

The Prefrontal Cortex in Reward-Based Decision Making

Aurélien Marchand and Adrien Lefebvre

Department of Cognitive Neuroscience, University of Lyon, 69007 Lyon, France, 2Neuroimaging Laboratory, INSERM U1219, 69008 Lyon, France

Citation: Aurélien Marchand and Adrien Lefebvre (2025). The Prefrontal Cortex in Reward-Based Decision Making. *Psychiatra*, 17(9), 72-106. <https://doi.org/10.5281/zenodo.19095189>

Abstract

The human prefrontal cortex (PFC) subserves cognitive control, that is, the ability to form behavioral strategies that coordinate actions and thoughts in relation to internal goals. Cognitive control involves the medial and lateral PFC but we still poorly understand how these regions guide strategy selection according to expected rewards. We addressed this issue using neuroimaging, computational modeling and model-based analyses of information flows between medial and lateral PFC. We show here that the (dorsal) medial PFC encodes and conveys to lateral PFC reward expectations driving strategy selection, while strategy selection originates in lateral PFC and propagates backward to medial PFC. This functional loop through lateral PFC enables strategy selection to further comply with learned rules encoded in lateral PFC rather than with reward expectations conveyed from medial PFC. Thus, the medial and lateral PFC are functionally coupled in cognitive control for integrating expected rewards and learned rules in strategy selection.

Key words: executive control, fMRI, human, prefrontal cortex, reinforcement learning

Introduction

The human prefrontal cortex (PFC) subserves cognitive control, that is, the ability to coordinate actions and thoughts in relation to internal goals (Rogers and Monsell 1995; Miller and Cohen 2001; Koechlin et al. 2003; Ridderinkhof et al. 2004). Cognitive control is a cornerstone of human higher cognition and especially enables the agent to switch between distinct behavioral strategies. The PFC subserves cognitive control by arbitrating between “task sets,” that is, consistent sets of internal representations specifying how to act in a given situation (Rogers and Monsell 1995; Koechlin et al. 2003; Sakai 2008; Collins and Frank 2013; Koechlin 2014). The lateral PFC is involved in selecting task sets according to learned rules through top-down selection processes operating from lateral PFC to premotor regions (Koechlin et al. 2003; Badre and D’Esposito 2007; Koechlin and Summerfield 2007; Badre 2008; Badre et al. 2009; Azuar et al. 2014; Nee and D’Esposito 2016). In agreement with the

review in Rushworth et al. 2011). The ACC thus has been proposed to predict action outcomes in the service of task-set selection (Alexander and Brown 2011). More specifically, several studies suggest that at decision time, the ACC contributes to cognitive control by encoding the reward advantage (referred to as the foraging value) to switch away rather than to stay with the default/ongoing task set, irrespective of actual choices (Kolling et al. 2012, 2014, 2016; Boorman et al. 2013). However, medial PFC activations in these studies have been argued to actually reflect “response difficulty” (measured by reaction times) and/or “choice ambiguity” (i.e., the proximity between control signals guiding task-set selection) rather than foraging values (Neta et al. 2014; Shenhav et al. 2014), supporting the idea that the medial PFC monitors choice ambiguity and selects task sets based on action outcomes (Shenhav et al. 2013, 2014). Yet, signals reflecting task-set selection based on outcome expectations

massive projections between lateral and medial PFC (Pandya and Yeterian

are found in multiple regions in both the medial and lateral PFC (Hampton et

© The Author 2017. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

For commercial re-use, please contact journals.permissions@oup.com

1996; Johansen-Berg et al. 2004; Beckmann et al. 2009; Medalla and Barbas 2009), furthermore, medial PFC regions and especially the presupplementary motor area (pre-SMA) are involved in switching between task sets with respect to learned rules presumably by inhibiting inappropriate behavioral responses (review in Aron et al. 2007; Nachev et al. 2008). The pre-SMA along with the dorsal anterior cingulate cortex (ACC) also regulate the engagement of lateral PFC in rule-based cognitive control (Gehring and Knight 2000; Botvinick et al. 2001; Holroyd and Coles 2002; Hyafil et al. 2009), notably according to rewards at stake (Kouneiher et al. 2009).

However, we still poorly understand how the PFC drives task-set selection according to reward expectations rather than learned rules. Previous studies suggest that while the ventromedial PFC is involved in encoding the rewarding values of action outcomes, the medial PFC including the ACC is involved in linking expected rewards to actions (Rudebeck et al. 2008; Camille et al. 2011;

al. 2006; Hampton and O’Doherty 2007). Thus, while both the medial and lateral PFC are clearly functionally distinct, central nodes in cognitive control (Dosenbach et al. 2007; Power and Petersen 2013; Gratton et al. 2017), the specific contribution of medial PFC to cognitive control remains largely unclear and no converging views have emerged yet regarding how the PFC drives task-set selection according to reward expectations.

To clarify this issue, we elaborated a behavioral protocol aiming at dissociating reward expectations, task-set selection, choice ambiguity and response difficulty in cognitive control. For that purpose, the protocol required subjects to repeatedly choose to carry out 1 of 2 behavioral strategies according to either learned rules or expected rewards. Reward expectations built from action outcomes observed in successive trials mixing rule-based and reward-based choices. Using functional magnetic resonance imaging (fMRI), we scanned 23 subjects in this protocol. We used computational modeling to

compare PFC activations and information flows between PFC regions associated with reward-based and rule-based choices of behavioral strategies. The mixture of rule- and reward-based choices allowed us to disentangle regional activations as well as interregional interactions reflecting the encoding of reward expectations, task-set selection processes, choice ambiguity and difficulty.

Materials and Methods

Subjects

Subjects (12 females and 11 males aged 20–31 years and righthanded) had no general medical, neurological, psychiatric, or addictive history as assessed by medical examinations. They provided written informed consent that was approved by the French National Ethics Committee. Subjects were paid for their participation.

Behavioral Protocol

Subjects responded to visually and successively presented letters. For every letter, subjects had to choose between 2 tasks: either performing a vowel/consonant discrimination task using 2 response buttons with left and right forefingers or an upper/lower case discrimination task using 2 other buttons with left and right middle fingers (Fig. 1). Subjects then received a stochastic reward varying from 1 to 9 monetary units (Gaussian-like distributed, SD = 2.5 monetary unit) for correct performance. One task was on average more rewarded than the other one (T+ = 6 vs. T- = 4 m.u.), but this advantage was systematically reversed after an unpredictable number of trials (from 16 to 32 trials). Critically, letters appeared in various colors with no influences on rewards except when letters appeared in green, subjects had to perform one task (e.g., the vowel/consonant task) to receive the reward and when letters were red, they had to perform the other task (e.g., the upper/lower case task) to receive the reward. We thus refer to the green and red letter trials as “rule-based” trials and to the others as “rule-free” trials. This design intermixing rule-based and rule-free trials permitted reward expectations and receptions to be matched in the 2 trial types. As rule-based trials randomly cued the 2 tasks, furthermore, subjects were induced to constantly switch between the most and least rewarded tasks (T+ and T-). This prevented rule-free (and rule-based) choices to turn into repetitive performances and cognitive control to remain engaged in both trial types. We refer to these series of rule-free and rule-based trials as “selection” blocks.

In order to identify brain regions involved in task selection compared with execution, selection blocks were intermixed with 4 “baseline” blocks. Baseline blocks consisted of “repeat”

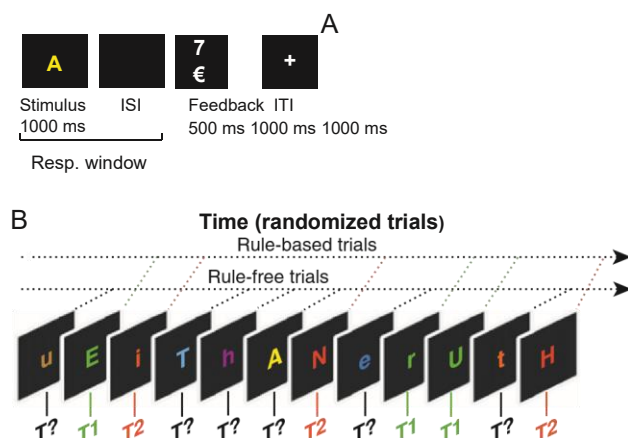
Figure 1. Behavioral protocol. (A) trial structure. In each trial, a colored letter randomly drawn from the set (A, E, I, U, a, e, i, u, H, N, R, T, h, n, r, t) was displayed during 1000 ms and subjects performed either the vowel/consonant (T1) or lower/upper case discrimination task (T2). After correct task performance, a variable monetary feedback was displayed during 1000 ms and indicated the reward subjects received (values ranging from 1 to 9 monetary units). Incorrect task performance or lacks of responses led to no rewards (a zero-feedback was displayed). Reward values were gaussian-like distributed (SD = 2.5). One task was on average better rewarded (referred to as T+ vs. T-, mean = 6 vs. 4 m.u., respectively). This advantage was reversed after an unpredictable number of trials (ranging from 16 to 32 trials). Interstimulus interval (ISI) from letter offsets to feedback onsets: 500 ms. Intertrial interval (ITI) from feedback offsets to next stimulus onsets: 1000 ms. (B) Rule-based trials consisted of letters presented in either green or red, with green instructing subjects to perform T1 and red to perform T2 (color-task assignments were counterbalanced across subjects). Departing from these instructions led to reward zero (a zero-feedback was displayed). Rule-free trials consisted of letters in colors randomly drawn from a set of 10 colors, except green and red. These color cues were uninformative. Rule-free and rule-based trials were in equal proportion and pseudorandomly intermixed independently of reward contingencies. Similarly, green and red cues were in equal proportion, pseudorandomly intermixed independently of reward contingencies. Catch trials were further inserted for appropriately separating fMRI bold responses to stimuli versus feedbacks and to rule-free versus rule-based trials (see Materials and Methods).

trials: visual instructions (duration: 5000 ms) preceding each baseline block instructed subjects to repeat one task over this block. Half of baseline blocks were assigned to one task, the other half to the other task. Task execution was rewarded as in selection blocks with no reversals: the low and high reward distribution was used for each repeated task but in distinct baseline blocks. Visual instructions terminated baseline blocks and informed subjects about selection block onsets. In order to remove sustained effects in comparing selection and baseline blocks as well as to appropriately separate fMRI bold responses to rule-free and rule-based trials, we also included additional “null” trials in both selection and baseline blocks (one-third of randomly intermixed trials). In null trials, stimuli were letters G and D presented in italic and gray color and subjects responded to letter G by pressing the 2 left buttons (G stands for L[eft] in French), and to letter D by pressing the 2 right buttons (D stands for R[ight] in French). Finally, for appropriately separating bold responses at decision and feedback times (Gogghari and MacDonald 2008), the display of monetary feedbacks was masked by a fixation cross on half of trials. Subjects were informed that these “catch” trials were rewarded exactly as the other trials and had no incidences on final pay-offs (the only difference was therefore that the actual reward could not be observed).

The experimental protocol consisted of 2 fMRI sessions administered a couple of days apart. Each session included 2 fMRI scanning runs and each run comprised 4 selection blocks and 4 baseline blocks. In total, each selection block comprised 72 trials on average and each baseline block comprised 36 trials. Order of blocks was arranged for controlling for task-by-reward combinations in baseline blocks, task-by-reward combinations preceding the first reversal in selection blocks, lengths and number of reversals in selection blocks. At the end of each scanning run, the monetary rewards cumulated within each block were averaged over blocks and displayed. Participants were informed that one of these amounts will be randomly drawn at the end of the protocol and added to their final pay-off. Finally, subjects were trained on the protocol during a training session a couple of days before fMRI sessions. The training session consisted of 2 behavioral runs built as scanning runs described above. This behavioral protocol was administered using the Psychtoolbox software package (<http://psychtoolbox.org>).

Computational Modeling and Model Fitting

We tested 2 computational models that describe how reward expectations are formed from feedbacks and drive task selection in rule-free trials. Specifically, both models describe how task values V_A' and V_B' involved in task selection in rule-free trials are updated according to monetary feedbacks r^i received in trial t (no updating occurs in “catch” trials, when no feedbacks were presented).



