Behavioral/Cognitive

Frontoparietal Representations of Task Context Support the Flexible Control of Goal-Directed Cognition

Michael L. Waskom,[1]* **Dharshan Kumaran,**[1,3]* **Alan M. Gordon,**[1] **Jesse Rissman,**[1,4] and **Anthony D. Wagner**[1,2]

[1]Department of Psychology, Stanford University, Stanford, California 94305, [2]Neurosciences Program, Stanford University, Stanford, California 94305, [3]Institute of Cognitive Neuroscience, University College London, London, United Kingdom WC1N 3AR, and [4]Department of Psychology, University of California, Los Angeles, Los Angeles, California 90095

Cognitive control allows stimulus-response processing to be aligned with internal goals and is thus central to intelligent, purposeful behavior. Control is thought to depend in part on the active representation of task information in prefrontal cortex (PFC), which provides a source of contextual bias on perception, decision making, and action. In the present study, we investigated the organization, influences, and consequences of context representation as human subjects performed a cued sorting task that required them to flexibly judge the relationship between pairs of multivalent stimuli. Using a connectivity-based parcellation of PFC and multivariate decoding analyses, we determined that context is specifically and transiently represented in a region spanning the inferior frontal sulcus during context-dependent decision making. We also found strong evidence that decision context is represented within the intraparietal sulcus, an area previously shown to be functionally networked with the inferior frontal sulcus at rest and during task performance. Rule-guided allocation of attention to different stimulus dimensions produced discriminable patterns of activation in visual cortex, providing a signature of top-down bias over perception. Furthermore, demands on cognitive control arising from the task structure modulated context representation, which was found to be strongest after a shift in task rules. When context representation in frontoparietal areas increased in strength, as measured by the discriminability of high-dimensional activation patterns, the bias on attended stimulus features was enhanced. These results provide novel evidence that illuminates the mechanisms by which humans flexibly guide behavior in complex environments.

Key words: attention; cognitive control; decision making; prefrontal cortex

Introduction

To behave intelligently in a complex world, humans must attend to relevant perceptual information and select actions that will attain goals, although what information is relevant and which actions are rewarded may change over time. The ability to flexibly interact with the environment, often termed cognitive control, is a hallmark of human behavior. Despite its ubiquity and centrality, though, the mechanisms that give rise to control remain poorly understood.

The prefrontal cortex (PFC) is thought to be critically involved in control. Injury to lateral PFC impairs performance when rules must be used to select from several possible stimulus-response associations (Milner, 1963; Keele and Rafal, 2000; Buck-

ley et al., 2009), and functional magnetic resonance imaging (fMRI) reveals increased activation in lateral PFC under similar manipulations (MacDonald et al., 2000; Braver et al., 2003; Brass and Cramon, 2004; Badre and D'Esposito, 2007). PFC neurons have complex response properties and can show selectivity to abstract task rules that generalize over specific cues, stimuli, and responses (Wallis et al., 2001; Buschman et al., 2012). This is thought to support control by maintaining a representation of task context: higher-order information that provides a source of bias on signals in sensorimotor circuits (Cohen et al., 1990; Miller and Cohen, 2001; Rougier et al., 2005).

Recent fMRI (Cole et al., 2011; Woolgar et al., 2011; Reverberi et al., 2012; Zhang et al., 2013) and electrophysiological (Mante et al., 2013; Rigotti et al., 2013; Stokes et al., 2013) results offer partial support for this view by using multivariate analyses to decode information about task context from PFC activity. Although there is some evidence that perceptual representations are enhanced during control (Sakai and Passingham, 2003; Egner and Hirsch, 2005; Chiu et al., 2011; Nelissen et al., 2013) and that the PFC is causally involved (Zanto et al., 2011), a direct association between the specific mechanism of prefrontal context representation and biased processing has yet to be demonstrated. The role of other cortical regions in control is also unclear. Lesion and nonhuman primate work has typically focused on PFC, but lateral parietal cortex is also consistently activated when control is
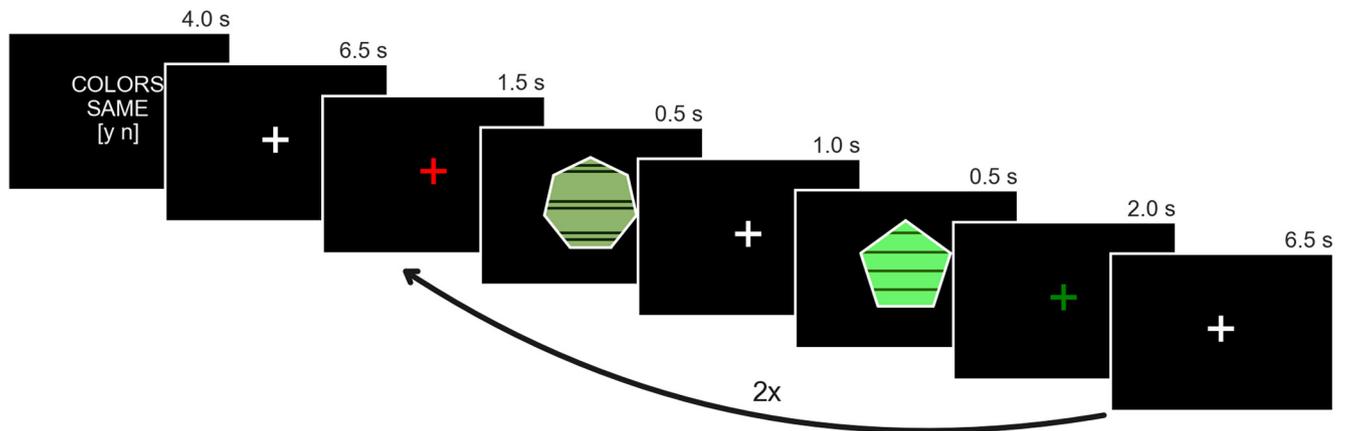
Figure 1. Experimental design. The task was organized into miniblocks of three consecutive trials with the same active rules, which were cued at the start of the block. The stimuli in this figure demonstrate the two possible feature values along the dimensions of shape, color, and pattern.

required (Braver et al., 2003; Brass and Cramon, 2004; Badre and D'Esposito, 2007). These observations align with an emerging perspective drawn from resting-state and task-based connectivity analyses (Power et al., 2011; Cole et al., 2013) that control is supported in part by a distributed frontoparietal network centered on the inferior frontal sulcus and intraparietal sulcus.

We used fMRI to measure brain activity while human participants used shifting rules to guide decisions. Our analyses used multivariate decoding methods to characterize the spatial and temporal expression of information in frontoparietal control regions. To understand the influences on and consequences of context representation, we related the strength of these measurements to task structure and signatures of attentional bias in visual cortex. Our results indicate that demands on cognitive control lead to increased discriminability of context representations in lateral frontoparietal cortex, which is in turn associated with enhanced processing of contextually relevant perceptual information.

