and breathing were made together with scanner slice synchronization pulses using the Spike2 data acquisition system (Cambridge Electronic Design). The cardiac pulse signal was measured using an MRI-compatible pulse

oximeter (model 8600 F0; Nonin Medical) attached to the subject's finger. The respiratory signal (thoracic movement) was monitored using a pneumatic belt positioned around the abdomen close to the diaphragm.

fMRI data analyses

Data were analyzed using SPM8 (Wellcome Trust Centre for Neuroimaging, University College London). Functional data were bias corrected for 32-channel head coil intensity inhomogeneities. Preprocessing involved realignment and unwarping using individual field maps, coregistration of EPI to T1-weighted images, and spatial normalization to the Montreal Neurological Institute (MNI) space using the segmentation algorithm on the T1-weighted image with a final spatial resolution of $1 \times 1 \times 1$ mm. Finally, data were smoothed with an 8 mm FWHM Gaussian kernel. The fMRI time series data were high-pass filtered (cutoff, 128 s) and whitened using a first-order auto-correlation (AR(1)) model.

For each subject, we used an in-house MATLAB toolbox (Hutton et al., 2011) to construct a physiological noise model to account for artifacts that take account of cardiac and respiratory phase, as well as changes in respiratory volume. This resulted in a total of 14 regressors that were sampled at a reference slice in each image volume to give a set of values for each time point. The resulting regressors were included as confounds in all first-level GLMs.

To identify brain areas sensitive to within-trial variations in choice prescribed by our model, we derived an offer-wise go threshold to use as a parametric modulator of offer onsets in all first-level GLMs. This model threshold (MT) represented an intercept value that increased linearly with every offer accepted and decreased linearly with every offer seen. The intercept and slopes were based on the mean posterior parameter fits across the group. If the offer value was higher than the MT, the preferable decision is accepting, otherwise rejecting is preferred.

Below we outline the GLM constructed for first-level analyses. All imaging analyses address time points when offers are within budget and the subject has a free choice. Results are reported whole-brain corrected at the cluster level (FWE $p < 0.05$) unless otherwise stated.

To enable us to explore a main effect of action constraint and value/MT (and their relevant interactions), we split offer onsets according to constraint (HC/LC) and offer (face) value (1, 2, 3, or 4), modeling each in a separate regressor parametrically modulated by MT. This resulted in 16 regressors of interest. The four scanning sessions were concatenated into one, and a binary matrix was included to encode the identity of each session. Additional regressors of no interest included six movement-related covariates (the three rigid-body translations and three rotations resulting from realignment), 14 physiological regressors (six respiratory, six cardiac, and two change in respiratory/heart rate), the onsets of the go responses (to explain away the effects of action), all offers outside of budget (for which "no-go" responses were enforced) parametrically modulated by offer value, and outcome onsets parametrically modulated by the relevant number of tokens won. All regressors were modeled as stick functions with duration of zero and convolved with a canonical form of the hemodynamic response function (HRF) combined with time and dispersion derivatives.

At the second level, we conducted a random-effects 2×4 ANOVA with factors condition (HC/LC) and offer value (1, 2, 3, or 4), using first-level contrast images corresponding to the onset regressors of inter-

est for each participant. This enabled us to explore main effects of condition and value and their interaction. We generated a second 2×4 random-effects ANOVA drawing on first-level contrast images from the eight MT parametric modulators, to explore an average effect of MT and a $MT \times$ value interaction. To obtain an average estimate of DLPFC activation in HC compared with LC, parameter estimates for offer values 1–4 were averaged in each condition, and LC was subtracted from HC.

Functional regions of interest

We used a functional regions of interest (f-ROI) approach to extract parameter estimates in a priori regions for a subset of analyses, including correlating neural and behavioral measures, comparing value representations between conditions and exploring functional connectivity patterns. f-ROIs were derived by identifying significant clusters of activation surrounding peak voxels from the relevant whole-brain mass univariate analysis. Given that these clusters often spanned multiple regions, activations were constrained to corresponding anatomical ROIs from the MarsBar toolbox (version 0.42) for SPM. For the ventral striatum (VS), activations were constrained to an anatomical ROI derived from a diffusion tensor imaging connectivity-based parcellation of the right nucleus accumbens (NA) in humans (taken from Baliki et al., 2013). The ROI consisted of both the core and shell subcomponents of NA, and the right region was flipped along the x -dimension in the MarsBar toolbox to obtain a bilateral accumbens mask.

Psychophysiological interaction

For each subject, we defined a volume of interest that included all active voxels (at $p = 0.2$) from a first-level contrast that specified a linear effect of model thresholds across offered value $\{-2 -1 1 2\}$ within f-ROIs derived from the same second-level contrast (see Fig. 5A, black arrows). This allowed us to define voxels active on a subject-by-subject basis but confined to the cluster active at the group level. We note that 1 of 21 subjects had no active voxels when specifying both the anterior cingulate cortex (ACC) and left DLPFC (BA46) as seeds, whereas 3 of 21 subjects had no active voxels when specifying dorsal vmPFC as a seed. These subjects were excluded from the corresponding psychophysiological interaction (PPI) analysis. We used the generalized PPI toolbox for SPM (gPPI; <http://www.nitrc.org/projects/gppi>) to create a new GLM in which the individual seed time course was deconvolved to construct a neuronal time course for multiplication with regressors modeling all task effects and then reconvolved with the HRF. Thus, the gPPI GLM includes a psychophysiological regressor for all conditions (McLaren et al., 2012). An indicator function for the relevant contrast, the original BOLD eigenvariate, and six motion and 12 physiological parameters were included as additional regressors. We first looked for regions in which connectivity with the seed region was modulated by MT, but when this modulation was greater for offers requiring adaptive control (values 1 and 2 in HC, and value 1 in LC > values 3 and 4 in HR, and values 2, 3, and 4 in LC). We also performed a second PPI restricted to offers requiring adaptive choice (values 1 and 2 in HC, and value 1 in LC) to ascertain whether connectivity increased (positive PPI) or decreased (negative PPI) with respect to increases in MT (compared with 0). One-sample t tests were performed on the relevant contrasts at the second level.