Reinforcement Learning Model

Values of chosen task V_{C_t} in trial t are updated according to reward prediction errors, that is, the discrepancy between received rewards r_t and values of chosen task V_{C_t} (Rescorla and Wagner 1972). To further account for the anticorrelated and reversal structure of reward contingencies in the protocol, the model assumes that task values of unchosen tasks (denoted V_{U_t}) are updated in the opposite direction according to reward prediction errors associated with chosen tasks (we verified that this RL model fitted better subjects' performances than the model without this assumption). Accordingly, task values are updated as follows:

$$\begin{aligned} V_{C_{t+1}} &= V_{C_t} + \alpha(-r_t V_{C_t}) \\ V_{U_{t+1}} &= V_{U_t} + \alpha(-r_t V_{C_t}) \end{aligned} \quad (1)$$

where α is the learning rate (treated as a free parameter). Equations (1) imply that the sum of task values

$V_{C_t} + V_{U_t} = V_A + V_B$ is an arbitrary constant that we set to 1 (as in the Bayesian model).

Bayesian Model

Task values V_A^t represent subjects' belief that task A is the actual best rewarding task in trial t . Accordingly, V_A^t writes as $V_A^t = (p_{z_t = A} / C_{1:t-1}, r_{1:t-1})$, that is, the inferred probability that current best rewarding task z_t is task A, given all previously chosen tasks $C_{1:t-1}$ and previously experienced rewards $r_{1:t-1}$. We naturally have $V_A^t + V_B^t = 1$. Task values V_A^t are then updated according to standard Bayesian inference:

$$V_A^{t+1} \propto (-) (1 + \tau P(r_{z_t} = A | C_{1:t}, r_{1:t})) V_A^t + \tau P(r_{z_t} = B | C_{1:t}, r_{1:t}) \quad (2)$$

where τ scales the perceived volatility of external contingencies, that is, the probability of reversals between 2 successive trials ($\tau = (p_{z_{t+1} \neq z_t})$ treated as a free parameter) (Behrens et al. 2007). Here, $P(r_{z_t} = A | C_{1:t}, r_{1:t})$ ($P(r_{z_t} = B | C_{1:t}, r_{1:t})$ resp.) is the likelihood of reward r_t , given that task A (task B resp.) is the best rewarding task, given chosen task C_t in trial t . Because subjects were trained on the protocol before the experiment, we assumed that the model encodes true likelihoods, that is, likelihoods were equal to actual reward distributions. This assumption was confirmed by fitting the Bayesian model that relaxes this assumption and optimally learns these likelihood distributions. This Bayesian learning model fit behavioral data less accurately than the Bayesian model based on true feedback likelihoods.

Task Selection in Rule-Free Trials

We assumed that in rule-free trials, choice probabilities $P^t(A)$ of choosing task A rather than B in trial t vary as a logistic function of subjective values $V_A^t - V_B^t$ of task A relative to B:

$$P^t(A) = \frac{1}{1 + \exp(-\beta(V_A^t - V_B^t))} \quad (3)$$

where β is the inverse temperature and ϵ the lapse rate (both β and ϵ are treated as free parameters).

Model Fitting

Both models include the same number of free parameters so that model (log-) likelihoods (LLHs) were used to compare model fits. Model parameters were computed for maximizing LLHs using grid searches associated with gradient

descents from multiple starting points (MATLAB optimization toolbox): for every subject, we computed model-free parameters that maximize the LLH of observing actual subjects' responses in every rule-free trial, given actual subjects' responses and feedbacks in either "all" preceding trials, only rule-free trials or only rule-based trials.

fMRI Data Acquisition and Processing

The experiment was carried out in 2 scanning sessions following the training session, each comprising 2 functional scanning runs and administrated in separated days within a week. MRI data were collected with a 3 T Siemens whole-body and ratio frequency coil scanner. Each functional scanning run included 792 T2 images (time of repetition: 2 000 ms; time of echos: 30 ms; flip angle: 90°; field of view: 192 × 192 mm; acquisition matrix: 64 × 64 × 37 voxels; voxel size: 3 × 3 × 3 mm). A structural T1 scan was also collected (acquisition matrix: 256 × 256 × 176 voxels; voxel size: 1 × 1 × 1 mm) at the end of scanning sessions.

Functional MRI data were processed and analyzed using SPM8 software package (<http://www.fil.ion.ucl.ac.uk>) using standard realignment, normalization to Montreal Neurological Institute echo planar imaging template (images resampled at 4 × 4 × 4 mm) and Gaussian spatial smoothing (isotropic 8-mm kernel). Temporal correlations were estimated using restricted maximum likelihood estimates of variance components using a first-order autoregressive model. The resulting nonsphericity was used to form maximum likelihood estimates of the activations.

Statistical parametric maps of local brain activations were computed in every subject using the standard general linear model (GLM). The model included separate event-related regressors for rule-free and rule-based trials, which convolved a series of delta function with the canonical hemodynamic response function that estimated blood oxygen level-dependent (BOLD) responses at stimulus onsets (decision time). Additional regressors modeled MR bold responses at trial onsets for repeat and null trials, trials with errors in discrimination tasks, no response trials, rule-based trials associated with choices departing from instruction cues (4% of rule-based trials), and instructions at the beginning of each block. Separate regressors modeled BOLD responses at feedback onsets for each trial type, along with the parametric modulation of feedback values. Finally, additional regressors at stimulus and feedback onsets were included to model task values derived from computational modeling as described below.

Univariate Statistical Analyses

Activations Associated with Task Selection at Decision Time We first identified prefrontal regions engaged in task selection rather than execution in rule-free and rule-based trials independently of variations in task values. We estimated a first GLM as defined above that factored out task values by using a binning procedure based on task values (see below). This binning approach was preferred to parametric regressors for factoring out the effects of task values without prior assumptions about the profile of value-based effects. At the group level, we obtained statistical maps of contrasts of parameter estimates, with subjects treated as a random factor (second level, random effect analyses). The voxel-wise significant threshold was set at $P < 0.05$ corrected for family-wise errors for multiple comparisons over search volumes (see below). The cluster-wise threshold was set at 10 voxels (0.64 cm³; $P < 0.005$, uncorrected).

Analysis 1. We identified within the prefrontal cortex the regions associated with task selection compared with execution by computing larger phasic activations at decision times in rule-free and rule-based trials compared with repeat trials. We used 2 interaction contrasts crossing blocks (selection vs. baseline) and trial types (task vs. null trials). The first contrast was (rule-free minus null trials) in selection blocks compared with (repeat minus null trials) in baseline blocks. The second contrast was (rule-based minus null trials) in selection blocks compared with (repeat minus null trials) in baseline blocks (these interaction contrasts include null trials for removing block effects). We then used a conjunction analysis of the 2 resulting interaction maps over the

whole frontal lobes (talairach coordinate $Y > -10$, search volume) to identify prefrontal regions involved in task selection in both rule-based and rule-free trials (yellow and red regions in Fig. 4A).

Analysis 2. Within the set of regions engaged in task selection identified in "Analysis 1" (search volume), we identified those more specifically involved in rule-free compared with rule-based trials as additionally showing larger activations in rule-free than rule-based trials (red regions in Fig. 4A). The converse analysis was performed for rule-based compared with rule-free trials but indicated no significant activations in prefrontal regions.

Analysis 3. We computed prefrontal regions involved in rule-free selection only. These regions were identified as significantly exhibiting both the contrast between rule-free and baseline trials used in "Analysis 1" (search volume) and larger activations in rule-free than rule-based trials (conjunction analysis), and by excluding regions significantly exhibiting the contrast between rule-based and baseline trials used in "Analysis 1" ($P > 0.05$, uncorrected) (purple region in Fig. 4A). The converse analysis for rule-based trials indicated no significant activations in prefrontal regions.

Activations Associated with Task Values

We then analyzed activations associated with task values within the network of prefrontal regions involved in task selection. In the following, we refer to $V_{i,t}^c - V_{i,t}^f$ as "relative chosen-task values" [rCV]. Thus $[rCV] = (V_{i,t}^c - V_{i,t}^f)^2$ measures the unsigned difference between task values, which we simply refer to as "relative task values." For that purpose, we evaluated a GLM constructed as described above, which estimated value-related effects on activations using a procedure that controls for variations of trial frequencies across task values and trial types, as well as variations of value ranges across subjects. For each subject and trial type, trials were sorted according to rCV computed from the RL model. Next, trials were binned with a fixed sampling rate of 20 trials, starting from the value $rCV = 0$ upward to the most extreme positive rCV value, as well as downward to the most extreme negative rCV value (the most extreme bins included 10 to 29 trials). Rule-based and rule-free trials were then modeled as series of event-related regressors corresponding to these bins.

For each subject, accordingly, we extracted estimates of BOLD responses on every bin. Each estimate was therefore associated with an rCV value corresponding to relative chosen-task values averaged across trials in the related bin. We averaged these estimates across voxels in each prefrontal region involved in task selection (Fig. 4A). The resulting mean estimates were then entered in a repeated measure, mixed generalized linear model (mixed GLM) including rCV and $[rCV]^2$ as within-subject covariates along with trial types and regions-of-interest as within-subject factors. Subjects were treated as a random factor and the mixed GLM also included the interaction terms across factors and covariates. Significant interactions including the factors of regions-of-interest or trial types were subsequently unraveled with mixed GLMs separately conducted on each region and trial type. All significant effects are reported in Results. Note that rCV and rCV^2 covariates were almost orthogonal in both rule-free and rule-based trials (mean Variance Inflation Factor across subjects = 1.87 and 1.02 in rule-free and rule-based trials, respectively), thereby ensuring the validity of the mixed GLM estimation. The results were virtually identical when the rCV^2 covariates were orthogonalized relative to rCV covariates in each trial type, and/or when instead of quadratic value rCV^2 , absolute value $|rCV|$ was used as covariate.

To further control for response-time, task switching and reversal effects on value-related brain activations, we estimated the same GLM as described above, except that at the subject level (first-level), we further included the following regressors: a parametric regressor modeling RTs on every trial; a categorical regressor modeling trials with task repetition versus switching (when subjects switched vs. repeated the task from the preceding trial); finally, a parametric regressor modeling trial order from reversals in task-reward contingencies.

Finally, we also considered $V_{i,t}^c - V_{i,t}^f$ measuring the relative value of one task relative to the other one, denoted task A relative-to-B values [rTV] where

task A is arbitrarily defined, regardless of task values, actual choices and trial-types. To factor out task effects, task A was chosen as the vowel/consonant discrimination task in scanning runs #1 and #4, and as the lower/upper case discrimination task in scanning runs #2 and #3. We then estimated the same GLM as described above, except that task A relative-to-B values rTV and rTV^2 were used as covariates instead of rCV and rCV^2 . Note that even though $rCV^2 = rTV^2$, these 2 covariates capture different effects because the first-order covariates rCV and rTV strongly differ between the 2 GLMs. As rCV captures the effect of task values contingent upon actual choices, rCV^2 in the multiple regression captures the encoding of relative task values, independently of task selection processes. By contrast, rTV measures the value of one arbitrary task relative to the other, independently of task selection. As a result, rTV^2 in the multiple regression captures the effect of relative task values in relation with task selection processes. The results of these parametric analyses are shown in Figures 4 and 5.

Brain Activations Associated with Response Times

To identify brain regions associated with response times, we computed 2 contrasts from the GLM described above: at the subject level, we contrasted the series of parameter estimates of BOLD responses over value-based bins, separately in rule-based and rule-free trials, with respect to RTs averaged over each bin and normalized across bins (i.e., z-scored RTs for each subjects used as contrast weights) (Fig. 9).

Functional Connectivity

We analyzed psychophysiological interactions (PPIs) using the standard SPM8 method (Gitelman et al. 2003) for investigating transfers of task value information between medial and lateral prefrontal regions (Fig. 6). We first focused on the 3 regions of interest (ROIs) located in the dorsomedial, left and right lateral PFC (red regions in Fig. 4A). We therefore conducted 3 analyses corresponding to the 3 ROIs treated successively as seed regions. Time series of seed-ROI activity were extracted from 5-mm radius spheres centered on activation peaks (MNI coordinates in Fig. 4A) and adjusted for factors of no interest (scanning runs, movements, errors). Each analysis thus corresponded to a GLM including the regressors used in the univariate statistical analysis described above along with seed-ROI activity regressors, PPI regressors modeling the interaction between, seed-ROI activity and relative chosen-task values: for each trial type, we binned trials according to 4 intervals of rCV ($-$, $-$, $+$, $++$) in order to get enough trials in every bin for reliable estimates of interregional correlations.