Materials and Methods

Participants and experimental design. Fifteen healthy, right-handed native English speakers (18–26 years old, 7 females) participated after giving informed written consent in accordance with the Stanford University Institutional Review Board. Subjects were paid $20/h for participating and additionally received a performance-based bonus of up to $30. Stimuli for the cued sorting task consisted of polygons with three feature dimensions: color, shape, and pattern. There were two possible feature values along each dimension: a bright (#6AF46B) or muted (#8FB46B) shade of green, pentagonal or heptagonal shape, and single or double narrow horizontal bands (Fig. 1). The task required participants to make rule-dependent judgments about the relationship between sequentially presented pairs of these stimuli. Responses were governed by two independent sets of rules: the dimension rules specified which of the three dimensions was relevant for that block (i.e., color, shape, or pattern), and the decision rules specified whether participants should respond positively if the stimulus features were the same or different along that dimension. Participants thus responded "yes" or "no" on each trial, with "yes" indicating a match in the same rule or a nonmatch in the different rule. Participants responded with the index and middle fingers of their right hand; the mapping of "yes" and "no" responses and button presses was counterbalanced across runs. There was no feedback provided during the main experiment.

The task was implemented with a slow event-related design, but it was logically organized into miniblocks of three trials (Fig. 1). At the beginning of each block, an orthographic cue indicating the rules for that block and reiterating the response mapping for that run was presented for 4 s; the rule cue was followed by 6.5 s of fixation on a white cross preceding

the 3 trial cycles. Each trial began when the fixation cross turned red for 1.5 s, which cued the impending stimulus onset. Subsequently, 2 stimuli were sequentially presented for 0.5 s each with 1 s of intervening fixation. At the offset of the second stimulus, the fixation cross turned green for 2 s. Participants were instructed to respond as quickly and accurately as possible when the second stimulus appeared, although they were given up to 2.5 s to indicate their response. After the response window on the first two trials of the block, the fixation cross turned back to white for 6.5 s, which provided temporal spacing between trials; 8 s of fixation on a yellow cross followed the final stimulus event in the block.

The experiment consisted of four scanning runs, each containing four miniblocks for each of the six rule pairings. Therefore, there were a total of 24 trials per dimension rule (collapsed across decision rule) and 36 trials per decision rule (collapsed across dimension rule) in each run. The attended features matched on half of the trials and were balanced across each set of rules. Both features on the two unattended dimensions either matched or differed on each trial; these features matched on half of the trials and unattended feature matches were evenly counterbalanced against attended feature matches. A different order, subject to these constraints, was used for each participant. Participants learned the task outside of the scanner and were required to reach a performance criterion before advancing to the scanning phase: training finished when at least nine correct responses were made on 10 consecutive trials after the 24th trial. Participants also performed a second short training session in the scanner to acclimatize to the magnet environment and to ensure proper understanding of the task. After completion of the experiment, one of the four runs was chosen randomly and a bonus payment was calculated based on response accuracy and average reaction time over that run.

Image acquisition and preprocessing. Whole-brain imaging was performed on a 3T Signa MRI system (GE Medical Systems). Functional images were obtained using a T2*-weighted two-dimensional gradient echo spiral-in/out pulse sequence (TR = 2 s, TE = 30 ms, flip angle = 75°, 30 slices, 3.28 × 3.28 × 4 mm, axial oblique sequential acquisition). In addition, a whole-brain high-resolution T1-weighted spoiled gradient recalled echo anatomical volume was acquired for cortical surface modeling and across-run alignment. Visual stimuli were projected onto a screen and viewed through a mirror; responses were collected through a magnet-compatible button box.

Imaging data were processed with a workflow of FSL (Smith et al., 2004) and Freesurfer (Dale et al., 1999) tools implemented in Nipype (Gorgolewski et al., 2011). Each time series was first realigned to its middle volume using normalized correlation optimization and cubic spline interpolation. To correct for differences in slice acquisition times, data were temporally resampled to the TR midpoint using sinc interpolation. Images with artifacts were automatically identified as those frames on which total displacement relative to the previous frame exceeded 0.5 mm or in which the average intensity across the whole brain deviated from the run mean by greater than three SDs. These frames were ex-
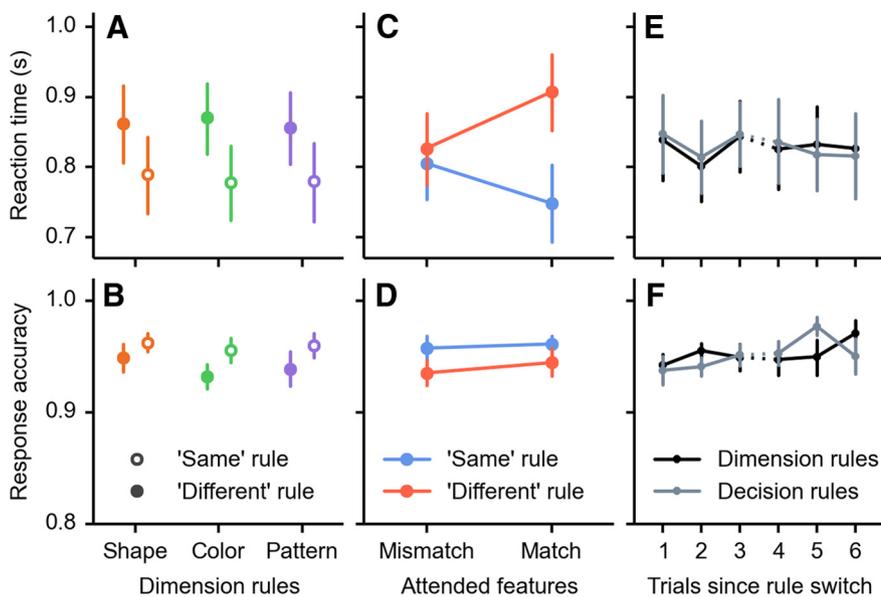
Figure 2. Main behavioral results. Mean within-subject median reaction time (top row) and mean within-subject proportion correct responses (bottom row) sorted by the two rule sets (*A*, *B*), by the decision rule and whether the attended features matched or differed (*C*, *D*), and by the number of trials since the rule switch plotted separately with respect to the dimension and decision rules (*E*, *F*). The time courses in *E* and *F* are dashed between trials 3 and 4 to indicate passage through a miniblock boundary. Error bars on all facets indicate bootstrapped SEs of the aggregate values across subjects. RTs are plotted only for trials included in the imaging analyses.

cluded from all decoding analyses. This procedure discarded 23 time points per subject on average, or 1% of the raw data. Finally, the time series data were high-pass filtered by fitting and removing Gaussian-weighted running lines with an effective cycle cutoff of 128 s. Functional data were not spatially smoothed.

Separately, the T1-weighted anatomical volume was processed using Freesurfer to segment the gray-white matter boundary and construct tessellated meshes representing the cortical surface (Dale et al., 1999). Functional data from each run were then registered to the anatomical volume with a six degree-of-freedom rigid alignment optimizing a boundary-based cost function (Greve and Fischl, 2009). Finally, runs 2–4 were resampled into the space of run 1 using cubic spline interpolation to bring the data into a common alignment. All decoding analyses were thus performed in a subject-specific space, avoiding normalization to a group template.

Region-of-interest definition. Our primary analyses were motivated by a priori interest in measuring context information in PFC. We thus pursued a hypothesis-driven region-of-interest (ROI) approach, which affords several advantages over a voxelwise analysis. Specifically, ROI analyses avoid the need for precise voxel-to-voxel alignment across subjects, the loss of power that is incurred when correcting for multiple comparisons over a large search space, and the difficulties of interpreting thresholded statistical maps. Furthermore, it is straightforward to directly test the differences in a dependent variable between ROIs, which is critical when investigating dissociations between regions.