Results

Subjects reject lower value offers when a go budget is scarce

Subjects were sensitive to both immediate (face) value and the delayed consequences arising from a budget constraint. Higher value offers were accepted more than lower value offers [a main effect of value: $F_{(1,68,33,63)} = 277.87$, mean squared error (MSE) = 379.38, $p < 0.001$], and more offers were accepted overall in LC compared with HC (a main effect of constraint: $F_{(1,20)} = 182.70$, MSE = 45.69, $p < 0.001$). Importantly, subjects were less willing to accept low-value offers in HC compared with LC (a budget constraint \times value interaction: $F_{(1,73,34,67)} = 30.41$, MSE = 136.19, $p < 0.001$; Fig. 1B).

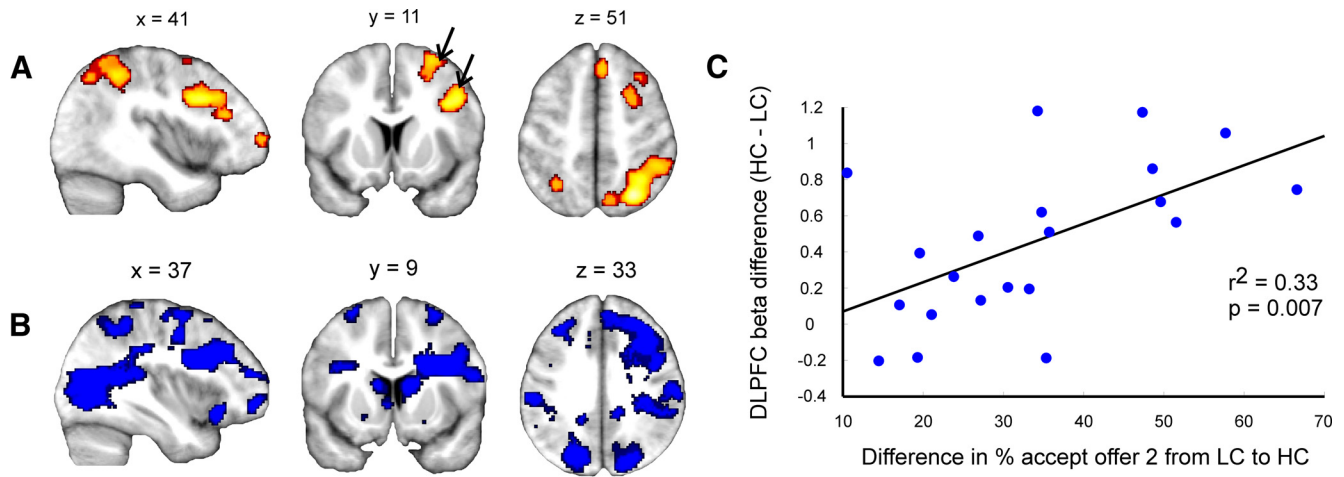
Table 1. Summary of fMRI second-level statistics

Name of region	Cluster FWE p value	MNI coordinates			Statistics	
		x	y	z	t value	Z score
HC > LC						
Right parietal	<0.001	28	−64	48	5.70	5.39
Right DLPFC	<0.001	40	12	30	4.75	4.56
Left parietal	0.002	−28	−56	46	4.53	4.37
LC > HC						
Right V1	<0.001	8	−76	6	7.38	6.77
Left V1		−8	−84	−8	6.08	5.72
Left parahippocampal	<0.001	−28	−28	−12	5.80	5.48
Left parietal	<0.001	−50	−24	24	5.77	4.46
Left insula		−40	−6	−2	4.59	4.42
vmPFC	<0.001	−6	44	−10	5.06	4.84
Left precuneus	<0.001	−10	−54	−12	4.56	4.39
Right parietal	<0.001	44	−34	24	4.49	4.33
Mid-cingulate	0.001	14	−20	46	4.08	3.96
Linear effect offer value						
vmPFC	<0.001	4	52	14	6.31	5.91
Bilateral accumbens		−4	14	−8	5.70	5.39
Left mid-temporal	<0.001	−52	−58	20	5.83	5.50
Left parietal		−54	−26	22	5.62	5.33
Left superior temporal		−58	−18	10	4.46	4.30
Left mid-occipital		−42	−74	32	4.15	4.02
Left parahippocampal	<0.001	−28	−32	−14	5.67	5.37
Right lingual		14	−44	2	5.45	5.18
Right cuneus	<0.001	16	−82	32	5.65	5.35
Right M1	0.004	56	−8	44	5.38	5.12
Left M1	0.001	−44	−16	58	4.91	4.71
Right hippocampus	0.008	24	−18	−16	4.89	4.69
Negative offer value						
Right insula	<0.001	30	22	−10	5.44	5.17
ACC	<0.001	6	24	48	5.26	5.02
Left insula	0.009	−34	18	−4	5.04	4.82
Right parietal	0.001	36	−50	50	4.47	4.31
Negative model thresholds						
Left caudate	<0.001	−12	−6	18	8.23	7.42
Right superior parietal		24	−56	52	7.95	7.21
Right IFGpt		32	18	28	7.92	7.19
Right thalamus		12	−10	18	7.77	7.08
Left mid-occipital		−38	−72	10	7.49	6.86
Right lingual		20	−74	4	7.16	6.60
Right DLPFC		32	8	26	6.99	6.46
Right frontal mid-orbital		34	52	−6	4.87	4.67
Left putamen	<0.001	−18	17	−6	4.25	4.11
Left M1	<0.001	−42	−2	52	6.40	5.99
Left DLPFC		−26	6	64	4.94	4.73
Right IFGpo	<0.001	34	26	−10	6.06	5.70
Right caudate		10	20	−8	4.27	4.13
Linear effect model thresholds						
ACC	<0.001	−6	28	22	5.29	5.04
Left DLPFC (BA46)	<0.001	−32	46	18	4.69	4.51
Dorsal vmPFC (BA10)	0.037	−8	56	2	4.50	4.34

Summary results for all second-level contrasts, reported whole-brain corrected at the cluster level, FWE $p \leq 0.05$. IFGpo, inferior frontal gyrus pars opercularis; IFGpt, inferior frontal gyrus pars triangularis; M1, primary motor cortex; V1, primary visual cortex.