In each analysis, parameter estimates of PPI regressors (i.e., interregional correlation values) were then computed for every voxel and averaged over every ROI considered as "target" region. Similarly to group-level univariate analyses described above, we entered all these averaged PPI estimates in repeated-measures mixed GLMs with rCV and rCV^2 as within-subject covariates, trial-types (rule-free vs. rule-based), hemispheres (left vs. right lateral PFC), and directionality (dorsomedial PFC as seed region and lateral PFC as target region vs. the converse directionality) as within-subject factors, and subjects as random factor. rCV corresponded to mean relative chosen-task values over trials included in each bin. We also included in the GLM the interaction terms across factors and covariates. Finally, the same PPI analyses were conducted between the other medial and lateral prefrontal regions associated with task execution in both rule-free and rule-based trials (Supplementary Fig. 2). As expected, all PPIs were independent of seed regions (all interactions, $F_s < 1$).

Effective Connectivity

We investigated dynamic causal models (DCMs) explaining activations of the medial-lateral PFC system using the DCM10 toolbox in SPM8 software package (Friston 2007). All DCMs were separately estimated on each hemisphere, in order to avoid making any additional assumptions regarding

neuronal interactions between left and right regions. DCMs were fitted on activation time series that was extracted from ROIs as in PPI analyses. To comply with independence of DCM analyses performed on each hemisphere, however, we extracted 2 time series from dorsomedial PFC activations corresponding to 2 symmetrical activated voxels located in the left and right hemisphere next to the activation peak ($x, y, z = -4, 24, 48$ and $x, y, z = 4, 24, 48$). Time series were adjusted to restrict variances to rule-based and rule-free trials. In these analyses, effects of relative task values rCV s and rCV^2 were estimated using a parametric regression analyses over all trials. These parametric analyses allowed us to further confirm the results of preceding analyses based on a binning procedure.

We compared a comprehensive collection of DCM models involving rCV and rCV^2 as covariates (see Supplementary Information for a detailed description) by using a family-wise Bayesian model comparison approach (Penny et al. 2010) (Fig. 7B, C, Materials and Methods). The analyses were conducted independently in each hemisphere with no prior assumptions regarding especially input regions and the functional contribution of MPC and LPC. (Fig. 7). We computed “exceedance probabilities” (i.e., Bayesian evidence from data supporting a model) for the left and right hemisphere separately and also jointly over both hemispheres. Results and Figure 7 report joint exceedance probabilities, given that similar results were observed in left and right hemispheres. Supplementary Tables 3 and 4 present the results for each hemisphere.

Results

Behavioral Performances

Subjects chose the more rewarded task (T+) in 75% of rule-free trials preceding reward reversals (Fig. 2A). This frequency consistently dropped to ~25% in the first trial following reversals (because T+ abruptly changed) but in the next/second-next trials, reached chance level (=50%) and again the plateau around ~75% about 10 trials later. This confirms that in rule-free trials, subjects chose the tasks according to reward expectations. In rule-based trials, subjects followed the rule and

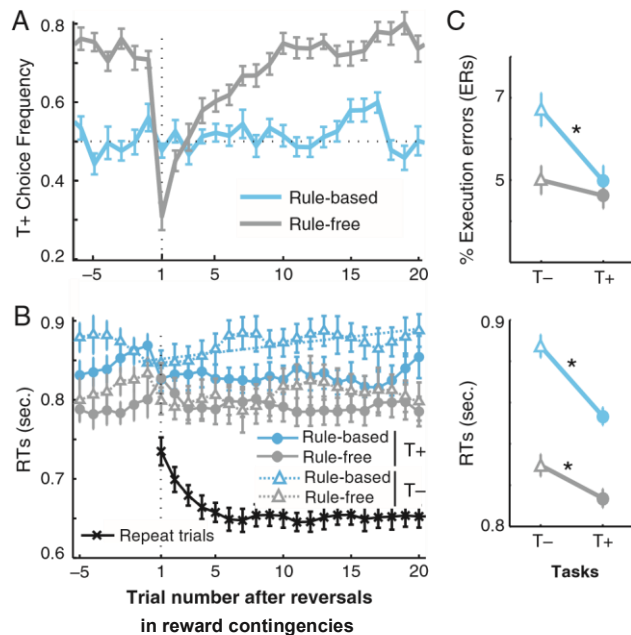


Figure 2. Human behavioral performances. (A) Frequencies of choosing the best rewarding task (T+) preceding and following reversals in reward contingencies for rule-free (gray lines) and rule-based (blue line) trials. Note that T+ changed from T1 to T2 (or the converse) when reversals occurred. In rule-based trials, subjects most often followed instruction cues occurring irrespective of reward contingencies and consequently performed T+ with about 50% frequency. Note however that these rule-based choices were slightly biased towards the betterrewarded task (see text). (B) Mean RTs preceding and following reversals in reward contingencies for T+ (circle) and T-

(triangle) performances in rule-free (blue) and rule-based (gray) trials. Note that in rule-based trials, RTs associated with T- performances significantly increased from reversals ($p < 0.05$, dashed lines). Black crosses show RTs in repeat trials from repeat block onsets. (C) 2×2 repeated measure ANOVA on execution error rates (% errors, ERs) and mean RTs crossing trial types (rule-free vs. rule-based) and tasks (T+ vs T-) as withinsubject factors. Both interactions are significant ($P < 0.05$). *Significant post hoc effect ($P < 0.05$). Error bars are SEM across subjects.

performed the cued task almost systematically (frequency = 96%), but slightly less frequently when the less rewarded task (T-) was cued (cued T-: 95%; cued T+: 97%; $F(1,22) = 25.6$, $P < 0.001$). Consistent with subjects' training, these rule-free and rule-based performances remained unchanged along experimental sessions (both $F_s(1,22) < 1.2$, $P_s > 0.31$). In all subsequent analyses, we then factored out the very rare rule-based trials when subjects departed from the rule.

As subjects followed the color rule and T+ and T- were randomly cued, subjects continually switched between tasks along series of rule-free and rule-based trials, thereby engaging cognitive control and preventing both rule-free and rule-based performances from turning into repetitive behaviors as in repeat trials. In baseline blocks of repeat trials, indeed, both correct response times (RTs) and error rates in task execution (ERs) decreased from block onsets and stabilized about 5 trials later (RTs: $T(22) = 6.04$, $P < 0.001$, ERs: $T(22) = 2.10$, $P = 0.05$) (Fig. 2B). In rule-free trials, by contrast, both RTs and ERs in performing T+ and T- exhibited no decreases from reversals (time courses: all $F_s(1,18) < 2.79$, $P_s > 0.12$) and remained considerably larger than those in repeat trials (all $T_s(22) > 4.84$, $P_s < 0.001$) (Fig. 2B). In rule-based trials, RTs and ERs were even larger with again no decreases and even a gradual increase from reversals when T- was cued (time courses for RTs: T+: $F < 1$; T-: $F(1,18) = 17.45$, $P < 0.001$; ERs: T+: $F < 1$; T-: $F(1,18) = 2.96$, $P = 0.10$).

Reflecting the processing of color cues, RTs were thus longer in rule-based than rule-free trials for both T- and T+ ($T_s(22) > 5.36$, $P_s < 0.001$). In both trial-types, furthermore, RTs were longer for T- than T+ ($T_s(22) > 3.70$, $P_s < 0.001$). However, this slow-down in T- compared with T+ performance was larger in rule-based than rule-free-trials (interaction T-/T+ by trial type: $F(1,22)=6.74$, $P = 0.016$, Fig. 2C). Importantly, these variations in RTs reflected no speed-accuracy trade-offs, as ERs exhibited the same interaction pattern (interaction T-/T+ by trial type: $F(1,22)=4.48$, $P = 0.046$): while ERs were similar in performing T+ in both trial types and T- in rule-free trials (both $T_s(22) < 1$), ERs was larger in performing T- in rule-based trials (T- in rulebased vs. rule-free trials: $T(22)=2.68$, $P = 0.014$). Overall, choosing T- compared with T+ was thus more difficult in rule-based than rule-free trials (Fig. 2C).

Altogether, these behavioral performances indicates that (1) at trial onsets, task selection was more often (and strongly) biased towards T+ than T-; (2) when colors signaled rule-free trials, these selection biases oriented task selection towards T+ or T-; (3) when colors signaled rule-based trials, color cues instead guided task selection according to the color rule. Compared with rule-free trials, thus, performing T- rather than T+ in rule-based trials was more often incongruent with selection biases, as the latter were more frequently oriented towards T+.

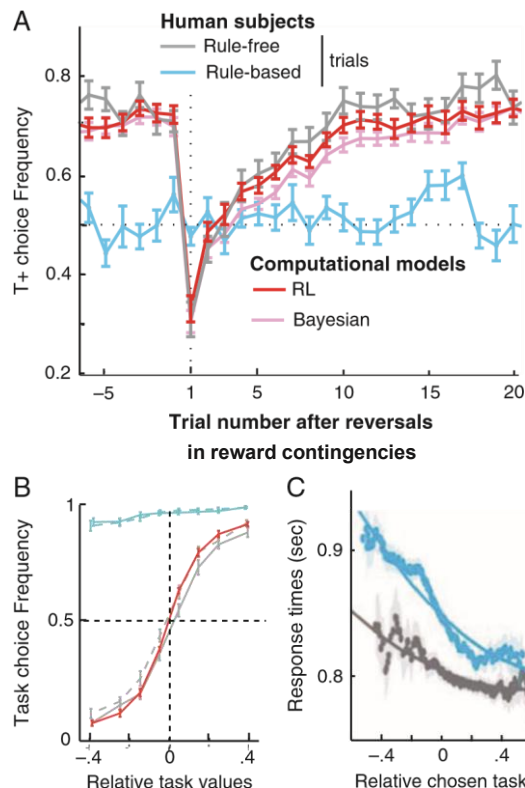
Computational Modeling

We then investigated 2 possible computational models describing how reward expectations are formed from feedbacks and drive task selection in rule-free trials: a “reinforcement learning” (RL) model and a “Bayesian Inference” model, both reflecting the protocol structure (see Materials and Methods). For both models, critically, subjects' choices in rule-free trials were best predicted when outcome expectations were updated from feedbacks delivered in “both” rule-free and rule-based trials (likelihood differences with alternative hypotheses: both $F_s(1,22) > 52.7$, $P_s < 0.001$). In the present protocol, indeed, task-reward contingencies were the same in both trial-types (provided that the color rule was followed): feedbacks in both trial-types were equally informative about the currently more rewarding task that could be selected in

subsequent rule-free trials. Consistently, the result shows that task selection in rule-free trials depended upon reward expectations computed from rewards delivered in “both” rule-free and rule-based trials. Furthermore, the RL model fitted these subjects’ choices better than the Bayesian model (differences in model likelihood and Bayesian Information Criteria: $F(1,22)=13.47$, $P = 0.001$) (Fig. 3A, B; Supplementary Table 1). All subsequent analyses were therefore based on the RL model. For clarity, we refer to reward expectations computed from this model as “task values,” which were passed and updated across “all” trials although they guided task selection in rule-free trials only.

As shown in Figure 3B, the probabilities of choosing one task computed from the RL model closely fitted the frequencies of task choices in rule-free trials, which increased as a sigmoid function of the task value relative to the other one. Critically, the empirical indifference point (choice frequency = 50%) closely matched the theoretical point of maximal ambiguity corresponding to identical task values. “Choice ambiguity” then decreased when the absolute difference between task values increased. Referring to the chosen minus unchosen task value as the rCV, we therefore found that choice ambiguity in rulefree trials varied as an inverted U-shaped function of rCVs centered on zero. By contrast, RTs monotonically increased when rCVs decreased from the largest positive values to the lowest negative values ($F = 33.9$, $P = 0.001$, Fig. 3C). This is consistent with sequential sampling models of selection processes predicting that in presence of reward biases along trial episodes, decreasing rCVs result in increased RTs (Supplementary Fig. 6; Bogacz et al. 2006; Milosavljevic et al. 2010). Thus, the protocol disentangled “response difficulty” reflected in the linear effect of rCV from “choice ambiguity” reflected in the inverted U-shaped function of rCV.