We derived ROIs from a population atlas of task-independent cortical networks (Yeo et al., 2011; see Fig. 3*A*), which provides a complete parcellation of the cortical surface into networks that show correlated spontaneous fluctuations while subjects are resting. This allowed us to test a specific hypothesis about the organization of PFC that has received support from analyses of the correspondence between task-independent networks and peak activations in task-based experiments (Nelson et al., 2011; Power et al., 2011) and network dynamics during flexible control (Cole et al., 2013). Specifically, it is thought that a distributed frontoparietal network with a PFC node centered on the inferior frontal sulcus is centrally important as a source of top-down influence on sensorimotor processes.

By using a population atlas, we were able to perform analyses that could be compared directly across experiments. However, a particular

virtue of the Yeo atlas (Yeo et al., 2011) is that it is defined in Freesurfer's common surface space, which allowed us to create ROIs with increased precision relative to a volume-based atlas. Our specific procedure for creating ROIs first warped region labels back to the individual subject surfaces by inverting the spherical normalization parameters obtained during cortical reconstruction (Fischl et al., 1999; see Fig. 3*B*). Vertex coordinates within each of these labels were then transformed into the native functional space by inverting the linear functional-to-anatomical transformation for the first run. Finally, voxels were identified for inclusion within each region's ROI mask by projecting half the distance of the cortical thickness at each vertex and labeling the intersected voxels. This method produced ROIs that reflected the underlying two-dimensional topology of the cortex and minimized the inclusion of voxels lying outside of gray matter.

Decoding analyses. Decoding analyses were performed using Scikit-Learn (Pedregosa et al., 2011) and other Python-based tools (Perez and Granger, 2007). Specifically, we used L2-penalized logistic regression models trained with LIBLINEAR (Fan et al., 2008). For all analyses, the regularization parameter was set to 1, its default value. These models were fit to preprocessed BOLD data from within the target ROIs; no additional feature selection was applied. Incorrect trials and those containing an fMRI artifact (see Image acquisition and preprocessing, above) were excluded from all decoding analyses. Decoding model matrices were normalized by run to zero mean and unit variance across samples. For all models aside from those during the cue period, the influence of reaction time was removed in a trialwise manner from each column of the model matrix using linear regression (Todd et al., 2013). In cases in which we decoded sets of task variables with more than two classes, we used a one-versus-rest strategy in which a set of binary models were fit to predict each class versus all others (Fan et al., 2008). For the time-resolved decoding analyses, we fit separate models for each of six consecutive time points beginning with the time point preceding the first stimulus onset in a trial. All other analyses were performed on BOLD data that had been averaged over the time points at 3 and 5 s after stimulus onset (accounting for the temporal interpolation during preprocessing). Decoding accuracy was computed with cross-validation over functional runs.

Within this general framework, we performed a number of secondary analyses to further understand the information underlying these decoding results, interactions between our ROIs, and how decoding performance was related to changes in the environment. Details of these analyses are explained in the following paragraphs. Briefly, we first attempted to quantify decoding performance in individual PFC regions. We then examined the spatial and temporal characteristics of the decoding models. Finally, we related continuous measures of classifier evidence in frontal and parietal ROIs to decoding performance in visual cortex and related decoding performance in all three regions to elements of the task structure that placed demands on cognitive control.

We used two related approaches to assess the significance of these model fits at the group and individual subject level. We first used a randomization procedure in which we randomly shuffled the class labels within run and refit the models 1000 times. This established an empirical distribution of accuracy scores under the null hypothesis where there is no association between BOLD activations and class labels (Ojala and Garriga, 2010). We used these null distributions to quantify the significance of the decoding results within each subject. To control for the presence of multiple comparisons across regions and time points, we used the same resampling to fit all models within a given iteration. We then constructed a distribution of the maximum accuracy across these
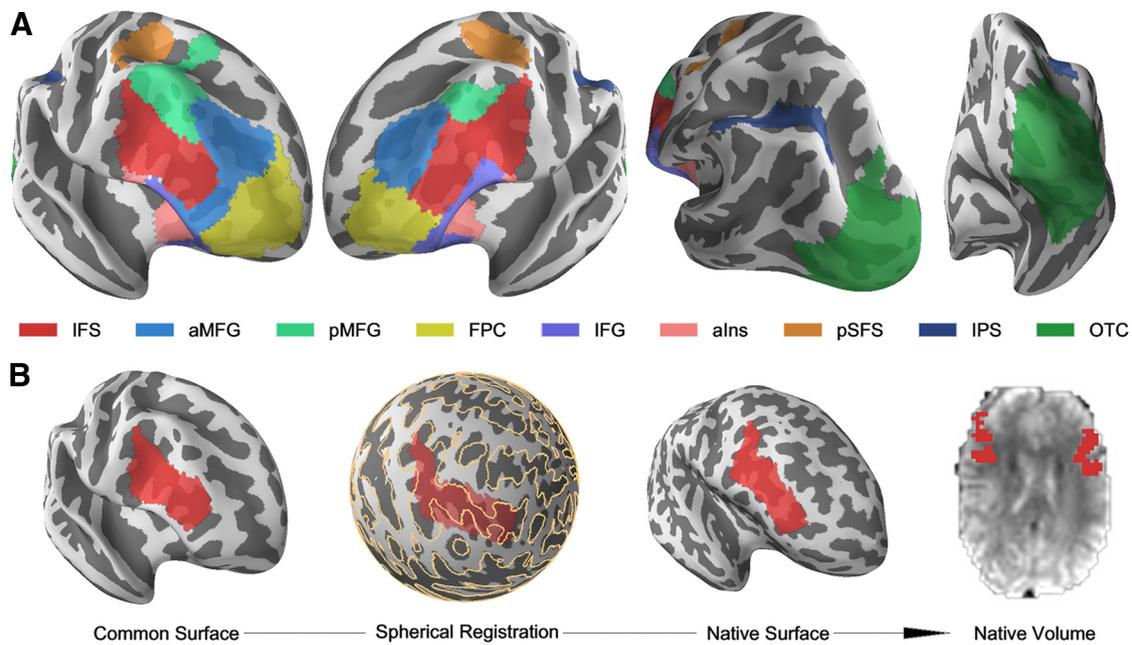
Figure 3. Targeted regions for decoding analyses. *A*, Source labels defining the ROIs on the Freesurfer average cortical surface. Please see the original reference (Yeo et al., 2011) for more information and additional visual perspectives. *B*, Schematic diagram of the procedure used to label voxels in native functional space for decoding analyses. The second panel shows contours of the template curvature plotted in the spherical coordinate system over the curvature from a representative subject. Darker areas indicate sulci. Cortical surfaces were visualized using PySurfer (http://pysurfer.github.io/).

Table 1. Region size and average BOLD signal intensity

	IFS	aMFG	pMFG	FPC	IFG	aIns	pSFS	IPS	OTC
No. of voxels									
Mean	616	560	461	584	528	352	275	407	1869
SD	84	68	60	62	57	38	57	48	224
Signal intensity									
Mean	116.51	118.78	117.50	102.56	99.55	118.01	116.80	138.11	117.82
SD	2.49	5.24	4.71	7.31	5.64	2.91	3.77	5.84	5.17

Shown is information about the ROIs used for decoding analyses. Means and SDs are across subjects. Signal intensities were scaled by run such that the median intensity across all four dimensions within the whole-brain mask is 100.

tests on each resample and used the percentiles in this distribution corresponding to our observed accuracies as *p*-values for those scores (Nichols and Holmes, 2002). The group tests used similar methods. Here, we first took the mean of the null distribution for each test as an empirical measure of chance and subtracted this from the observed accuracy. We then computed a one-sample *t* statistic across the group for each test. To obtain *p*-values for these statistics, we performed a randomization procedure with 100,000 iterations in which we multiplied all scores for each subject by either 1 or −1 before computing a *t* statistic for each test. We then used the distribution of maximum statistics across tests within each resample to determine the corrected *p*-value associated with our observed group mean accuracy scores (Nichols and Holmes, 2002).