Dynamic versus fixed control

Subjects dynamically adjusted their responses when delayed consequences fluctuated within a trial. These consequences depended on both the number of offers already seen and the number previously accepted/rejected in a trial. Figure 2 illustrates that subjects used both these components to adjust their responses. We quantified this effect by comparing models accounting for the number of previous offers, number of previous accepts, or both (see Materials and Methods). We found strong evidence that the integrated sliding model, wherein both components contribute to choice, fitted subject data best at the group level (lowest iBIC score; Fig. 1C). Although the sliding offer



displayed at $p < 0.001$ uncorrected

Figure 3. Distinct but overlapping frontoparietal networks are recruited when action constraints increase and when the expected long-term value of an option increases. **A**, A frontoparietal network spanning right DLPFC and bilateral parietal cortex was more active in HC compared with LC trials during offers subject to go/no-go. The black arrows indicate two DLPFC clusters that were combined to form a DLPFC f-ROI responding to HC > LC. **B**, Model thresholds, denoting the long-term component of expected value, correlated negatively with BOLD in an overlapping fronto-subcortical-parietal network, including ACC, bilateral DLPFC, parietal cortex, and striatum. Activity in these regions was highest when the value of conserving a unit of budget (rejecting) was low. **C**, Subjects with greater right DLPFC recruitment (see **A**, black arrows, for DLPFC f-ROI) in HC compared with LC showed a larger adjustment in willingness to accept value 2 offers between conditions ($r^2 = 0.33, p = 0.007$). Each point represents one participant.

model performs well (in which only the number of offers seen is used to adjust choice), an addition of tracking the number of accepts/rejects improved the maximum likelihood across every subject (Wilcoxon's signed rank test, $p = 5.96 \times 10^{-5}$). Consistent with the notion that subjects used a dynamic control strategy, reaction times were faster when action (model) thresholds from the winning model were low (mean β HC = 192.5, $p < 0.0001$; mean β LC = 320.1, $p < 0.0001$), controlling for the immediate (face) value of the current offer (mean β HC = -107.3 , $p < 0.0001$; mean β LC = -92.0 , $p < 0.0001$).

fMRI neuroimaging

As in other control paradigms (Kerns et al., 2004; Barber and Carter, 2005), we first performed a categorical comparison to identify brain regions more active when the overall demand for control is increased (HC > LC), averaging across offer values (see Materials and Methods, GLM; Table 1). We found greater whole-brain corrected activity in right DLPFC and bilateral superior parietal lobule in HC overall compared with LC (Fig. 3A). These regions are associated with model-based planning (Owen, 1997; van den Heuvel et al., 2003; Wunderlich et al., 2012), task switching and cognitive control (Botvinick et al., 2001; Liston et al., 2006; Badre, 2008), the resolution of uncertainty (Yoshida and Ishii, 2006), and working memory (WM; Curtis and D'Esposito, 2003; Narayanan et al., 2005; Barbey et al., 2012).

We next hypothesized that greater right DLPFC recruitment in HC compared with LC would result in a larger behavioral adjustment between conditions. We focused on value 2 offers for which we observed the largest change in behavior between HC and LC. We derived an average parameter estimate for an HC > LC contrast in a right DLPFC functional ROI, combining two activated right DLPFC clusters (1078 total voxels; Fig. 3A, middle, black arrows), averaging the β values for our four value regressors and then subtracting LC from HC. A between-subject correlation revealed a positive association between parameter estimates in right DLPFC for an HC > LC contrast and the change in propensity to accept value 2 offers between HC and LC ($r^2 =$

0.33, $p < 0.007$; Fig. 3C). Thus, right DLPFC is instrumental in the categorical adjustment of action control in our task.

To identify correlates of value for guiding choice, we tested for a positive average linear effect of offer (face) value across both HC and LC conditions, revealing a value-dependent response in regions that included vmPFC and VS (including NA; Fig. 4A; for all regions, see Table 1). Importantly, this value signal was independent of any motor response because go responses were modeled as separate onsets in our GLM. Thus, offer values were tracked in regions involved in value representation (Schultz, 2000). Furthermore, because participants' choices were sensitive to action constraint, we anticipated that the representation of offer value would be modulated in one or more regions accordingly. We tested for a value \times constraint interaction (LC more linear than HC) but did not detect any voxels that survived whole-brain correction. Nonetheless, for exploratory purposes, we conjectured that vmPFC and VS, both widely implicated in value-based choice (Guitart-Masip et al., 2012; Hunt et al., 2012; De Martino et al., 2013), might demonstrate an interaction when using a less stringent ROI approach. We derived f-ROIs (see Materials and Methods) by defining voxels (within whole-brain corrected clusters) in vmPFC (928 voxels; Fig. 4B) and VS (56 voxels; Fig. 4C) that showed a linear effect of offer value on average (as above) and then tested for an orthogonal value \times condition (HC or LC) interaction. We found a significant interaction in vmPFC ($F_{(2,38,47,62)} = 5.34$, MSE = 1.67, $p = 0.005$) but not in VS (Fig. 4B). In LC, vmPFC was more responsive to value 2 than value 1 ($p = 0.02$) and value 3 than value 2 ($p = 0.02$), whereas in HC, neither value 2 ($p = 0.35$) nor value 3 ($p = 0.38$) induced greater BOLD than value 1.