Note that in rule-based trials, rCVs simply measure the cued relative to uncued task values, given that subjects systematically chose the cued tasks (the very rare trials with the opposite performance being factored out). Consistent with the behavioral performances reported above, cued tasks were chosen slightly less frequently, when rCVs decreased (from 98 to 94%; $F(5, 110) = 15.05$, $P < 0.001$). Also, RTs monotonically increased when rCVs decreased ($F > 33.9$, $P < 0.001$) and these increases



values (rCV)

Figure 3. Model fits of human choices in rule-free trials. (A) Frequencies of choosing the best rewarding task (T+) preceding and following reversals in reward contingencies. Pink and red lines show predicted choices probabilities in rule-free trials according to subjects’ previous choices from the Bayesian and RL model, respectively. Gray and blue lines show subjects’ choice frequencies in rule-free and rule-based trials, resp. (data from Fig. 2). (B) Frequencies of subjects’ choices and predicted choice probabilities (red) in rule-free trials, for one task plotted against the values of this task relative to the other one, computed from the RL model (best-fitting model). Frequencies of subjects’ choices in rulebased trials for this task are shown in blue. Dashed and solid lines are for the lower/upper case and vowels/consonant discrimination task, respectively. Error bars in A, B are SEM across subjects. (C) Mean RTs plotted against relative chosen-task values (rCV) computed from the RL model. In this graph, for display purpose, all subjects are pooled together and data points are RTs smoothed over sliding windows (range = 0.2, sliding step = 0.006). Shaded areas show s.e.m. over trials within sliding windows.

were again steeper in rule-based than rule-free trials (interaction: $F = 188.9$, $P < 0.001$, Fig. 3C).

fMRI Regional Activations

We found that in both rule-free and rule-based trials, task selection rather than execution engaged the expected bilateral PFC network involved in cognitive control (Dosenbach et al. 2006), including dorsomedial and lateral PFC regions (Fig. 4A,

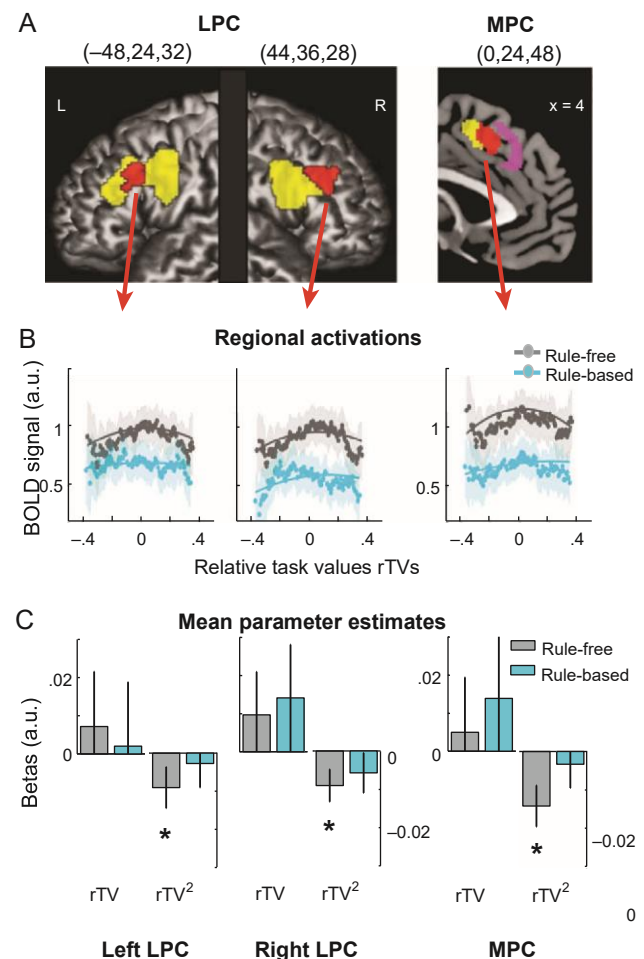


Figure 4. Prefrontal activations associated with task selection. (A) 3D rendering of prefrontal phasic activations (voxel-wise threshold: $P < 0.05$, corrected for family-wise errors over search volumes) in task selection compared with execution (baseline condition). Yellow: regions showing larger activations in rule-free and rule-based trials relative to repeat trials. Red: regions showing larger activations in rule-free and rule-based trials relative to repeat trials (as yellow regions) and larger

activations in rule-free than rule-based trials. Purple: regions showing larger activations in rule-free compared with repeat trials (no significant activations in rule-based compared with repeat trials: $P > 0.05$, uncorrected). Number x, y, z in parentheses are MNI coordinates of activation peaks for red regions (full data in Supplementary Table 2). MPC: medial prefrontal cortex. LPC: lateral prefrontal cortex. (B) Regional activations averaged over red regions plotted against task A-relative-to-B values (rTVs) in rule-free (gray) and rule-based (blue) trials. In this graph, for display purpose, subjects are pooled together and data points are activations smoothed over sliding windows on rTVs (range: 0.3). Vertical origins are mean activations in repeat trials. Shaded areas show SEM across trials within sliding windows. Lines show the second order polynomial regression presented in bottom panel. C, mixed generalized linear model analyses with subjects treated as random factor, rTV, rTV² and trial-types treated as within-subject factors. Histograms show betas for rTV and rTV² in rule-free (gray) and rule-based (blue) trials. Error bars are SEM across subjects. *Significant effects ($P < 0.05$).

yellow and red regions, Supplementary Table 2). Dorsomedial activations were located in the pre-SMA and slightly extended into the dorsal ACC. All these regions were engaged in both trial-types except that in rule-free trials, dorsomedial activations extended slightly more rostrally in the dorsal ACC (Fig. 4A, purple regions). Critically, among the regions activated in both trial-types, the anterior pre-SMA adjacent to dorsal ACC (denoted MPC) and bilaterally, the middle frontal gyrus (denoted LPC, BA 9) further exhibited differential activations between trial-types, with larger activations in rule-free than rule-based trials (Fig. 4A, red regions). Consistent with previous studies (Hampton et al. 2006; Hampton and O'Doherty 2007), thus, the MPC, left and right LPC were more specifically involved in task-set selection based on task values. Within the dorsomedial PFC, finally, the MPC was the region exhibiting the largest effect of task selection relative to execution in both trial-types.

To understand the role of MPC and bilateral LPC in task-set selection based on task values, we first analyzed how MPC and LPC activations varied according to task A-relative-to-B values rTVs, where task A is arbitrarily defined but controlled for task identity (see Materials and Methods). We reasoned that if such activations reflect value-based selection processes, then these regions should exhibit maximal activations when choice ambiguity is maximal: that is, activations should vary in rule-free but not rule-based trials as an inverted U-shape function of rTVs. We tested the prediction by analyzing MPC and LPC activations using a mixed generalized second-order polynomial regression model, that is, a linear model including rTV and rTV² as within-subject parametric covariates along with the other relevant categorical regressors. Covariate rTV² was introduced to capture the predicted variations of activations as U-shape functions of rTVs (see Materials and Methods).

As covariate rTV was controlled for task identity and reflected an arbitrary asymmetry, we logically found no activations exhibiting any linear effects of rTVs (MPC, left and right LPC in rule-free or rule-based trials: all $F_s < 1$, Fig. 4C). As predicted, however, we found that both MPC and bilateral LPC activations exhibited a negative effect of rTV² in rule-free but not rule-based trials (MPC, left and right LPC in rule-free trials: all $F_s > 4.20$, $P < 0.043$; in rule-based trials: all $F_s < 1$, Fig. 4C): in rule-free trials, all these activations varied as an inverted U-shape function of rTVs centered on zero, when choice ambiguity was maximal (Fig. 4B). This result supports the hypothesis that MPC and LPC activations reflect value-based task selection processes. To further assess the hypothesis, we investigated whether these activations were consistent with sequential sampling models of selection processes accounting for RTs in rule-free trials as observed above (Bogacz et al. 2006; Ploran et al. 2007; Milosavljevic et al. 2010; Frank et al. 2015; Gratton et al. 2017; Domenech et al. 2017). The standard neuronal implementation of these models indeed predicts that activations reflecting value-based selection processes should increase in rule-free trials when relative chosen-task values rCVs decrease (Wang 2012; Domenech et al. 2017; see Supplementary Fig. 6), while remaining independent of rCVs in rule-based trials. We tested the prediction by analyzing MPC and LPC activations using a mixed generalized second-order polynomial regression model, that is, a linear model including rCV and rCV² as within-subject parametric covariates along with the other relevant categorical regressors (see Materials and Methods). We introduced covariate rCV² to capture the variations of activations, when the effects of value-based selection are presumably captured by the negative linear effect of rCV and consequently, factored out. We reasoned that if the MPC and LPC further

“encode” task values passing through every trial and guiding task selection in rule-free trials (as shown above from computational modeling), then activations in both trial-types should increase when value information increases, that is, when the entropy across task values decreases or equivalently, when task values increasingly differ. Thus, if task values are encoded in every trial, then the polynomial regression model should reveal a “positive quadratic” effect of rCV, that is, a positive effect of rCV² similar in rule-free and rule-based trials, in addition to the negative linear effect of rCVs predicted in rule-free trials. In other words, factoring out this linear effect of rCVs in rule-free trials should reveal activations varying as a U-shaped function of rCVs similarly in rule-free and rule-based trials.

We found that both MPC and bilateral LPC activations exhibited negative effects of rCV only in rule-free trials (rule-free trials: both $F_s > 11.58$, $P_s < 0.001$. Rule-based trials MPC: $F = 3.27$, $P = 0.07$; left and right LPC: both $F_s < 1$; interaction rCV \times trial type in MPC, left and right LPC: all $F_s > 4.49$, $P_s < 0.03$) (Fig. 5B, D). In both trial-types, additionally, MPC unlike LPC activations further exhibited “positive” effects of rCV² (MPC: both $F_s > 4.85$, $P_s < 0.028$; left and right LPC: all $F_s < 1$) so that in both trial-types, MPC activations decreased when rCVs were closer to zero. This quadratic effect was similar between trialtypes (interaction rCV² \times trial type: $F < 1$, Fig. 5C, D). None of these effects significantly differed between the left and right PFC (all interactions with hemisphere: $F_s < 1$, Fig. 5D).

Importantly, this activation pattern remained virtually unchanged and all the value effects described above remained significant when the regression analysis further factored out the effects of response difficulty (RTs), task switching and reversals (Supplementary Fig. 4). Although MPC and bilateral LPC activations correlated with RTs and increased with task switching (all $F_s > 9.26$, $P_s < 0.003$), all these activations still exhibited negative effects of rCV in rule-free trials (all $F_s > 6.65$; $P_s < 0.01$). Moreover, MPC activations still exhibited a “positive” effect of rCV², which was again similar between trial-types (MPC: both $F_s > 4.69$, $P_s < 0.03$). Thus, none of these value effects reflected a general effect of response difficulty (Shenhav et al. 2014), task switching or reversals. The same results were also obtained when the regression analysis also included the relative value of the task “chosen in the preceding trial” as additional regressor. In all these regions, actually, this additional regressor captured no significant amounts of activation variances (MPC: $F < 1$; left and right LPC: both $F_s < 2.4$, $P_s > 0.12$, full variance analysis), even when the regression included only this value-based regressor (all $F_s < 1$). These results indicate that in the MPC, the effect of rCV² was unrelated to relative values of previously chosen tasks. Finally, similar activation patterns were observed in the other medial and lateral regions within the PFC network identified above (Supplementary Fig. 1) while elsewhere in the PFC and especially in the ventromedial PFC, no significant activations associated with task values were found at decision time (Supplementary Fig. 3). At feedback times in contrast, we found vmPFC activations to be correlated with feedback values rather than task values (Supplementary Fig. 3).