To characterize how the information in our decoding models was organized spatially, we calculated the spatial autocorrelation of the model coefficients. These analyses were performed on the models in which features corresponded to BOLD activation averaged over the time points at 3 and 5 s after stimulus onset. For the multiclass analyses, each feature was associated with three weights corresponding to the three binary classifiers. We projected these weight maps back into the native functional image space masked by the ROI labels. For each binary model, we then computed a Pearson correlation between the original weight map and each of three maps that had been shifted one voxel in the *x*, *y*, and *z* directions. This produced nine correlation coefficients for each subject and ROI; we then took the mean of these nine values and fit mixed-effects regressions with the average correlation coefficient as a dependent measure to test for differences in autocorrelation across regions.

We also sought a precise estimate of when our decoding performance was greatest relative to stimulus onset. To compute this measure, we first upsampled the original BOLD time series to 500 ms resolution using cubic spline interpolation and repeated the decoding analysis on each new time point within the original window. We then fit gamma probability density functions (PDFs) to the resulting accuracy time courses using constrained least-squares optimization over five free parameters: the shape and scale parameters of the gamma distribution, onset time, baseline accuracy, and a single vertical scaling coefficient. These models were fit separately for subjects and ROIs and temporal inferences were made with mixed-effects regression using each peak time as a dependent variable.

To measure interactions between context representations in frontoparietal cortex and signatures of attentional influence in visual cortex, we exploited the fact that logistic regression provides a probabilistic estimate of the class labels for each sample in the testing dataset. In the multiclass case, the probabilistic estimates for each positive class in the set of binary classifiers were normalized to sum to 1. To obtain a measure of representational strength, we extracted the probability for each sample's target label; that is, on a "shape" trial, we used the probability associated with the "shape" label even though the classifier may have predicted "color" or "pattern." Because probabilities are bounded between 0 and 1, we applied a logit transform to these values before including them in our statistical models. We then fit mixed-effects logistic regressions predicting the accuracy of classifier predictions in visual cortex with the target logit values from frontoparietal cortex as independent variables. Note that this analysis is inherently correlative and does not establish the directionality of the relationship.
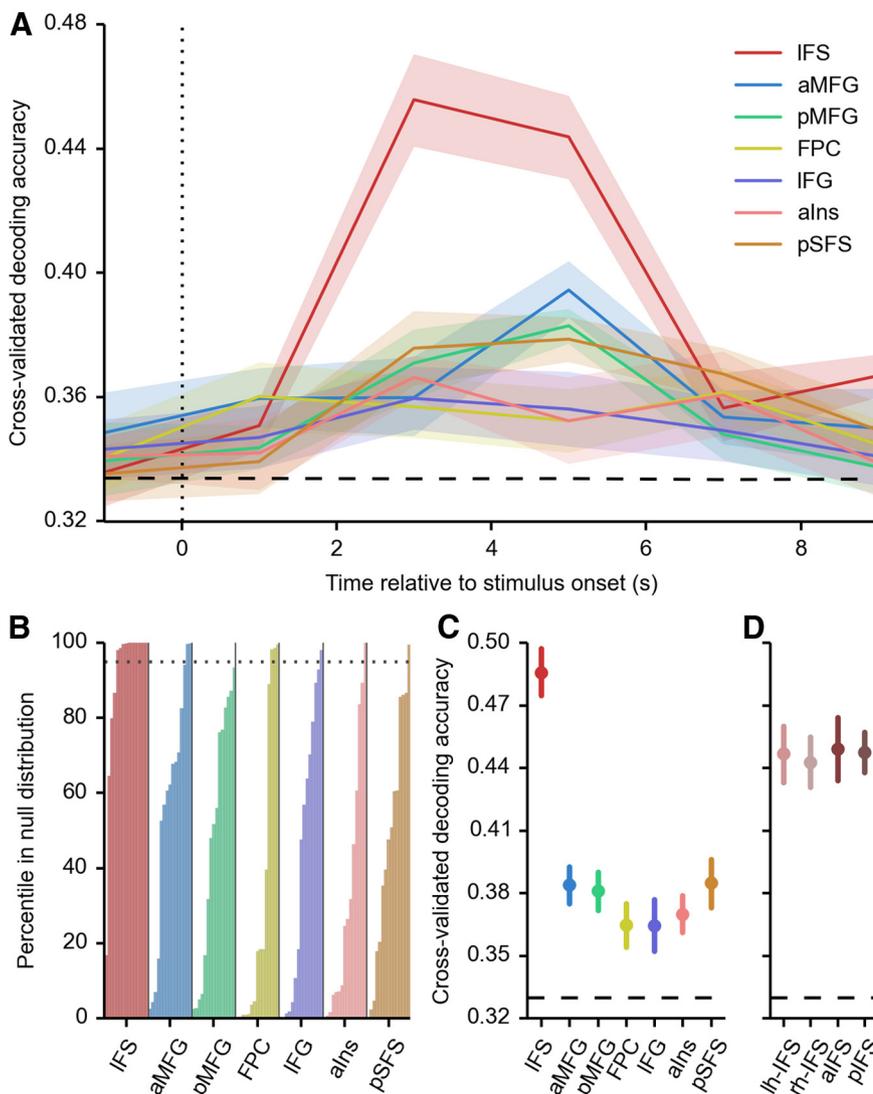
Figure 4. Decoding results for the dimension rules in prefrontal cortex. *A*, Time-resolved decoding results. Solid lines indicate mean decoding accuracy within each time bin and error bands denote bootstrapped SE across subjects. The horizontal dashed line shows the empirical measure of chance performance derived from the permutation analysis. Vertical dashed line is placed at the onset of the first stimulus. *B*, Height of each bar shows the percentile in the shuffled null distribution corresponding to the observed accuracy for each subject and region. Horizontal dashed line demarcates the criterion of significance at a (corrected) $\alpha = 0.05$. Bars are sorted by height within region. *C, D*, Decoding accuracies fit to data averaged across the time points at 3 and 5 s after stimulus onset. Points and error bars represent the mean and bootstrapped SE across subjects, respectively.

Results

Behavioral results

To ensure a high level of task performance during the experiment, participants were trained on the task until it was well learned outside of the scanner and were provided with monetary incentives to motivate performance during scanning. Collapsing across conditions, overall response accuracy during the main experiment was high (mean 95%, SD 3.2%), with all participants performing >87%. Throughout the fMRI analyses, we considered only trials with a correct response; the mean within-subject median reaction time (RT) on these trials was 826 ms (SD = 209 ms) with a mean interquartile range of 307 ms (SD = 101 ms).