Given behavioral and computational evidence that subjects used trial structure to evaluate options, we conjectured that within-trial adaptive choice would manifest as a dynamic modulation of value representations in vmPFC, analogous to that observed between HC and LC trials. To test this, we constructed a summary measure reflecting a time-varying decision threshold, as prescribed by the winning model, that then provided an offer-

wise parametric regressor (see Materials and Methods). In effect, this MT represented the value of carrying one more unit of budget (the number of accepts endowed for a trial) into the next offer, independent of the immediate value of the current offer. The overall value of accepting was thus the difference between offer value and MT. Note, however, in contrast to the downregulation of value 2 offers in HC, the time-variant adaptation in choice prescribed by the winning model require an upregulation of low-value offers when the future benefit of conserving a unit of budget is low.

We first tested for regions in which BOLD signal correlated with MTs across both conditions (see Materials and Methods, GLM) finding that a fronto-subcortical–parietal network was modulated negatively, with no regions modulated positively. This is consistent with BOLD signal being highest when the expected utility of carrying a unit of budget forward was low, and thus a go response was more favorable. This network, which includes ACC, bilateral DLPFC, parietal cortex, and striatum (Fig. 3B; for all regions, see Table 1), is partially overlapping with that seen in the contrast of HC > LC (Fig. 3A), implying that similar regions of PFC are recruited when action control is reliant on internal valuations versus external cues. We note that similar networks are engaged during WM (Curtis and D’Esposito, 2003; Barbey et al., 2012) and in goal-directed and/or cognitive control paradigms (Yoshida and Ishii, 2006; Badre, 2008; Hare et al., 2009; Rushworth et al., 2011).

In our task, the immediate reward gained from accepting value 3 or 4 offers is higher than the maximum MT value, and thus these offers should always be accepted. In contrast, the difference between the immediate reward obtainable from value 1 and 2 offers and their corresponding MTs fluctuates ~ 0 , signifying that choice policy, consistent with the observed behavior, should shift in response to trial state. Consequently, we hypothesized that an independent network tracked MTs differentially dependent on offered value. To test this, we looked for brain regions showing a linearly increasing effect of MTs across both conditions. Because MTs were tracked negatively, this tested an hypothesis that they would correlate more strongly with BOLD as offer value decreased. We found clusters in ACC, left DLPFC (BA46), and a dorsal region of vmPFC (BA10) (Fig. 5A; for details, see Table 1) that were increasingly more responsive to changes in MTs as offered value decreased. The ACC cluster was particularly striking, with *post hoc* exploratory one-sample *t* tests revealing MT representations solely for offers requiring adaptive choice, that is offer value 1 for both conditions (HC, $p = 0.002$; LC, $p = 0.01$) and a trend for offer value 2 for HC alone ($p = 0.09$; Fig. 5B).

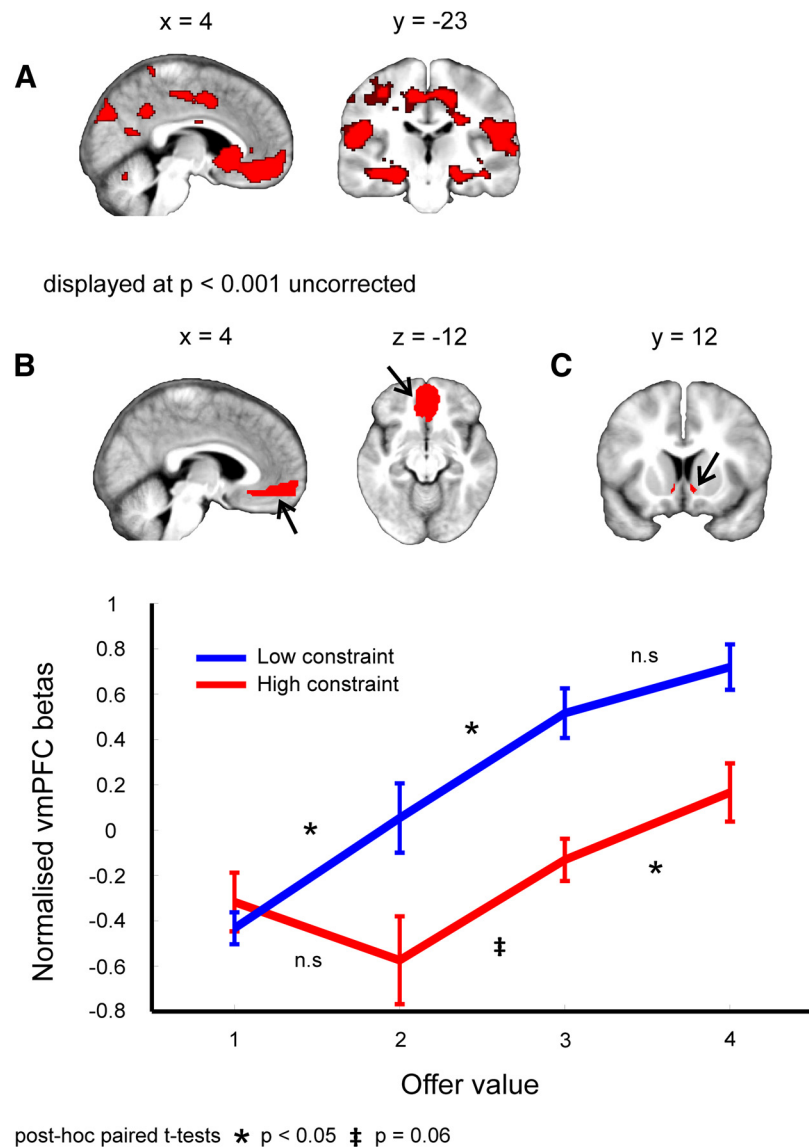


Figure 4. Value representations modulated by context. *A*, The BOLD signal in vmPFC, VS, right amygdala, and precuneus/posterior cingulate covaries with offer value. *B*, vmPFC tracks value linearly in LC but with a depressed slope for HC. The representation of value 2 offers is particularly degraded, mirroring behavioral data. Vertical lines represent SEM. *C*, An f-ROI confined to the VS was used in a constraint (HC/LC) \times value (1, 2, 3, or 4) interaction analysis.