In summary, both MPC and LPC activations exhibited a negative linear effect of rCV confined to rule-free trials. Only MPC activations further exhibited a positive effect of rCV², which was further present in both trial-types: factoring out the negative linear effect of rCV in rule-free trials, MPC activations varied as a Ushape function of rCVs similarly in rule-free and rule-based trials. By contrast, LPC activations increased in rule-free trials when rCVs decreased, while being virtually independent of task

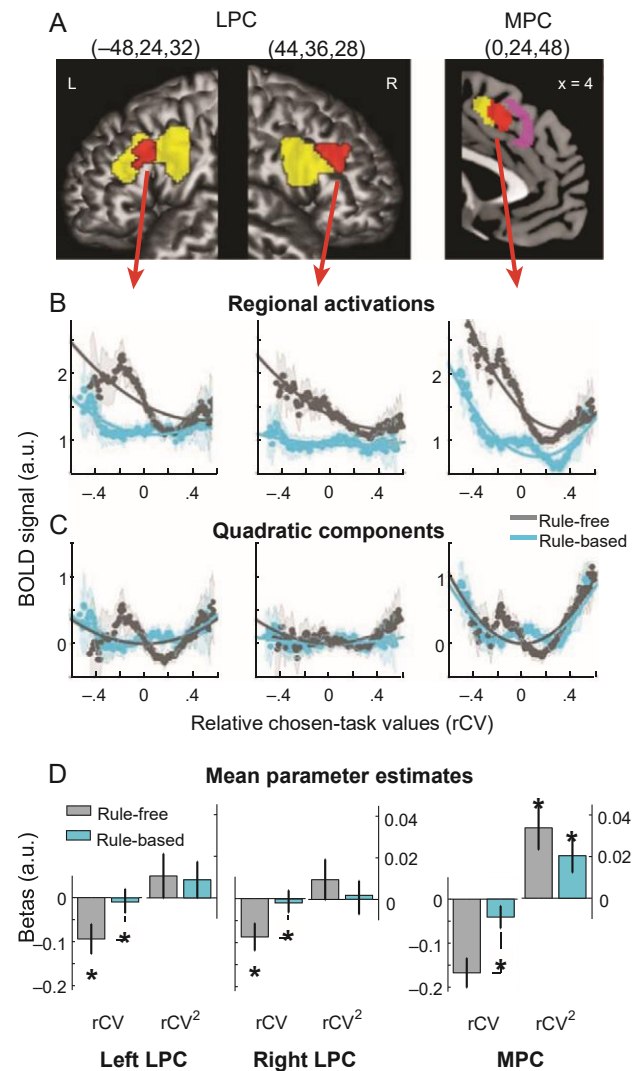


Figure 5. Prefrontal activations and value-based selection processes. (A) Same data and legend as in Figure 4A shown here for convenience. (B) Regional activations averaged over red regions plotted against relative chosen values (rCV) in rule-free (gray) and rule-based (blue) trials. In this graph, for display purpose, subjects are pooled together and data points are activations smoothed over sliding windows on rCV (range: 0.2, sliding step: 0.006). Vertical origins are mean activations in repeat trials. Shaded areas show SEM across trials within sliding windows. Lines show the second order polynomial regression on these data points. (C) Same data as in B but linearly detrended (only the quadratic and higher-order components remained). (D) Mixed generalized linear model analyses with subjects treated as random factor, rCV, rCV² and trial-types treated as within-subject factors. Histograms show betas for rCV and rCV² in rule-free (gray) and rule-based (blue) trials. Error bars are SEM across subjects. *Significant effects (P < 0.05). Supplementary Fig. 4 shows the same analyses factoring out the effects of RTs, task switching and reversals.

values in rule-based trials. This activation pattern supports the 2 following hypotheses: (1) the MPC encodes task values passing through every trial, although these values guided task selection only in rule-free trials and (2) both the MPC and LPC are involved in value-based task selection occurring in rule-free trials.

Functional Connectivity

The results above suggest that dorsomedial and lateral PFC regions are functionally coupled as guiding task selection in rulefree trials. To analyze this functional coupling, we computed the

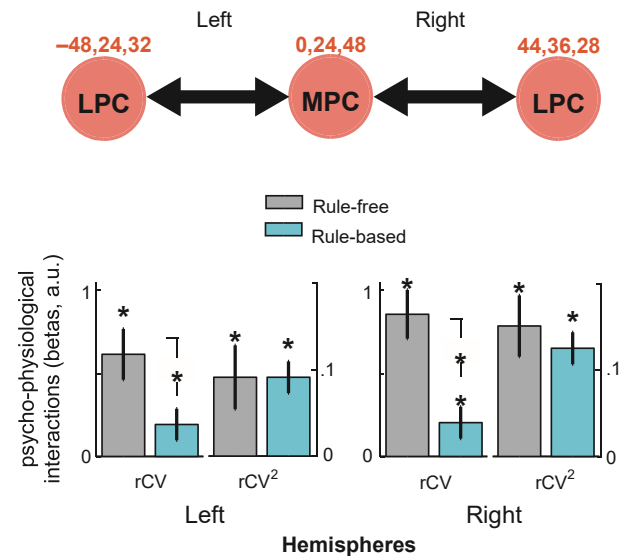


Figure 6. Functional connectivity between LPC and MPC activations. Histograms show psychophysiological interactions between LPC and MPC in the left and right hemisphere associated with relative chosen-task values rCV, and rCV². LPC and MPC regions are red regions in Figure 4 (red numbers are MNI coordinates x, y, z of activation peaks). Error bars are SEM across subjects. *P < 0.05.

Physiological correlations are shown in Supplementary Fig. 5.

PPIs (Gitelman et al. 2003) measuring how correlations between MPC and LPC activations vary with rCV and rCV² (see Materials and Methods).

In both hemispheres, MPC–LPC correlations increased significantly with rCVs in rule-free trials but marginally in rule-based trials (Fig. 6, Interaction trial type × rCV in each hemisphere: both Fs > 22.60, Ps < 0.001; rule-free trials, each hemisphere: Fs > 39.88, Ps < 0.001; rule-based trials: left F = 1.87, P = 0.17; right F = 4.15, P = 0.043.). There were no significant lateralization effects (all interactions with hemispheres: Fs < 1.67, Ps > 0.19). In rule-free rather than rule-based trials, thus, the functional coupling between the MPC and LPC increased with the reward advantages of actual choices or with the consistency between actual choices and relative task values. In both trial types, furthermore, MPC–LPC correlations also varied positively with the quadratic component rCV² (left and right hemisphere, rule-free and rule-based trials: all Fs > 12.66, Ps < 0.001). The quadratic effect was similar in both trial types and both hemispheres (all interactions: Fs < 1). Accordingly, the MPC–LPC functional coupling also reflected the encoding of relative task values irrespective of trial type and chosen tasks. A similar functional coupling was observed between the other medial and lateral PFC regions identified above (Supplementary Fig. 2)

Effective Connectivity

As MPC rather than LPC activations encoded relative task values passing through every trial, the PPIs reported above suggest that relative task values are conveyed from MPC to lateral PFC regardless of trial-types. This hypothesis then suggests that consistent with both regional activations and PPIs, task selection processes in rule-free trials propagate reciprocally from LPC to MPC. To directly test this putative model, we used Dynamic Causal Modeling (Friston 2007) for analyzing the reciprocal effective connectivity between MPC and LPC. In agreement with anatomical studies (Pandya and Yeterian 1996; Beckmann et al. 2009; Medalla and Barbas 2009), we modeled the MPC–LPC system as a coupled neuronal system whereby regional activity stems from intraregional, MPC-to-LPC and LPC-to-MPC interactions. The system responds to extrinsic signals that reflect stimulus occurrences conveying color cues: the system responds according to internal task value variables modulating neuronal influences within and between regions (see Materials and Methods). The DCM analysis allowed us to assess the

directionality of functional interactions observed in preceding PPI analyses between MPC and LPC in relation to task values, that is, whether relative task values rCV^2 are conveyed from MPC to LPC, and whether value-based selection processes reflected in linear rCV effects occur reciprocally from LPC to MPC. This putative model (Fig. 7A) thus predicts that the influence MPC exerts onto LPC (referred to as the M-to-LPC effective connectivity) should reflect the encoding of relative task values irrespective of actual choices, that is, vary with rCV^2 regardless of trial-types. Reciprocally, the influence LPC exerts to MPC (referred to as the L-to-MPC effective connectivity) is predicted to reflect value-based selection processes and consequently, to vary in rule-free trials according to the consistency between actual choices and relative task values or equivalently, to vary with rCV s in rule-free trials only.

First, we compared our putative model to alternative models assuming different coupling structures between MPC and LPC (see Supplementary Information for a detailed description). The results revealed that our putative model fitted neuronal activations better than (1) the “null” model assuming no MPC-LPC effective couplings (Exceedance probability over both hemispheres: $P > 0.999$, Fig. 7B, Supplementary Table 3); (2) “unidirectional” models assuming that either M-to-LPC or L-to-MPC effective connectivity varies with task-value variables rCV and rCV^2 ($P > 0.922$); and (3) bidirectional, “symmetric” models assuming that M-to-LPC and L-to-MPC effective connectivity are identical and associated with rCV and/or rCV^2 ($P > 0.998$). Across all these models, furthermore, evidence was that color cues directly influenced LPC rather than MPC activations ($P = 0.981$).

Second, we entered our putative model in a factorial analysis comprising the comprehensive collection of DCM models based on the same coupling structure for directly testing the directionality of task value effects and variations of value effects across trial-types. Thus, the factorial analysis comprised a collection of bidirectional, “asymmetric” coupling models assuming opposite directionalities of rCV and rCV^2 -related effective connectivity between MPC and LPC. The analysis included directionality (Mto-LPC and L-to-MPC connectivity associated with rCV^2 and rCV resp., vs. the converse), trial type (rule-free vs. rule-based) and input regions (LPC vs. MPC) as within-subject factors (Fig. 7C, Materials and Methods, full results in Supplementary Table 4). The results confirm the predictions from the proposed model. The main effect of directionality revealed that rCV^2 modulated M-to-LPC, while rCV modulated L-to-MPC effective connectivity ($P = 0.995$ against the converse directionality). Consistent with PPI results, furthermore, the directionality \times trial type interaction showed that rCV^2 modulated M-to-LPC connectivity “equally” across trial types, whereas rCV modulated L-to-MPC connectivity “differentially” across trial types ($P = 0.98$ against alternative interaction effects). Finally, evidence was again that color cues directly influenced LPC rather than MPC activations across all these models ($P = 0.83$) and notably in the proposed model ($P = 0.999$).