We next considered how behavior was influenced by the dimension and decision rules (Fig. 2) using linear mixed-effects regression as implemented in the R package lme4 (http://lme4.r-forge.r-project.org/). For these models and all mixed-effects models reported herein, we used a maximal additive random-

effects structure and inferred on the fixed effects using likelihood ratio tests (Barr et al., 2013). We first examined response speed with a linear regression model predicting RT as a function of the dimension and decision rules (Fig. 2A). We found no evidence for an interaction between rule sets on RT ($\chi^2_2 = 0.57$; $p = 0.75$), so we inferred on the main effects within an additive model. There was no support for a main effect of dimension rule ($F = 0.42$; $\chi^2_2 = 1.52$; $p = 0.47$), but decision rules strongly influenced RT ($F = 26.05$; $\chi^2_1 = 15.09$; $p = 0.0001$), with a slower mean median RT of 94 ms (95% CI: 66.2 ms, 123 ms) when making a decision under the "different" rule relative to the "same" rule. We next examined the influence of task rules on response accuracy (Fig. 2B) using mixed-effects logistic regression. Here, we coded incorrect responses and failures to respond as errors. There was no support for an interaction between the two rule sets ($\chi^2_2 = 0.15$; $p = 0.972$), and we did not find evidence for a main effect of dimension rule on accuracy ($\chi^2_2 = 2.64$; $p = 0.27$); however, there was a weak main effect of decision rule ($\chi^2_1 = 3.85$; $p = 0.05$), with average accuracy decreasing by 1.9% (95% CI: 0.3%, 3.7%) when participants made a decision under the "different" rule relative to the "same" rule.

Although we found no evidence that average RT across the group varied with the dimension rules, it is possible that each subject exhibited idiosyncratic differences that could confound our within-subject fMRI analyses despite disappearing when pooling across subjects in the behavioral models (Todd et al., 2013). To assess whether this was the case, we performed a separate one-way ANOVA over the dimension rules for each subject. The median F score for these tests was 0.67 and only 2 of 15 subjects reached significance at $p < 0.1$, which indicated that the different dimension rules did not generally influence RT in an asymmetric manner. We nevertheless removed the effects of RT from the imaging data before undertaking decoding analyses (see Materials and Methods for details).

We also investigated whether response speed differed when the attended features of the stimuli matched relative to when they differed (Fig. 2C). We first tested the effect of attended feature matches directly; responses on these trials were numerically, but not significantly, slower ($\beta = 0.013$; SE = 0.0092; $\chi^2_1 = 1.99$; $p = 0.16$). We then regressed RT in an interactive model with predictors specifying the decision rule and whether the attended stimulus features matched. This model revealed a strong interaction ($\chi^2_1 = 79.04$; $p < 0.0001$) such that responses were fastest when the features matched under the "same" rule and slowest when they matched under the "different" rule, with nonmatch events falling in between.
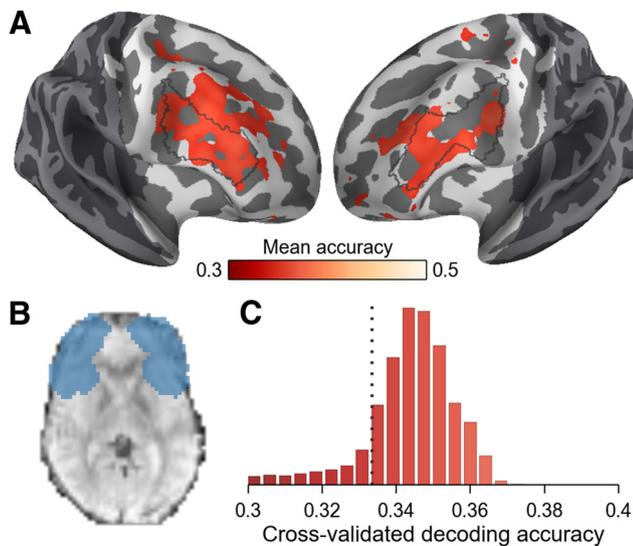
Figure 5. Searchlight analysis of the dimension rules in prefrontal cortex. *A*, Map of group-average searchlight accuracy after surface-based normalization and smoothing. The map is thresholded at $p < 0.005$ (uncorrected) from a group t test against expected chance. Voxels falling outside of the lateral PFC mask are dimmed and the IFS region used in ROI-based analysis is outlined in gray. *B*, Example slice through the native functional volume showing the search space. This region was defined by combining the masks for each individual PFC ROI and dilating the result by three voxels. Decoding models were fit within spheres of 10 mm radius and then the resulting accuracy maps were projected onto the surface for group testing. *C*, Distribution of searchlight accuracies within the PFC mask after group averaging.

Because trials were grouped into miniblocks, participants made decisions using the same rules at least three times before switching to a new set. We investigated whether rule repetition influenced response speed by fitting mixed models predicting RT with the number of trials since the last rule switch (Fig. 2E). These analyses did not reveal an effect of dimension rule repetition on RT ($\beta < 0.001$; SE $= 0.002$; $\chi^2_1 = 0.07$; $p = 0.80$), but there was a small effect for the decision rules ($\beta = -0.005$; SE $= 0.002$; $\chi^2_1 = 3.88$; $p = 0.049$). We also tested the influence of task switching directly by analyzing RT on the first trial of each miniblock as a function of whether the rules had switched. We found no evidence to support an interaction between rule type and task switches ($\chi^2_1 = 0.61$; $p = 0.44$) and neither the main effects of dimension ($\beta = 0.011$; SE $= 0.017$; $\chi^2_1 = 0.42$; $p = 0.52$) nor decision ($\beta = 0.020$; SE $= 0.017$; $\chi^2_1 = 1.39$; $p = 0.24$) switches were significant. Therefore, behavior was generally not influenced by rule repetition or task switching.

Context decoding in lateral PFC
To focus our decoding analyses, we used independent ROIs defined in a population atlas of task-independent cortical networks (Yeo et al., 2011). Although this parcellation was driven entirely by a clustering analysis of functional connectivity patterns, we will refer to the ROIs with approximate anatomical names for clarity. We first considered seven bilateral regions in lateral PFC (Fig. 3A, Table 1): inferior frontal sulcus (IFS), anterior and posterior middle frontal gyrus (aMFG and pMFG), frontopolar cortex, inferior frontal gyrus, anterior insula/frontal operculum (aIns), and posterior superior frontal sulcus (pSFS). These common-space ROIs were reverse-normalized to individual subject surfaces using a spherical transformation and projected into the native volume space to select features for the decoding datasets (Fig. 3B). The use of

surface-based ROIs and native space analyses provided increased precision in our efforts to understand the functional organization of the PFC.

We used linear classifier models (L2-penalized logistic regression) to predict task variables from patterns of BOLD signal in these ROIs. To characterize the temporal profile of context representation, we leveraged the slow event-related design in a time-resolved decoding analysis. Initial analyses focused on decoding the dimension rules from BOLD activation elicited in each PFC ROI during the stimulus period. That is, we aimed to decode from PFC whether the participants were attending to the color, shape, or pattern of the stimuli when they judged their relationship. As depicted in Figure 4, we achieved robust decoding of the dimension rules from the IFS; decoding performance in the other prefrontal ROIs was substantially lower, although in most cases, we observed a stimulus-evoked increase in accuracy (Fig. 4A).