Note that we found behavioral evidence of adaptive choice corresponding to these three offers (Fig. 2).

Finally, we used a connectivity analysis to ask whether brain regions tracking MTs for offers requiring policy switches were modulating value representations in vmPFC to instigate adaptive switches in choice. We selected physiological responses from three f-ROI seed regions, showing a linear effect of MTs (reflecting the long-term component of value), that included ACC (739 voxels in group-level ROIs), left DLPFC/BA46 (502 voxels in group-level ROIs), and dorsal vmPFC/BA10 (179 voxels in group-level ROIs) (Fig. 5A, black arrows). Interestingly, the PFC has been implicated previously in flexible action control, and, in the case of DLPFC, top-down modulation of value signals (Walton et al., 2007; Hare et al., 2009). We performed a PPI to test a hypothesis that coupling would be modulated by fluctuations in MTs and that this change would be greater for low-value offers requiring adaptive choice (values 1 and 2 in HC, and value 1 in LC) than for high-value offers (in which choice is not dependent

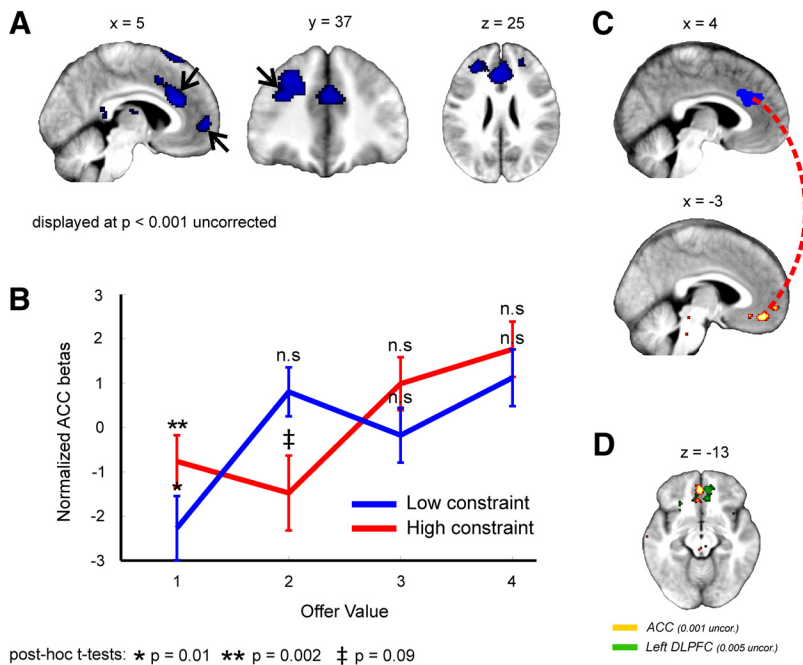


Figure 5. MTs selectively tracked in a prefrontal network. **A**, BOLD signal in ACC, left DLPFC (BA46), and dorsal vmPFC (BA10) increases as MTs decrease (and action is most favorable), only for offers mandating adaptive control. **B**, Parameter estimates from the ACC cluster shown in **A** illustrate model thresholds are tracked for offers requiring adaptive control (value 1 in HC and LC, and a trend for value 2 in HC). Red corresponds to HC and blue to LC. Vertical lines represent SEM. **C**, A whole-brain voxel-based gPPI analysis revealed that ACC is more functionally connected with the vmPFC when actions cost are high and low offers should be rejected. This region of vmPFC overlaps with a cluster that tracks offer value (Fig. 4A) and is sensitive to categorical changes in context (Fig. 4B). **D**, Comparison of functional connectivity patterns between ACC (yellow; displayed at 0.001 uncorrected) or left DLPFC/BA46 (green; displayed at 0.005 uncorrected) and vmPFC. As with ACC, the left DLPFC demonstrates a functional coupling with vmPFC when accepting an option offering only a small immediate reward is unfavorable, but this effect only emerges at a more liberal threshold.

on MT). The regions identified by the ensuing PPI correspond to regions whose connectivity with the relevant seed region depends on both the immediate value and MT of the current offer.

We found a functional coupling between ACC and vmPFC that was sensitive to fluctuations in MTs that was larger on average for offers requiring adaptive choice. This effect was significant when using small volume correction for the vmPFC f-ROI that tracked offer value. Given that directionality cannot be determined when comparing parametric effects across conditions, we performed a second PPI analysis, now confined to offers requiring adaptive choice, enabling us to assess whether connectivity was positively or negatively modulated by increasing MTs. A vmPFC f-ROI approach revealed that ACC and vmPFC were more functionally coupled when MTs were high (mean PPI = 3.04, $p = 0.005$), in other words, when low-value offers need to be rejected. Thus, connectivity between ACC and vmPFC was dependent on both immediate value and MT. Although the left DLPFC did not demonstrate functional coupling with vmPFC that depended on both MT and offer value, qualitatively we observed an effect in vmPFC at a more liberal threshold ($p = 0.005$ uncorrected). In fact, we did not detect any significant difference in the magnitude of the PPI effect (two-sample t test, $p = 0.68$) between ACC and DLPFC when using a vmPFC f-ROI, implying that, despite a more prominent contribution of ACC, DLPFC also contributes to the observed connectivity. When dorsal vmPFC was used as a seed, no significant results were observed.