Parameter estimates in the proposed model confirm the preceding results in both hemispheres (all interactions with hemispheres: $F_s(1,22) < 1.45$, $P_s > 0.24$) (Fig. 8; Supplementary Table 5). The M-to-LPC effective connectivity varied with rCV^2

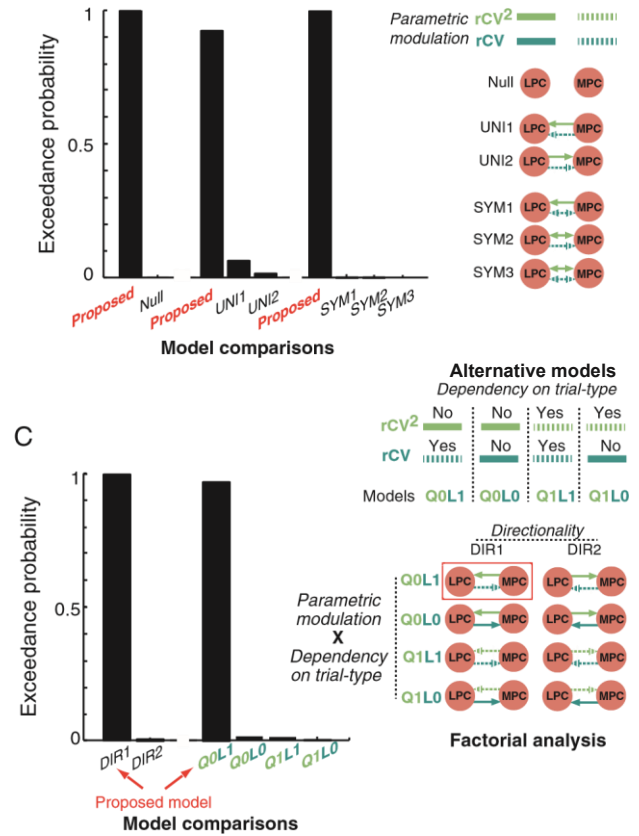
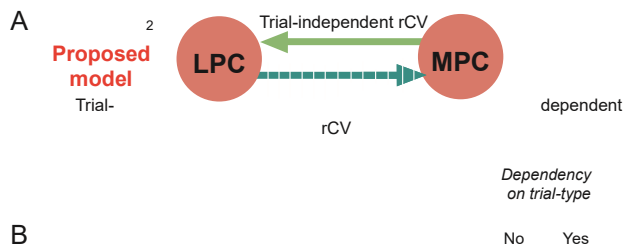


Figure 7. Bayesian model comparisons of effective connectivity between MPC and LPC. Histograms show joint evidence (exceedance probabilities) over both hemispheres for a comprehensive collection of dynamic causal models of MPC–LPC connectivity. (A) The proposed model (red) assumes that rCV^2 modulate Mto-LPC connectivity independently of trial types (solid, light-green arrows), while rCV modulates L-to-MPC connectivity in rule-free trials only (dashed dark-green arrows). (B) The proposed model is compared with various control models including null (no connectivity modulations), unidirectional (UNI1,2) and symmetrical bidirectional (SYM1,2,3) models. (C) Factorial analysis of asymmetrical models with directionality, trial type and value-related modulation as within-subject factors. The factorial analysis includes 8 families of dynamic causal models crossing these factors. In all these analyses, each model actually pools 2 distinct (sub-)models corresponding to distinct input regions for stimuli. See Supplementary Information for a detailed description of all models and Supplementary Tables 3, 4 for the full results.

“equally” in both trial types (both $T_s(22) > 2.87$, $P_s < 0.009$; interaction with trial type: $F < 1$). Conversely, the L-to-MPC effective connectivity varied with rCV in rule-free trials “only” ($T(22) = 3.23$, $P = 0.004$; rule-based trials: $T < 1$; interaction with trial type: $F(1,22) = 8.41$, $P < 0.009$) (Fig. 8). Consistently, intra-MPC effective connectivity varied with rCV^2 “equally” in both trial types ($T(22) > 2.31$, $P_s < 0.03$; interaction with trial type: $F < 1$), while intra-LPC effective connectivity varied with rCV in rulefree trials only ($T(22) = 2.95$, $P = 0.007$; rule-based trials: $T < 1$) (Supplementary Table 5).

Finally, additional analyses reveal that irrespective of value effects, the effective connectivity from LPC to MPC increased in task selection relative to execution similarly in rule-based and rule-free trials ($P = 0.98$ relative to differential or no increases) (Supplementary Information).

Discussion

The results confirm the involvement of lateral and dorsomedial PFC in rule-based cognitive control. In rule-based trials, the color rule rather than reward expectations associated with tasks (referred to as “task values”) guided task-set selection. Task values were simply maintained and updated during rulebased

trials, even though these values guided task-set selection in rule-free trials only. In rule-based trials, task-set selection compared with execution was associated with activations in the lateral and dorsomedial PFC. Dynamic Causal Modeling (Friston et al. 2003) further shows that the color rule driving task-set selection primarily involved lateral rather than dorsomedial PFC. Consistently, lateral PFC activations and their influence onto dorsomedial PFC in rule-based trials were independent of task values. In these trials, by contrast, dorsomedial PFC activations and their influence onto lateral PFC varied as an approximately zero-centered U-shaped function of relative chosen-task values (i.e., chosen/cued relative to unchosen/ uncued task values denoted rCVs) or equivalently, varied with relative task values irrespective of actual choices. Unlike these PFC activations, reaction times (RTs) in rule-based trials monotonically increased when relative chosen-task values rCVs decreased. Consequently, neither the dorsomedial PFC activations nor their influences onto lateral PFC activations reflect response difficulty measured by RTs. The present results thus provide evidence that irrespective of task-set selection, the dorsomedial PFC encodes and conveys to lateral PFC reward expectations associated with task sets. Consistent with these findings, previous studies suggest that (1) the lateral PFC drives task-set selection according to learned rules (Koechlin et al. 2003; Badre and D'Esposito 2007; Koechlin and Summerfield 2007; Badre 2008; Badre et al. 2009; Azuar et al. 2014; Nee and D'Esposito 2016) and (2) the dorsomedial PFC including the pre-SMA

M-to-LPC: rCV² modulation

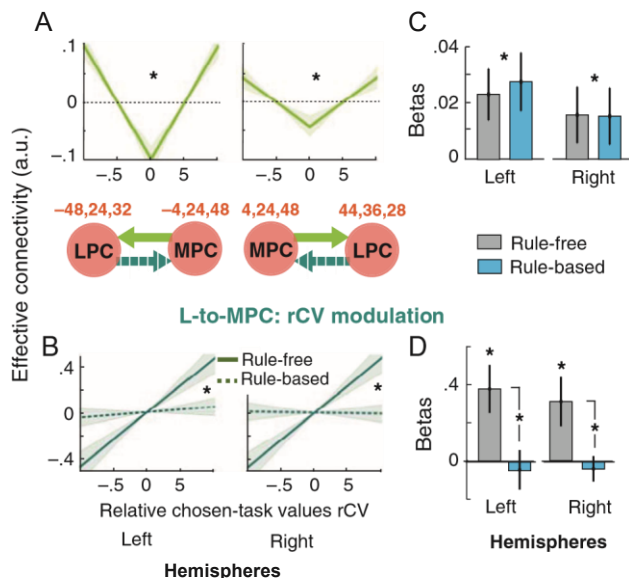


Figure 8. Parameter estimates of effective connectivity between MPC and LPC. (A, B) Proposed model of effective connectivity in both hemispheres (best-fitting model see Fig. 6): light-green, rCV² modulating M-to-LPC connectivity independently of trial types (A); dark-green, rCV modulating L-to-MPC connectivity in rule-free trials only (B). Red numbers indicate activation peaks used for dynamic causal modeling of effective connectivity. MPC and LPC refer to red regions in Figure 4. Mean parameter estimates of rCV² (C) and rCV (D) modulations in rule-free (gray) and rule-based (blue) trials in both hemispheres. Shaded areas and error bars are s.e.m. across subjects. *Significant effects (all *P*s < 0.004). See Supplementary Information and Supplementary Table 5 for the full data set (including parameter estimates of intraregional connectivity not shown here for clarity).

than rule-free trials, thereby ruling out any interpretations of these effects in general terms of response difficulty or conflict. Altogether, these findings thus provide convergent evidence that the dorsomedial and lateral PFC jointly implement value-based task selection processes.

Furthermore, when the effect of value-based task selection deriving from sequential sampling models was factored out from activations, dorsomedial but not lateral PFC activations varied in rule-free as in rule-based trials as a U-shaped function of relative chosen-task values rCVs. Thus, evidence was that in both rule-free and rule-based trials, the dorsomedial PFC encode relative task values irrespective of actual choices. The same result was observed in the

regulates the engagement of lateral PFC regions in rule-based cognitive control according to rewards at stake (Kouneiher et al. 2009).

In rule-free trials, by contrast, relative task values guided task-set selection. Univariate, multivariate, model-free and model-based analyses together indicate that value-based task selection also involved dorsomedial and lateral PFC: (1) at decision time, activations were larger on rule-free than rule-based trials specifically in dorsomedial and lateral PFC; (2) dorsomedial and lateral PFC exhibited maximal activations in rule-free trials when value-based choices were most ambiguous, that is, when the difference in task values was minimal and subjects' chose either task indifferently; (3) consistent with RTs, dorsomedial and lateral PFC activations matched the predictions of sequential sampling models of neuronal selection processes (Bogacz et al. 2006; Ploran et al. 2007; Milosavljevic et al. 2010; Frank et al. 2015; Gratton et al. 2017; Domenech et al. 2017), namely activations in rule-free trials increased when relative task-chosen values rCVs decreased; and (4), dorsomedial and lateral PFC activations increasingly correlated to each other when subject's choices in rule-free trials became more congruent with relative task values. Critically, all these effects were found to be unrelated to RTs. RTs were further longer and increased more with rCVs in rule-based

correlation between dorsomedial and lateral PFC activations, suggesting that relative task values encoded in the dorsomedial PFC are conveyed to lateral PFC. The analyses of effective connectivity between the dorsomedial and lateral PFC confirm this interpretation. In both rule-free and rule-based trials, the influence of dorsomedial PFC onto lateral PFC activations varied as a U-shape function of rCVs, that is, according to relative task values irrespective of actual choices, while the reciprocal influence of lateral PFC onto dorsomedial PFC activations varied as a linear function of rCV, that is, according to the consistency between relative task values and actual choices. Altogether, these findings suggest that task-set selection originates from lateral PFC based on relative task values conveyed from dorsomedial PFC and propagate backward to dorsomedial PFC. In rule-free trials, accordingly, dorsomedial PFC activations vary as the superimposition of a quadratic Ushaped function and a negative linear function of rCVs. The quadratic term is independent of actual choices and reflect the encoding of relative task values conveyed to lateral PFC for guiding downstream task-set selection, while the negative linear term is contingent upon actual choices and reflect the propagation of value-based selection processes from lateral PFC. In rulebased trials, by contrast, task selection is contingent upon color cues rather than relative task values. Consistently, dorsomedial PFC activations vary only as the quadratic U-shaped function of rCVs, as still reflecting the encoding of relative task values conveyed to lateral PFC, while both lateral and dorsomedial PFC activations exhibit no more linear dependences upon rCVs. In rule-based trials consistently, the influence of lateral PFC onto dorsomedial PFC activations reflected task selection irrespective of relative chosen-task values rCVs.

In rule-based trials as noted above, no prefrontal activations showed a significant effect of rCVs, even though RTs monotonically increased when rCVs decreased, with even stronger RT increases than in rule-free trials. This RT pattern is also consistent with current sequential sampling models of option selection (Bogacz et al. 2006; Milosavljevic et al. 2010; Frank et al. 2015). These models indeed predict that preceding trial onsets, larger relative task values orient task preparation processes more frequently (and strongly) towards the most rewarding task. In rulefree trials, such preparation processes continue to evolve to ultimately reach selection thresholds, so that task preparation and selection are most often congruent. In rule-based trials, by contrast, color cues guide task selection independently of such preparation processes. This implies the cued task to be increasingly incongruent with task preparation when rCVs decrease. As observed consequently, lower rCVs are associated with much larger RTs in rule-based compared with rule-free trials. We then searched for activations associated with this reaction time pattern

reflecting the incongruence between task preparation and selection. We found a unique frontal region located in the left premotor cortex exhibiting this activation pattern (Fig. 9): in both trial-types, left premotor activations increased when RT increased (or equivalently, rCVs decreased) and this effect was stronger in rule-based than rule-free trials. As lateral premotor regions presumably code for sensorimotor associations constituting task sets (Koechlin et al. 2003; Koechlin and Summerfield 2007; Badre 2008), these activations are likely to reflect the incongruence effects between the sensorimotor associations deriving from task preparation described above and the actual sensorimotor associations composing the ultimately selected task. Noticeably, the pattern of premotor activations was left lateralized, consistent with the use of verbal material and previous studies reporting the involvement of left premotor cortex in both left and right hand movement (Rushworth et al. 1998; O'Shea et al. 2007; Domenech et al. 2017).

The present results have several implications regarding current theories of medial PFC function and cognitive control in the PFC. First, we found evidence that the ventromedial PFC encodes the rewarding values of outcomes contingent upon task-set execution, the dorsomedial PFC encodes reward expectations associated with task sets, while the lateral PFC encodes learned rules linking external cues to task sets. These findings support the distributed view of PFC function and cognitive control, whereby the different PFC regions operate along distinct dimensions of cognitive control (Dosenbach et al. 2007, 2008; Rudebeck et al. 2008; Kounieher et al. 2009; Camille et al. 2011; Stuss 2011; Kolling et al. 2012; Rushworth et al. 2012; Power and Petersen 2013; Koechlin 2014, 2015; Domenech and Koechlin 2015).