To quantify the statistical significance of these results, we first compared the distribution of observed accuracies across the sample of participants against an empirical measure of chance performance. This analysis revealed that group mean accuracy reached significance at $p < 0.05$ (one tailed; corrected) for at least one time point in the IFS (mean $= 0.46$; time $= 3$ s; $t = 7.95$; $p < 0.0001$), aMFG (mean $= 0.39$; time $= 5$ s; $t = 6.60$; $p = 0.0002$), pMFG (mean $= 0.38$; time $= 5$ s; $t = 8.69$; $p < 0.0001$), aIns (mean $= 0.37$; time $= 3$ s; $t = 4.48$; $p = 0.0081$), and pSFS (mean $= 0.38$; time $= 5$ s; $t = 5.92$; $p = 0.0008$). Importantly, we also assessed the significance of the within-subject model fits using the results of our permutation analysis. Here, we found that the decoding models reached significance for 11 of 15 participants in the IFS, but for no more than three participants in any of the other lateral PFC regions (Fig. 4B).

We next verified that decoding performance in the IFS was statistically superior to that in other prefrontal regions. For these analyses, we averaged the BOLD signal across two time points at 3 and 5 s after stimulus onset before fitting new models to obtain a more robust estimate of the information in each ROI (Fig. 4C). We then performed paired t tests on decoding accuracy for all pairs of PFC regions. These tests showed that decoding was significantly greater in the IFS compared with all other PFC regions (all $p < 0.001$, corrected); we found no other differentiation at $p < 0.05$ (corrected). In addition, we compared accuracy in the IFS model with decoding performance from a region combining all seven lateral PFC regions. Peak accuracy in this collective ROI was significantly reduced relative to the more targeted IFS region (mean: 0.44, 95% CI: 0.41, 0.47; paired $t_{14} = 4.71$; $p = 0.0003$). Therefore, the representation of information related to the dimension rules within the PFC was primarily localized to the region spanning the IFS.

We furthermore found no evidence for a finer scale of organization within the IFS region (Fig. 4D). We examined the left and right ROIs separately and found similar performance in both hemispheres (left: mean $= 0.45$, 95% CI $= 0.42$, 0.47; right: mean $= 0.44$, 95% CI $= 0.42$, 0.47; paired $t_{14} = 0.29$; $p = 0.77$); accuracy in each lateralized ROI was relatively weaker than that for the full region (left: paired $t_{14} = 4.25$; $p = 0.0008$; right: paired $t_{14} = 5.31$; $p = 0.0001$). We also examined a possible rostrocaudal organization by evenly dividing each hemisphere perpendicular to its principle eigenaxis. Decoding performance between the anterior (aIFS) and posterior (pIFS) ROIs did not significantly differ (aIFS: mean $= 0.45$, 95% CI $= 0.42$, 0.48; pIFS: mean $= 0.45$; 95% CI $= 0.43$, 0.47; paired $t_{14} = 0.13$; $p = 0.90$), and each
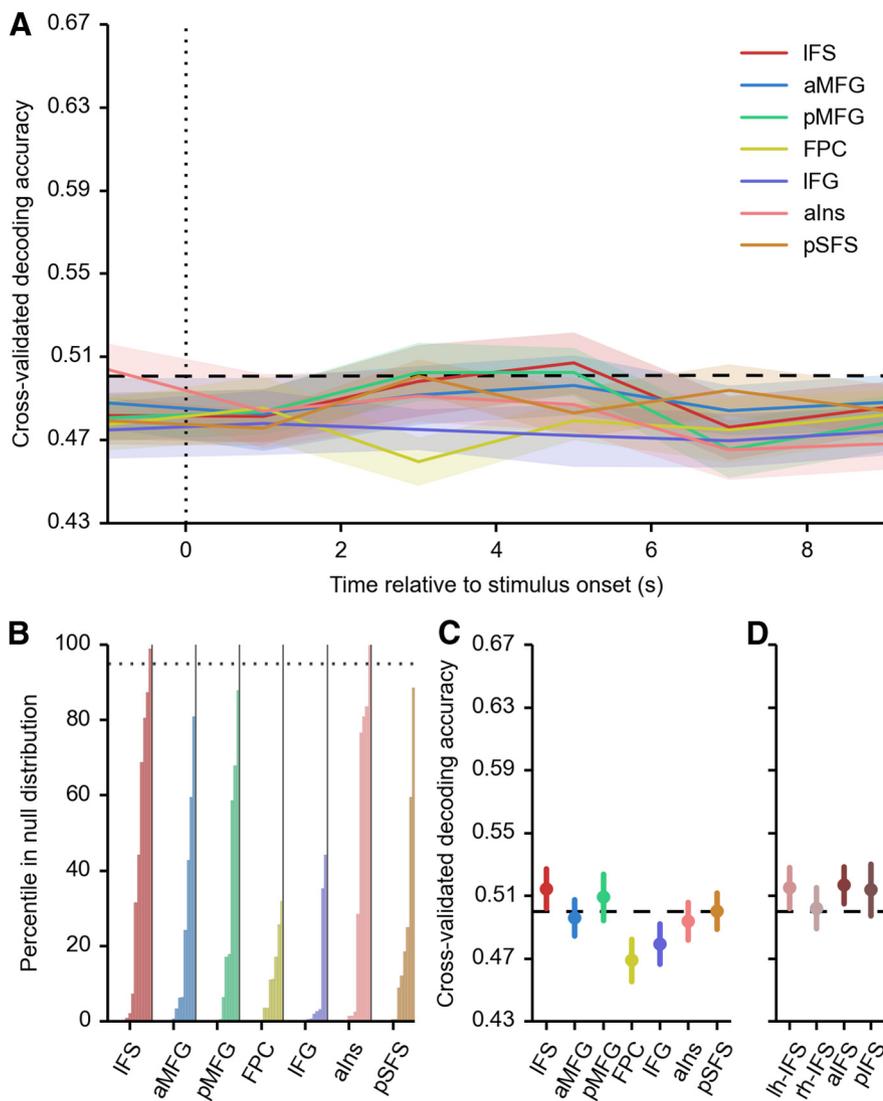
Figure 6. Decoding results for the decision rules in prefrontal cortex. Plot conventions are identical to those in Figure 4. The y-axis is scaled to span similar binomial probabilities as in Figure 4.

Participants may have verbally rehearsed the task rules during the stimulus period. If this were the case, our results could be due to phonological processes rather than abstract representations of task information. To assess this explanation, we reasoned that phonological processing would also be active during the cue period when subjects processed the instructions for that miniblock. We therefore conducted a time-resolved decoding analysis within the IFS during the cue period. This decoding was not successful, obtaining a maximum group mean accuracy of 0.36 at 5 s after cue onset ($p = 0.12$, corrected). The context decoding results thus depend on information that is specifically active during rule implementation and are unlikely to arise from verbal rehearsal of task rules.

We also considered the set of decision rules that specified whether participants should make "same" or "different" decisions about the relationship between the pair of stimuli on each trial. The ability to decode the decision rules was considerably diminished relative to the dimension rules (Fig. 6). All group mean tests were nonsignificant ($p > 0.9$, corrected) and only two individual subject models (one in the IFS and one in the aIns) reached significance (Fig. 6B). We thus found no evidence that distributed patterns of BOLD signal in the PFC carried information about the kind of decision that subjects made on each trial. For this reason, we focused on the dimension rules in all subsequent analyses.

Control signals in posterior neocortex

Prefrontal regions are coupled with other neocortical areas, and a central hypothesis about mechanisms of goal-directed behavior is that control emerges in part from interactions within these networks (Goldman-Rakic, 1987; Petersen and Posner, 2012). Guided by the particular focus on frontoparietal interactions (Goldman-Rakic, 1987), we next investigated whether patterns of activation in posterior parietal cortex discriminate task contexts dependent on the dimension rules. Returning to the atlas of task-independent networks (Yeo et al., 2011), we selected the lateral parietal component of the frontoparietal network, which is located within the intraparietal sulcus (IPS; Fig. 3A, Table 1).