Discussion

Our study addressed the computational implementation of context-specific action control in value-guided choice. We show

that subjects incorporate both extrinsic constraints on action and intrinsic fluctuations in opportunity to adaptively switch between a go/no-go response. Mechanistically, a fronto-subcortical-parietal network tracks the downstream consequence of spending a limited action budget, whereas ACC couples to vmPFC to shift the representation of value in favor of long-term profit.

In our task, subjects track the number of offers already seen and number already accepted/rejected in a trial to compute the future value of expending a unit of budget. This model fits behavior better than simpler candidates in which action is driven solely by immediate reward or when only a restricted set of environmental features is consequential. Of interest, the winning model produces behavior that closely approximates optimal choice, which relies on backpropagating through a decision tree of all future moves in a trial. Although this strategy is computationally taxing (given the depth of the search tree in this game), subjects could be computing long-term value by recruiting a model-based system that searches through future states “on the fly” (Dayan, 2008). Alternatively, a player could track aspects of the environment to index stored values or to update values under a model-free regimen. Although our task cannot arbitrate between these possibilities,

we note the circuitry that tracks the MTs from our winning model overlaps with that implicated in model-based reinforcement learning (Gläscher et al., 2010; Daw et al., 2011; Wunderlich et al., 2012).

Influential accounts of ACC propose a myriad of roles, including conflict monitoring (Botvinick, 2007), error monitoring (Rushworth et al., 2004), overriding prepotent responses (Kerns et al., 2004), evaluating outcomes (Gehring and Willoughby, 2002), and action-outcome learning for negative feedback (Rushworth et al., 2004). Although our task lacked explicit negative feedback, the finding that ACC tracks the MTs necessary for implementing adaptive choice is consistent with the conflict monitoring account but not with a role in error monitoring, given that choices were closely aligned with optimality. Unlike previous paradigms in which switches in contingency are explicitly cued (Kerns et al., 2004), we show that conflict in ACC can arise endogenously via tracking fluctuations in downstream consequence.

ACC is also implicated in foraging (Kolling et al., 2012) in which it is proposed to track the value of alternative choice options during a tradeoff between exploration and exploitation. We found that ACC activity was highest when exploiting a low-value offer was more optimal. However, in our task, ACC only tracks MTs corresponding to offers that are routinely rejected. In this light, our findings can be construed as in keeping with the former role. These findings hint that a conflict monitoring account of ACC can be reinterpreted as reflecting a need to switch behavior from the current default response, as opposed to encoding a non-specific conflict signal (Shenhav et al., 2013). Indeed, recent work

further supports the notion that ACC assumes a default frame of reference by adapting choice from the best long-running option (Boorman et al., 2013).

A number of studies propose that ACC expresses a prediction error (Ide et al., 2013) that can be used to update internally generated models (O'Reilly et al., 2013). This may explain why high-conflict or high-volatility trials, often confounded with surprise, also induce responses in ACC. However, our data indicate that surprise cannot fully account for the ACC activation we observe, because stimuli are presented with equal frequency such that surprise does not vary within a trial. Instead, a response to low-value offers switches in line with changes in delayed consequence. Thus, in the context of the current study, it is likely that ACC plays a more general role in a strategic adjustment of behavior that is rooted in processing or initiating atypical stimulus or action requirements, which also includes surprising events.

A dynamic coupling between ACC and vmPFC was seen when MTs dictate action costs are high, with the greatest change in coupling evident in offers in which action requirement is most dependent on MT. One interpretation is that ACC suppresses the representation of low-value offers in vmPFC when the future value of conserving a unit of budget is high and the optimal decision is to reject. Conversely, when MTs are low, decoupling between ACC and vmPFC may reflect a disinhibition of value signals relating to previously unfavorable offers. This contrasts with other suggestions that ACC signals a need for control but plays no causal role in conflict resolution (Kerns et al., 2004) or that dissociable decision variables are computed in vmPFC and ACC that compete for behavioral output (Boorman et al., 2013). Because ACC activity in our task is not sensitive to changes in MTs corresponding to high-value offers, it is unlikely to represent an unrelated correlate of trial time or WM content.

In contrast to the selectivity implemented by ACC, we found that MTs were tracked indiscriminately within an extensive fronto-subcortical–parietal network. Although planned choice has only been studied recently in a value domain, a finding that this network tracks computations related to future value is consistent with previous work from the model-based reinforcement learning literature (Daw et al., 2005; Gläscher et al., 2010; Wunderlich et al., 2012). Interestingly, recent evidence suggests that PFC neurons can adapt their tuning profiles to accommodate changes in behavioral context (Stokes et al., 2013), a mechanism that could underlie a network-level implementation of the adaptive responses observed in our task. We note that this frontoparietal network also encompasses regions implicated in executive control (Wallis and Miller, 2003; Barber and Carter, 2005; Hare et al., 2009), exploratory behavior (Daw et al., 2006; Yoshida and Ishii, 2006), intertemporal choice (McClure et al., 2004), and WM (Curtis and D'Esposito, 2003; Barbey et al., 2012).

One limitation of our task is that it cannot characterize a neural correlate of the fully integrated value derived from our computational model (the difference between the current offer and the associated MT) because this is correlated with the immediate value of the offer. However, the observed fronto-subcortical–parietal activity may reflect a value comparison between offer value and MT. As MTs decrease, the difference in value between go and no-go shifts in favor of a go response, whereas when MTs increase, they approach the average worth of the offer value range (2.5), making the decision to accept or reject harder. Alternatively, given that MTs trended downward as trials progressed (although not exclusively because they are also a function of the current budget), they are anticorrelated with WM demand; following the contents of trial history become harder to maintain

(and update) through time. Because we found that activity in this fronto-subcortical–parietal network tracked MTs across all offers, this profile may reflect a WM signature. Interestingly, it has been shown that goal-directed choice is dependent on WM (Otto et al., 2013). In this regard, there is considerable debate as to whether delay-period DLPFC activity, classically interpreted as a correlate of WM, reflects the pure maintenance of information, or instead whether WM is merely an emergent property of executive and attentional functions implemented in DLPFC (Postle, 2006).