Second, we found dorsomedial PFC activations encoding task values and involved in task-set selection to be centered in the anterior pre-SMA (Fig. 4A). Pre-SMA is commonly associated with the inhibition of irrelevant responses or habitual prepotent responses (Aron et al. 2007, 2014; Isoda and Hikosaka 2007; Nachev et al. 2007, 2008; Hikosaka and Isoda 2010). Pre-SMA activations are also found (although overlooked) to be associated

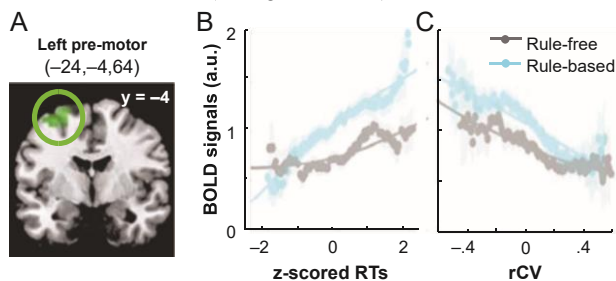


Figure 9. Frontal activations associated with RTs. (A) Coronal slices showing the only frontal activations (in left premotor cortex) associated with RTs in both rule-free and rule-based trials ($P < 0.001$, uncorrected) with RTs effects larger in rule-based than rule-free trials ($P < 0.05$, uncorrected). MNI coordinates of activation peaks are in parentheses. Activations averaged in the premotor region shown in (A), plotted against z-scored RTs (B) and rCV (C). Gray: rule-free trials. Blue: rule-based trials. In this graph, for display purpose, subjects are pooled together and data points are activations averaged across trials within sliding windows (B: range = 1; C: range = 0.2; sliding step: 0.006). Vertical origins show mean activations in repeat trials. Shaded areas are s.e.m. across trials within sliding windows. Lines show second order polynomial regression on these data points.

with reward expectations in decision-making tasks (review in Bartra et al. 2013). Importantly, we found these medial PFC activations to minimally involve the dorsal ACC. This might be viewed as surprising knowing that the dorsal ACC is commonly viewed as central in linking reward expectations to actions (Rushworth et al. 2011). Our finding however is consistent with previous results showing that the pre-SMA conveys motivational signals to lateral PFC in relation with immediate task switching, whereas the dorsal ACC conveys such signals for maintaining sustained control throughout behavioral episodes (Kounieher et al. 2009). The present protocol indeed mixed rule-free and rule-based trials inducing subjects to constantly switch back and forth between concurrent task sets. Moreover, previous studies show that the dorsal

ACC especially encodes the reward advantage to switch away from a default behavior and to inhibit the corresponding task set for exploring alternative behaviors (Kolling et al. 2012, 2014, 2016; Boorman et al. 2013; Donoso et al. 2014). In the present protocol, consistently, there was presumably neither exploration nor default behavior, as subjects constantly switch between 2 concurrent task sets: as a result, one task (the more rewarding one) was on average performed in 6 over 10 trials and the other performed in the remaining 4 trials. One might still consider the last performed task set to form the default behavior but medial PFC activations observed here were unrelated to the reward advantage to switch away from the last performed task set. Furthermore, preSMA rather than dorsal ACC activations have been associated with response conflict and inhibition (Rushworth et al. 2004; Aron et al. 2007; Nachev et al. 2008). In the present protocol, consistently, subjects switched between 2 concurrent task sets comprising incongruent sensorimotor associations, thereby requiring the inhibition of inappropriate ones. This inhibition process might be precisely achieved thanks to task-set selection signals that were found here to propagate from lateral PFC to the pre-SMA. Thus, the present results support the idea that the pre-SMA encodes reward expectations associated with concurrent behavioral strategies for driving immediate selection while according to previous results (Kolling et al. 2012, 2014, 2016; Boorman et al. 2013; Donoso et al. 2014), the dorsal ACC encodes the reward advantage to switch away from an ongoing/default course of action.

Third, the present results provide evidence that through reinforcement learning, task sets acquire appetitive values in the preSMA (referred to as “cached values”) so that task-set selection in the PFC may also be driven according to such cached values rather than to the appetitive values of inferred task-set outcomes. This conceptual distinction refers to what is usually named “model-free” versus “model-based” reinforcement learning (Dolan and Dayan 2013), which in the present study corresponds to the RL and Bayesian model described in the Results section, respectively. We found indeed that the RL model predicts behavioral performances better than the Bayesian model. Moreover, the pre-SMA encoded relative reward values associated with concurrent task sets regardless of actual choices, even when color cues rather than reward expectations guided task-set selection. In the present protocol, furthermore, the ventromedial PFC, known to encode inferred action outcomes and the relative reward values of chosen-action outcomes (review in Stalnaker et al. 2015), responded only to feedback reward values: no ventromedial PFC activations at decision or outcome time were found to encode relative rewards expected from “chosen” task sets (rCVs). Altogether, these findings provide convergent evidence that in the present protocol, concurrent task sets acquire “cached values” in the pre-SMA that drive task-set selection. The present study thus shows that cognitive control in the PFC may also operate through model-free reinforcement learning based on cached values acquired at the level of abstract task sets rather than simple sensorimotor associations. Consistent with this result, hierarchical reinforcement learning has been recently shown to involve model-free mechanisms (Cushman and Morris 2015). The result thus suggest that the distinction between model-free vs. model-based reinforcement learning is not equivalent to the distinction between habitual, automatic behaviors versus reflective, controlled behaviors involving basal ganglia and the PFC, respectively. Instead, the result suggests that the dorsomedial PFC, and more specifically the preSMA, contributes to task-set selection through model-free reinforcement learning, whereas according to previous results (Rushworth et al. 2011; Donoso et al. 2014; Stalnaker et al. 2015), task-set selection based on model-based reinforcement learning involves more rostral regions comprising the dorsal ACC and ventromedial PFC. The present protocol is likely to elicit model-free rather than model-based cognitive control mechanisms, because the protocol requires subjects to constantly switch back and forth between task sets and consequently, may prevent them from forming more costly modelbased inferences.

Fourth, the present results provide evidence that task-set selection originates in lateral rather than dorsomedial PFC. Our comprehensive analysis of the reciprocal effective interactions between medial and lateral PFC activations clearly favors the model, whereby the medial PFC encodes and conveys to lateral PFC choice-independent rather than choice-contingent

representations of reward expectations. This scales well with recent findings based on higher temporal resolution techniques like EEG and MEG (Cavanagh and Frank 2014; Jha et al. 2015). Moreover, our analysis indicates that dorsomedial PFC activations scaling with relative task-chosen values rCVs in rule-free trials reflected the influence of task-set selection processes propagating from lateral to medial PFC rather than the converse. The present findings thus provide little support to the view that the dorsomedial PFC collects all control signals relevant for task-set selection, including especially reward expectations associated with task sets, achieves task-set selection according to such signals and conveys this choice to lateral PFC for task-set implementation (Botvinick 2007; Shenhav et al. 2013, 2014). Instead, our findings are consistent with the view that the dorsomedial PFC generates control signals based on reward values of task sets and carries out performance monitoring (like revising reward values according to choice signals conveyed from lateral PFC as observed here), while the lateral PFC accumulates evidence for task-set selection and adjusts task-set control according to the behavioral context (like the presence/absence of rules and selection processes originating in lateral PFC observed here) (Dosenbach et al. 2007, 2008; Ploran et al. 2007; Neta et al. 2015, 2017; Gratton et al. 2017).

Fifth, we found that consistent with the massive and reciprocal direct projections between medial and lateral PFC (Pandya and Yeterian 1996; Johansen-Berg et al. 2004; Beckmann et al. 2009; Medalla and Barbas 2009), the lateral PFC exhibits the activation and effective connectivity profile combining rules and rewards in decision-making: namely, dorsomedial PFC influenced lateral PFC activations according to relative task values irrespective of actual choices; lateral rather than dorsomedial PFC received external cues as inputs guiding rule-based selection; finally, lateral PFC activated and influenced dorsomedial PFC activations as reflecting task-set selection according to relative task values and the color rules in rule-free and rule-based trials, respectively. These results provide evidence that the lateral PFC integrates learned rules with reward expectations conveyed from dorsomedial PFC for arbitrating between concurrent task sets, so that selection processes originate in lateral PFC and propagate backward to dorsomedial PFC. Accordingly, the dorsomedial PFC and more specifically the pre-SMA gains access to the chosen task set. This presumably allows this region to inhibit incongruent sensorimotor associations associated with unchosen task sets and to appropriately update reward expectations according to actual action outcomes. Reciprocally, conveying reward expectations from dorsomedial to lateral PFC may provide appropriate signals for learning potential rules in lateral PFC.

Finally, the present results suggest a mechanistic equivalence between the notion of free choices (internally driven as in rule-free trials) vs. instructed choices (externally driven as in rule-based trials). These 2 notions stem from 2 independent lines of research. Although they are both associated with overlapping dorsomedial and lateral PFC activations similar to those observed here (reviews in Ridderinkhof et al. 2004; Rushworth et al. 2011; Shenhav et al. 2013; Dixon and Christoff 2014), there have been very few attempts to connect the 2 notions and their neural mechanisms (Coutlee and Huettel 2012). The present results suggest unifying the 2 notions as stemming from the same functional coupling between dorsomedial and lateral PFC. Accordingly, we propose free choices to emerge from rewards expectations conveyed from dorsomedial to lateral PFC, driving selection processes that originate in lateral PFC and propagate backward to dorsomedial PFC (among other brain regions). Instructed choices then simply emerge from rules intervening in lateral PFC and overriding the influence of reward expectations onto selection processes originating in lateral PFC. This functional loop through lateral PFC enabling rules to intervene in the selection process may explain that in response to instruction cues, humans change free choices more flexibly than previously instructed choices (Fleming et al. 2009). The mechanistic equivalence between the notion of free and instructed choices is observed here at the level of task-set selection, supporting the idea that task sets constitute an abstract, common representation format across distinct PFC regions for collectively controlling behavior (Domenech and Koechlin 2015).

Authors' contributions

Both authors equally contribute to the work.

Supplementary Material

Supplementary data is available at Cerebral Cortex online.

Funding

This work was supported by a European Research Council Grant (ERC-2009-AdG #250106) to E.K.

Notes

We thank Muriel Ekovich and Jan Drugowitsch for their help.

References

- Alexander WH, Brown JW. 2011. Medial prefrontal cortex as an action-outcome predictor. *Nat Neurosci.* 14:1338–1344.
- Aron AR, Behrens TE, Smith S, Frank MJ, Poldrack RA. 2007. Triangulating a cognitive control network using diffusionweighted magnetic resonance imaging (MRI) and functional MRI. *J Neurosci.* 27:3743–3752.
- Aron AR, Robbins TW, Poldrack RA. 2014. Inhibition and the right inferior frontal cortex: one decade on. *Trends Cogn Sci.* 18:177–185.
- Azuar C, Reyes P, Slachevsky A, Volle E, Kinkingnehun S, Kounieher F, Bravo E, Dubois B, Koechlin E, Levy R. 2014. Testing the model of caudo-rostral organization of cognitive control in the human with frontal lesions. *Neuroimage.* 84: 1053–1060.
- Badre D. 2008. Cognitive control, hierarchy and the rostrocaudal organization of the frontal lobes. *Trends Cogn Sci.* 12: 193–200.
- Badre D, D'Esposito M. 2007. Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. *J Cogn Neurosci.* 19:2082–2099.
- Badre D, Hoffman J, Cooney JW, D'Esposito M. 2009. Hierarchical cognitive control deficits following damage to the human frontal lobe. *Nat Neurosci.* 12:515–522.
- Bartra O, McGuire JT, Kable JW. 2013. The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage.* 76:412–427.
- Beckmann M, Johansen-Berg H, Rushworth MF. 2009. Connectivity-based parcellation of human cingulate cortex and its relation to functional specialization. *J Neurosci.* 29: 1175–1190.
- Behrens TEJ, Woolrich MW, Walton ME, Rushworth MFS. 2007. Learning the value of information in an uncertain world. *Nat Neurosci.* 10:1214–1221.
- Bogacz R, Brown E, Moehlis J, Holmes P, Cohen JD. 2006. The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol Rev.* 113:700–765.
- 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.
- Botvinick MM. 2007. Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cogn Affect Behav Neurosci.* 7:356–366.
- Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD. 2001. Conflict monitoring and cognitive control. *Psychol Rev.* 108: 624–652.
- Camille N, Tsuchida A, Fellows LK. 2011. Double dissociation of stimulus-value and action-value learning in humans with orbitofrontal or anterior cingulate cortex damage. *J Neurosci.* 31:15048–15052.
- Cavanagh JF, Frank MJ. 2014. Frontal theta as a mechanism for cognitive control. *Trends Cogn Sci.* 18:414–421.