We also sought to test the hypothesis that representations of context provide a source of bias on goal-relevant perceptual information (Desimone and Duncan, 1995; Miller and Cohen, 2001). We reasoned that biased processing might give rise to discriminative patterns in the visual cortex BOLD signal by amplifying the representation of the attended feature, allowing us to further investigate control processes as they play out across the neocortex. We therefore also extracted BOLD data from one of the two visual regions in the Yeo atlas (Yeo et al., 2011). This region covers the foveal components of early visual regions (V1–V3) and extends to midlevel areas on the lateral and ventral sur-

subregion again demonstrated relatively weaker accuracy compared with the full IFS (aIFS: paired $t_{14} = 3.82$; $p = 0.0019$; pIFS: paired $t_{14} = 4.14$; $p = 0.0010$). We furthermore fit a decoding model after collapsing the BOLD data to a single feature by taking the mean across voxels. Group mean decoding accuracy did not significantly exceed the permutation-estimated chance value (mean = 0.35; 95% CI = 0.34, 0.36; chance = 0.34; $p = 0.075$, one-tailed), and only two individual subject models significantly exceeded chance in a permutation test. The performance of the reduced dimensionality model was also markedly inferior to that of the full model (paired $t_{14} = 9.24$; $p < 0.0001$). Together, these analyses indicate that the decoding models exploited pattern information distributed across the IFS region.

To further validate these results, we performed a searchlight analysis within lateral PFC (Fig. 5). This provided additional evidence that information about the dimension rules was generally centered on, but not completely circumscribed by, the IFS region (Fig. 5A). Decoding performance in the searchlight analysis was considerably weaker than what we found using ROI-based methods (Fig. 5C), with the distribution of decoding accuracy in searchlight centers falling well below what we achieved when considering the entire IFS region.

faces; we will refer to it as the occipital temporal cortex (OTC; Fig. 3A, Table 1). In contrast to our expectations about frontoparietal cortex, we did not have a priori predictions about how informative signals might be distributed within or across subregions of visual cortex. We thus chose to consider visual cortex as a single unit and allow the classifiers to learn any discriminative information that could serve as a signature of top-down influence over visual processing. That information could take the form of changes either in the average activation across visual subregions or in distributed patterns across larger expanses of visual cortex.

We then repeated the time-resolved decoding analysis to determine whether the different contexts imposed by the dimension rules influenced processing within these regions. This revealed substantial context information (Fig. 7A,B), with significant group tests for both IPS ($p < 0.0001$) and OTC ($p < 0.0001$) and significant individual model performance for every subject in both ROIs. In our conception of goal-directed attention, control-related processes differ between these regions: frontoparietal areas maintain a more abstract representation of context that leads to biased processing of stimulus features within visual cortex. We thus sought to understand to what extent the information underlying our decoding results differed between these three regions.

We first considered the spatial scale of this information by computing the autocorrelation of the model coefficients when projected back into the functional image space. In regions with higher first-order autocorrelation, voxels that influence the classifier's prediction toward one particular dimension are more likely to be located adjacent to other voxels with similar preferences. An omnibus mixed-effects regression showed a significant difference across the three ROIs ($F = 42.6$; $\chi_2^2 = 28.5$; $p < 0.0001$; Fig. 7C). Pairwise comparisons determined that this effect was driven by higher autocorrelation in the OTC (mean $r = 0.25$; SD $= 0.34$) relative to the IFS (mean $r = 0.15$; SD $= 0.35$; paired $t_{14} = 8.84$; $p < 0.0001$) and IPS (mean $r = 0.15$; SD $= 0.43$; paired $t_{14} = 6.50$; $p < 0.0001$). In contrast, autocorrelation did not differ between the IFS and IPS (paired $t_{14} = 0.13$; $p = 0.90$). Visual inspection of the classifier weights aligned with this statistical analysis. Clusters of voxels with similar preferences were more easily observed in visual cortex than in frontoparietal cortex and these clusters appeared to be consistently organized across participants only in visual cortex. Therefore, information about the attended stimulus dimension was distributed more coarsely in visual cortex relative to frontoparietal regions.

We next considered the temporal profile of context decoding within this network. In the group-average data, the IFS and IPS decoding accuracy time courses peaked at 3 s after stimulus onset,
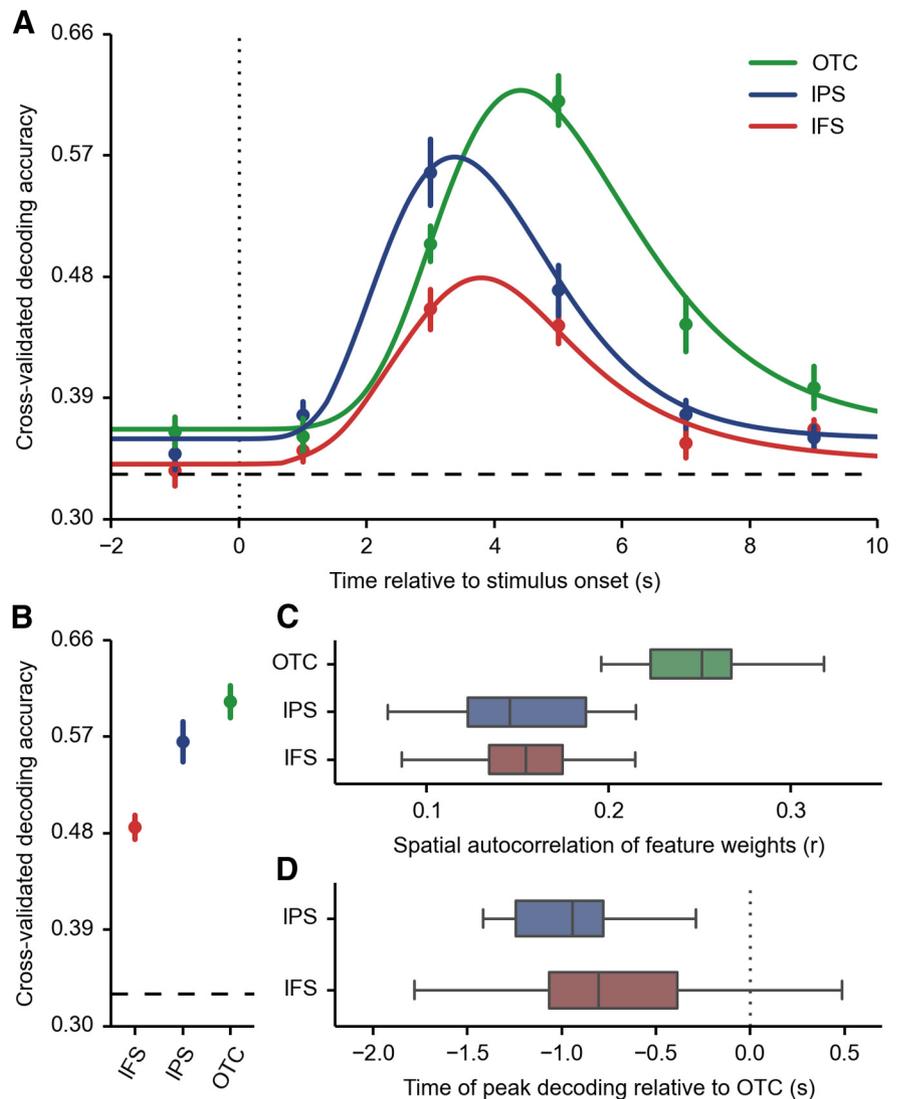