Our paradigm also incorporated HC and LC environments, and, in the former, subjects reject lower-value options to increase the probability of capitalizing from larger later rewards. We found categorically switching from LC to HC correlated with the fMRI signal in a similar frontoparietal network. Within this network, the more DLPFC was recruited in HC compared with LC, the more a subject would modulate their behavioral response to value 2 offers between conditions. In addition, we found widespread correlates of offer value in regions linked previously to value computations, including vmPFC (Hare et al., 2009), VS (Guitart-Masip et al., 2012) and posterior cingulate/precuneus (Litt et al., 2011). Importantly, value representations were altered in HC in vmPFC, a key value-coding region.

Interestingly, a comparable frontoparietal network is reliably upregulated in conditions requiring cognitive control or overcoming response conflict in task-switching paradigms (Kerns et al., 2004; Badre, 2008; Pochon et al., 2008; Mansouri et al., 2009). This likeness suggests that participants may be engaging cognitive control mechanisms to appropriately reject appetitive, though relatively less valuable, offers in light of increasing environmental demands in HC trials. In this framework, our data corroborate previous ideas of interplay between PFC and value regions, suggestive of a scheme whereby value signals are modulated directly to achieve adaptive choice (Hare et al., 2009; Diekhof and Gruber, 2010). However, as with previous control paradigms, we note that a categorical difference in activity profiles between conditions does not pose any properties that allow attribution of specific computational roles.

References

- Badre D (2008) Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends Cogn Sci* 12:193–200. [CrossRef Medline](#)
- Baliki MN, Mansour A, Baria AT, Huang L, Berger SE, Fields HL, Apkarian AV (2013) Parceling human accumbens into putative core and shell dissociates encoding of values for reward and pain. *J Neurosci* 33:16383–16393. [CrossRef Medline](#)
- Balleine BW, Dickinson A (1998) Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* 37:407–419. [CrossRef Medline](#)
- Barber AD, Carter CS (2005) Cognitive control involved in overcoming prepotent response tendencies and switching between tasks. *Cereb Cortex* 15:899–912. [CrossRef Medline](#)
- Barbey AK, Koenigs M, Grafman J (2013) Dorsolateral prefrontal contributions to human working memory. *Cortex* 49:1195–1205. [CrossRef Medline](#)
- 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](#)
- Boorman ED, Rushworth MF, Behrens TE (2013) Ventromedial prefrontal and anterior cingulate cortex adopt choice and default reference frames during sequential multi-alternative choice. *J Neurosci* 33:2242–2253. [CrossRef Medline](#)
- Botvinick MM (2007) Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cogn Affect Behav Neurosci* 7:356–366. [CrossRef Medline](#)
- Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD (2001) Conflict