- Collins AG, Frank MJ. 2013. Cognitive control over learning: creating, clustering, and generalizing task-set structure. *Psychol Rev.* 120:190–229.
- Coutlee CG, Huettel SA. 2012. The functional neuroanatomy of decision making: prefrontal control of thought and action. *Brain Res.* 1428:3–12.
- Cushman F, Morris A. 2015. Habitual control of goal selection in humans. *Proc Natl Acad Sci USA.* 112:13817–13822.
- Dixon ML, Christoff K. 2014. The lateral prefrontal cortex and complex value-based learning and decision making. *Neurosci Biobehav Rev.* 45:9–18.
- Dolan RJ, Dayan P. 2013. Goals and habits in the brain. *Neuron.* 80:312–325.
- Domenech P, Koechlin E. 2015. Executive control and decisionmaking in the prefrontal cortex. *Curr Opin Behav Sci.* 1: 101–106.
- Domenech P, Redouté J, Koechlin E, Dreher JC. 2017. The neurocomputational architecture of value-based selection in the human brain. *Cereb Cortex.* 1:1–17.
- Donoso M, Collins AG, Koechlin E. 2014. Human cognition. Foundations of human reasoning in the prefrontal cortex. *Science.* 344:1481–1486.
- Dosenbach NU, Fair DA, Cohen AL, Schlaggar BL, Petersen SE. 2008. A dual-networks architecture of top-down control. *Trends Cogn Sci.* 12:99–105.
- Dosenbach NU, Fair DA, Miezin FM, Cohen AL, Wenger KK, Dosenbach RA, Fox MD, Snyder AZ, Vincent JL, Raichle ME, et al. 2007. Distinct brain networks for adaptive and stable task control in humans. *Proc Natl Acad Sci USA.* 104: 11073–11078.
- Dosenbach NU, Visscher KM, Palmer ED, Miezin FM, Wenger KK, Kang HC, Burgund ED, Grimes AL, Schlaggar BL, Petersen SE. 2006. A core system for the implementation of task sets. *Neuron.* 50:799–812.
- Fleming SM, Mars RB, Gladwin TE, Haggard P. 2009. When the brain changes its mind: flexibility of action selection in instructed and free choices. *Cereb Cortex.* 19:2352–2360.
- Frank MJ, Gagne C, Nyhus E, Masters S, Wiecki TV, Cavanagh JF, Badre D. 2015. fMRI and EEG predictors of dynamic decision parameters during human reinforcement learning. *J Neurosci.* 35:485–494.
- Friston KJ. 2007. Dynamic causal models for fMRI. In: Friston KJ, Ashburner JT, Kiebel SJ, Nichols TE, Penny WD, editors. *Statistical parametric mapping.* London: Academic Press. p. 541–560.
- Friston KJ, Harrison L, Penny W. 2003. Dynamic causal modelling. *Neuroimage.* 19:1273–1302.
- Gehring WJ, Knight RT. 2000. Prefrontal-cingulate interactions in action monitoring. *Nat Neurosci.* 3:516–520.
- Gitelman DR, Penny WD, Ashburner J, Friston KJ. 2003. Modeling regional and psychophysiological interactions in fMRI: the importance of hemodynamic deconvolution. *Neuroimage.* 19:200–207.
- Gogghari VM, MacDonald AW. 2008. Effects of varying the experimental design of a cognitive control paradigm on behavioral and functional imaging outcome measures. *J Cogn Neurosci.* 20:20–35.
- Gratton C, Neta M, Sun H, Ploran EJ, Schlaggar BL, Wheeler ME, Petersen SE, Nelson SM. 2017. Distinct stages of moment-to-moment processing in the cinguloopercular and frontoparietal networks. *Cereb Cortex.* 27:2403–2417.
- Hampton AN, Bossaerts P, O'Doherty JP. 2006. The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *J Neurosci.* 26: 8360–8367.
- Hampton AN, O'Doherty JP. 2007. Decoding the neural substrates of reward-related decision making with functional MRI. *Proc Natl Acad Sci USA.* 104:1377–1382.
- Hikosaka O, Isoda M. 2010. Switching from automatic to controlled behavior: cortico-basal ganglia mechanisms. *Trends Cogn Sci.* 14:154–161.
- Holroyd CB, Coles MG. 2002. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol Rev.* 109:679–709.
- Hyafil A, Summerfield C, Koechlin E. 2009. Two mechanisms for task-switching in the prefrontal cortex. *J Neurosci.* 29: 5135–5142.
- Isoda M, Hikosaka O. 2007. Switching from automatic to controlled action by monkey medial frontal cortex. *Nat Neurosci.* 10:240–248.
- Jha A, Nachev P, Barnes G, Husain M, Brown P, Litvak V. 2015. The frontal control of stopping. *Cereb Cortex.* 25:4392–4406.
- Johansen-Berg H, Behrens TE, Robson MD, Drobnjak I, Rushworth MF, Brady JM, Smith SM, Higham DJ, Matthews PM. 2004. Changes in connectivity profiles define functionally distinct regions in human medial frontal cortex. *Proc Natl Acad Sci USA.* 101:13335–13340.
- Koechlin E. 2014. An evolutionary computational theory of prefrontal executive function in decision-making. *Philos Trans R Soc Lond B Biol Sci.* 369:20130474.
- Koechlin E. 2015. Prefrontal executive function and adaptive behavior in complex environments. *Curr Opin Neurobiol.* 37:1–6.
- Koechlin E, Ody C, Kouneiher F. 2003. The architecture of cognitive control in the human prefrontal cortex. *Science.* 302: 1181–1185.
- Koechlin E, Summerfield C. 2007. An information theoretical approach to prefrontal executive function. *Trends Cogn Sci.* 11:229–235.
- Kolling N, Behrens T, Wittmann MK, Rushworth M. 2016. Multiple signals in anterior cingulate cortex. *Curr Opin Neurobiol.* 37:36–43.
- Kolling N, Behrens TE, Mars RB, Rushworth MF. 2012. Neural mechanisms of foraging. *Science.* 336:95–98.
- Kolling N, Wittmann M, Rushworth MF. 2014. Multiple neural mechanisms of decision making and their competition under changing risk pressure. *Neuron.* 81:1190–1202.
- Kouneiher F, Charron S, Koechlin E. 2009. Motivation and cognitive control in the human prefrontal cortex. *Nat Neurosci.* 12:939–945.
- Medalla M, Barbas H. 2009. Synapses with inhibitory neurons differentiate anterior cingulate from dorsolateral prefrontal pathways associated with cognitive control. *Neuron.* 61:609–620.
- Miller EK, Cohen JD. 2001. An integrative theory of prefrontal cortex function. *Annu Rev Neurosci.* 24:167–202.
- Milosavljevic M, Malmaud J, Huth A, Koch C, Rangel A. 2010. The drift diffusion model can account for the accuracy and reaction time of value-based choices under high and low time pressure. *Judgm Decis Mak.* 5:437–449.
- Nachev P, Kennard C, Husain M. 2008. Functional role of the supplementary and pre-supplementary motor areas. *Nat Rev Neurosci.* 9:856–869.
- Nachev P, Wydell H, O'Neill K, Husain M, Kennard C. 2007. The role of the pre-supplementary motor area in the control of action. *Neuroimage.* 36(Suppl 2):T155–T163.
- Nee DE, D'Esposito M. 2016. The hierarchical organization of the lateral prefrontal cortex. *Elife.* 5:e12112.
- Neta M, Miezin FM, Nelson SM, Dubis JW, Dosenbach NUF, Schlaggar BL, Petersen SE. 2015. Spatial and temporal characteristics of error-related activity in the human brain. *J Neurosci.* 35:253–266.
- Neta M, Nelson SM, Petersen SE. 2017. Dorsal anterior cingulate, medial superior frontal cortex, and anterior insula show performance reporting-related late task control signals. *Cereb Cortex.* 27:2154–2165.
- Neta M, Schlaggar BL, Petersen SE. 2014. Separable responses to error, ambiguity, and reaction time in cingulo-opercular task control regions. *Neuroimage.* 99:59–68.
- O'Shea J, Johansen-Berg H, Trief D, Göbel S, Rushworth M. 2007. Functionally specific reorganization in human premotor cortex. *Neuron.* 54:479–490.
- Pandya DN, Yeterian EH. 1996. Morphological correlations of the human and monkey frontal lobe. In: Damasio AR, Damasio H, Christen Y, editors. *Neurobiology of decisionmaking.* Berlin Heidelberg: Springer-Verlag. p. 13–46.

- Penny WD, Stephan KE, Daunizeau J, Rosa MJ, Friston KJ, Schofield TM, Leff AP. 2010. Comparing families of dynamic causal models. *PLoS Comput Biol*. 6:e1000709.
- Ploran EJ, Nelson SM, Velanova K, Donaldson DI, Petersen SE, Wheeler ME. 2007. Evidence accumulation and the moment of recognition: dissociating perceptual recognition processes using fMRI. *J Neurosci*. 27:11912.
- Power JD, Petersen SE. 2013. Control-related systems in the human brain. *Curr Opin Neurobiol*. 23:223–228.
- Rescorla RA, Wagner AR. 1972. A theory of pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: Black AH, Prokasy WF, editors. *Classical Conditioning II: Current research and theory*. New York: Appleton-Century-Crofts. p. 64–99.
- Ridderinkhof KR, Ullsperger M, Crone EA, Nieuwenhuis S. 2004. The role of the medial frontal cortex in cognitive control. *Science*. 306:443–447.
- Rogers RD, Monsell S. 1995. Costs of predictable switch between simple cognitive tasks. *J Exp Psychol Gen*. 124:207–231.
- Rudebeck PH, Bannerman DM, Rushworth MF. 2008. The contribution of distinct subregions of the ventromedial frontal cortex to emotion, social behavior, and decision making. *Cogn Affect Behav Neurosci*. 8:485–497.
- Rushworth M, Nixon PD, Wade DT, Renowden S, Passingham RE. 1998. The left hemisphere and the selection of learned actions. *Neuropsychologia*. 36:11–24.
- Rushworth MF, Kolling N, Sallet J, Mars RB. 2012. Valuation and decision-making in frontal cortex: one or many serial or parallel systems? *Curr Opin Neurobiol*. 22:946–955.
- Rushworth MF, Noonan MP, Boorman ED, Walton ME, Behrens TE. 2011. Frontal cortex and reward-guided learning and decision-making. *Neuron*. 70:1054–1069.
- Rushworth MF, Walton ME, Kennerley SW, Bannerman DM. 2004. Action sets and decisions in the medial frontal cortex. *Trends Cogn Sci*. 8:410–417.
- Sakai K. 2008. Task set and prefrontal cortex. *Annu Rev Neurosci*. 31:219–245.
- Shenhav A, Botvinick MM, Cohen JD. 2013. The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron*. 79:217–240.
- Shenhav A, Straccia MA, Cohen JD, Botvinick MM. 2014. Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value. *Nat Neurosci*. 17:1249–1254.
- Stalnaker TA, Cooch NK, Schoenbaum G. 2015. What the orbitofrontal cortex does not do. *Nat Neurosci*. 18:620–627.
- Stuss DT. 2011. Functions of the frontal lobes: relation to executive functions. *J Int Neuropsychol Soc*. 17:759–765.
- Wang X-J. 2012. Neural dynamics and circuit mechanisms of decision-making. *Curr Opin Neurobiol*. 22:1039–1046.