- monitoring and cognitive control. *Psychol Rev* 108:624–652. [CrossRef Medline](#)
- Buschman TJ, Denovellis EL, Diogo C, Bullock D, Miller EK (2012) Synchronous oscillatory neural ensembles for rules in the prefrontal cortex. *Neuron* 76:838–846. [CrossRef Medline](#)
- Curtis CE, D'Esposito M (2003) Persistent activity in the prefrontal cortex during working memory. *Trends Cogn Sci* 7:415–423. [CrossRef Medline](#)
- Daw ND, Niv Y, Dayan P (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci* 8:1704–1711. [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](#)
- Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ (2011) Model-based influences on humans' choices and striatal prediction errors. *Neuron* 69:1204–1215. [CrossRef Medline](#)
- Dayan P (2008) The role of value systems in decision making. In: *Better than conscious? Decision making, the human mind, and implications for institutions* (C. Engel, S. Wolf, eds), pp 51–70. Cambridge, MA: Massachusetts Institute of Technology.
- De Martino B, Fleming SM, Garrett N, Dolan RJ (2013) Confidence in value-based choice. *Nat Neurosci* 16:105–110. [CrossRef Medline](#)
- Diekhof EK, Gruber O (2010) When desire collides with reason: functional interactions between anteroventral prefrontal cortex and nucleus accumbens underlie the human ability to resist impulsive desires. *J Neurosci* 30:1488–1493. [CrossRef Medline](#)
- Gehring WJ, Willoughby AR (2002) The medial frontal cortex and the rapid processing of monetary gains and losses. *Science* 295:2279–2282. [CrossRef Medline](#)
- Gläscher J, Daw N, Dayan P, O'Doherty JP (2010) States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron* 66:585–595. [CrossRef Medline](#)
- Guitart-Masip M, Huys QJ, Fuentemilla L, Dayan P, Duzel E, Dolan RJ (2012) Go and no-go learning in reward and punishment: interactions between affect and effect. *Neuroimage* 62:154–166. [CrossRef Medline](#)
- Hare TA, Camerer CF, Rangel A (2009) Self-control in decision-making involves modulation of the vmPFC valuation system. *Science* 324:646–648. [CrossRef Medline](#)
- Hare TA, Malmaud J, Rangel A (2011) Focusing attention on the health aspects of foods changes value signals in vmPFC and improves dietary choice. *J Neurosci* 31:11077–11087. [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, S1–3. [CrossRef Medline](#)
- Hutton C, Josephs O, Stadler J, Featherstone E, Reid A, Speck O, Bernarding J, Weiskopf N (2011) The impact of physiological noise correction on fMRI at 7 T. *Neuroimage* 57:101–112. [CrossRef Medline](#)
- Huys QJ, Cools R, Gölzer M, Friedel E, Heinz A, Dolan RJ, Dayan P (2011) Disentangling the roles of approach, activation and valence in instrumental and pavlovian responding. *PLoS Comput Biol* 7:e1002028. [CrossRef Medline](#)
- Ide JS, Shenoy P, Yu AJ, Li CS (2013) Bayesian prediction and evaluation in the anterior cingulate cortex. *J Neurosci* 33:2039–2047. [CrossRef Medline](#)
- Kable JW, Glimcher PW (2007) The neural correlates of subjective value during intertemporal choice. *Nat Neurosci* 10:1625–1633. [CrossRef Medline](#)
- Kass R, Raftery A (1995) Bayes factors. *J Am Stat Assoc* 90:773–795.
- Kerns JG, Cohen JD, MacDonald AW 3rd, Cho RY, Stenger VA, Carter CS (2004) Anterior cingulate conflict monitoring and adjustments in control. *Science* 303:1023–1026. [CrossRef Medline](#)
- Kolling N, Behrens TE, Mars RB, Rushworth MF (2012) Neural mechanisms of foraging. *Science* 336:95–98. [CrossRef Medline](#)
- Liston C, Matalon S, Hare TA, Davidson MC, Casey BJ (2006) Anterior cingulate and posterior parietal cortices are sensitive to dissociable forms of conflict in a task-switching paradigm. *Neuron* 50:643–653. [CrossRef Medline](#)
- Litt A, Plassmann H, Shiv B, Rangel A (2011) Dissociating valuation and saliency signals during decision-making. *Cereb Cortex* 21:95–102. [CrossRef Medline](#)
- Mansouri FA, Tanaka K, Buckley MJ (2009) Conflict-induced behavioural adjustment: a clue to the executive functions of the prefrontal cortex. *Nat Rev Neurosci* 10:141–152. [CrossRef Medline](#)
- McClure SM, Laibson DI, Loewenstein G, Cohen JD (2004) Separate neural systems value immediate and delayed monetary rewards. *Science* 306:503–507. [CrossRef Medline](#)
- McLaren DG, Ries ML, Xu G, Johnson SC (2012) A generalized form of context-dependent psychophysiological interactions (gPPI): a comparison to standard approaches. *Neuroimage* 61:1277–1286. [CrossRef Medline](#)
- Narayanan NS, Prabhakaran V, Bunge SA, Christoff K, Fine EM, Gabrieli JD (2005) The role of the prefrontal cortex in the maintenance of verbal working memory: an event-related fMRI analysis. *Neuropsychology* 19:223–232. [CrossRef Medline](#)
- O'Reilly JX, Schüffelgen U, Cuell SF, Behrens TE, Mars RB, Rushworth MF (2013) Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proc Natl Acad Sci U S A* 110:E3660–E3669. [CrossRef Medline](#)
- Otto AR, Gershman SJ, Markman AB, Daw ND (2013) The curse of planning: dissecting multiple reinforcement-learning systems by taxing the central executive. *Psychol Sci* 24:751–761. [CrossRef Medline](#)
- Owen AM (1997) Cognitive planning in humans: neuropsychological, neuroanatomical and neuropharmacological perspectives. *Prog Neurobiol* 53:431–450. [CrossRef Medline](#)
- Plassmann H, O'Doherty JP, Rangel A (2010) Appetitive and aversive goal values are encoded in the medial orbitofrontal cortex at the time of decision making. *J Neurosci* 30:10799–10808. [CrossRef Medline](#)
- Pochon JB, Riis J, Sanfey AG, Nystrom LE, Cohen JD (2008) Functional imaging of decision conflict. *J Neurosci* 28:3468–3473. [CrossRef Medline](#)
- Postle BR (2006) Working memory as an emergent property of the mind and brain. *Neuroscience* 139:23–38. [CrossRef Medline](#)
- Rangel A, Camerer C, Montague PR (2008) A framework for studying the neurobiology of value-based decision making. *Nat Rev Neurosci* 9:545–556. [CrossRef Medline](#)
- Rangel A, Hare T (2010) Neural computations associated with goal-directed choice. *Curr Opin Neurobiol* 20:262–270. [CrossRef Medline](#)
- Rushworth MF, Walton ME, Kennerley SW, Bannerman DM (2004) Action sets and decisions in the medial frontal cortex. *Trends Cogn Sci* 8:410–417. [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](#)
- Schultz W (2000) Multiple reward signals in the brain. *Nat Rev Neurosci* 1:199–207. [CrossRef Medline](#)
- Shenhav A, Botvinick MM, Cohen JD (2013) The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron* 79:217–240. [CrossRef Medline](#)
- Stokes MG, Kusunoki M, Sigala N, Nili H, Gaffan D, Duncan J (2013) Dynamic coding for cognitive control in prefrontal cortex. *Neuron* 78:364–375. [CrossRef Medline](#)
- van den Heuvel OA, Groenewegen HJ, Barkhof F, Lazeron RH, van Dyck R, Veltman DJ (2003) Frontostriatal system in planning complexity: a parametric functional magnetic resonance version of Tower of London task. *Neuroimage* 18:367–374. [CrossRef Medline](#)
- Wallis JD, Miller EK (2003) Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur J Neurosci* 18:2069–2081. [CrossRef Medline](#)
- Walton ME, Croxson PL, Behrens TE, Kennerley SW, Rushworth MF (2007) Adaptive decision making and value in the anterior cingulate cortex. *Neuroimage* 36 [Suppl 2]:T142–154. [CrossRef](#)
- Wunderlich K, Dayan P, Dolan RJ (2012) Mapping value based planning and extensively trained choice in the human brain. *Nat Neurosci* 15:786–791. [CrossRef Medline](#)
- Yoshida W, Ishii S (2006) Resolution of uncertainty in prefrontal cortex. *Neuron* 50:781–789. [CrossRef Medline](#)