Figure 7. Decoding results for the dimension rules in posterior neocortex. **A**, Points and error bars show the mean and bootstrapped SE, respectively, for decoding accuracy in the original time bins. The solid traces show the gamma PDF models used to derive temporal information averaged across subjects. Plot conventions are otherwise as in Figure 4A. **B**, Plot conventions are as in Figure 4C. **C**, Boxplots showing the distribution of model coefficient autocorrelation across subjects sorted by region. **D**, Boxplots showing the distribution of relative differences in the time of peak decoding accuracy for the IFS and IPS models relative to the OTC models. Negative numbers indicate later peaks in the OTC.

whereas the OTC time course peaked at 5 s. For a more precise estimate, we fit gamma PDFs to upsampled decoding time courses (Fig. 7A). Overall, these models characterized the data well (IFS: median $R^2 = 0.69$; 95% CI $= 0.62, 0.87$; IPS: median $R^2 = 0.87$; 95% CI $= 0.72, 0.92$; OTC: median $R^2 = 0.95$; 95% CI $= 0.91, 0.96$). We then took the time at the peak of each PDF as an estimate of when there was maximal information about task rules relative to stimulus onset (IFS: mean $= 3.71$ s; 95% CI $= 3.47$ s, 3.95 s; IPS: mean $= 3.47$ s; 95% CI $= 3.22$ s, 3.74 s; OTC: mean $= 4.42$ s; 95% CI $= 4.19$ s, 4.64 s). Paired comparisons (Fig. 7D) confirmed that both the IFS (mean $= 0.71$ s; 95% CI $= 0.41$ s, 0.99 s; paired $t_{14} = 4.67$; $p = 0.0004$) and IPS (mean $= 0.95$ s; 95% CI $= 0.77$ s, 1.11 s; paired $t_{14} = 10.36$; $p < 0.0001$) peaked earlier than the OTC; there was not strong evidence for a difference between timing in the IFS and IPS (mean $= -0.24$ s, 95% CI $= -0.53$ s, 0.04 s; $t_{14} = 1.58$; paired $p = 0.14$). The OTC time course peaked later than the IPS time course in every subject and peaked later than the IFS time course in 13 subjects.
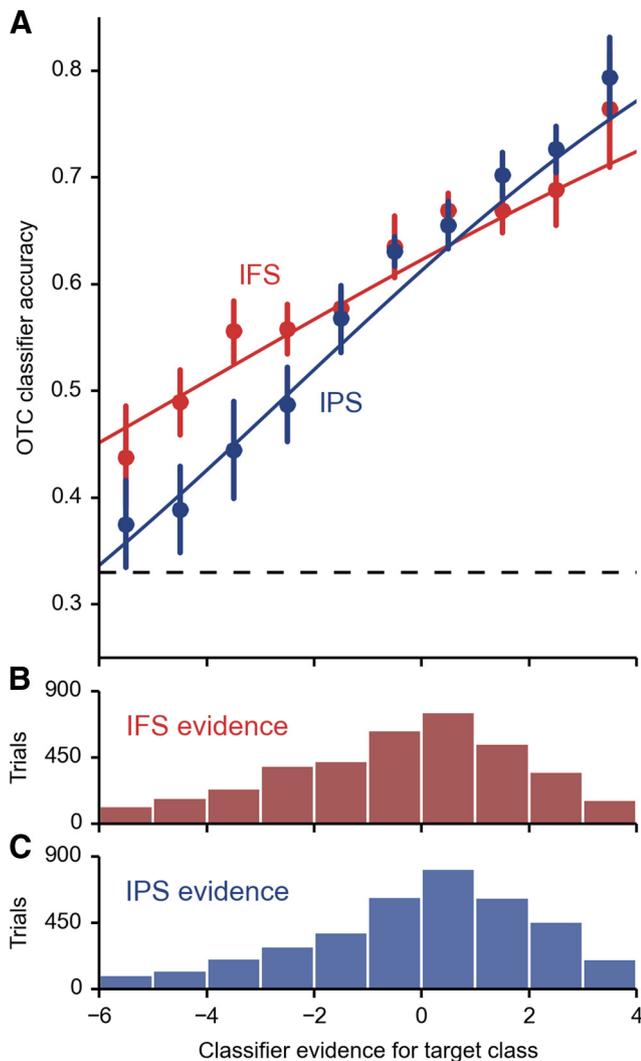
A

Figure 8. Relationship between frontoparietal classifier evidence and visual classifier performance. ***A***, Points show the mean OTC classifier accuracy sorted by binned IFS and IPS classifier evidence (the logit-transformed probability of the target class). Error bars represent the bootstrapped SE across subjects. Solid traces show the predictions of a logistic regression model fit to all data points. Horizontal dashed line shows chance performance. ***B***, ***C***, Histograms showing the distribution of classifier evidence in IFS and IPS for all trials and subjects; the *x*-axis and bins correspond across all three panels.

Interactions between frontoparietal and sensory representations

To directly test the theory that context representations in frontoparietal cortex support the biased processing of goal-relevant sensory information, we next measured the relationship between these regions with respect to our decoding measures. We reasoned that, if frontoparietal cortex supports top-down control over processing in visual regions, then the strength of the frontoparietal representations should predict the degree of attentional bias in visual cortex. We thus conducted a mixed-effects logistic regression that predicted whether the OTC classifier was accurate on each trial with the logit-transformed target class probabilities from the IFS and IPS models (Fig. 8). In this analysis, both the IFS ($\beta = 0.082$; SE $= 0.013$; $\chi_1^2 = 24.49$; $p < 0.0001$) and the IPS ($\beta = 0.18$; SE $= 0.019$; $\chi_1^2 = 29.10$; $p < 0.0001$) predictors were significantly greater than 0. Because we included both the IFS and IPS logit vectors in a single model, these parameter estimates reflect the unique ability of classifier strength in each area to predict

classifier accuracy in the OTC. Furthermore, we tested a linear contrast of these parameter estimates and found that that the regression coefficient for the IPS was significantly larger than that for the IFS ($\beta = 0.094$; SE $= 0.025$; $z = 3.82$; $p = 0.0001$). As a control, we used the realignment parameters to compute a root-mean-squared frame-to-frame displacement measure for the two time points on each trial underlying this analysis. Adding this measure of motion to the model did not change the results for the parameters of interest ($\beta_{IFS} = 0.082$; $\beta_{IPS} = 0.18$), suggesting that the relationship between frontoparietal and visual decoding was not being driven by motion-related image confounds. Therefore, the discriminability of context information in the IFS and IPS appears to predict the strength of attentional bias on goal-relevant stimulus features in visual cortex.

Control demand effects on context representation

Our final analyses concerned the relationship between demands on cognitive control and the neural measures identified in the preceding results. Control processes must be responsive to the environment and they are thought to be more necessary in the presence of shifting task demands (Keele and Rafal, 2000; Rossi et al., 2007). To understand how different levels of demand on control influence the information within this network of regions, we examined the effects of making repeated decisions under the same dimension rule. We extracted the trialwise prediction accuracies from the dimension rule classifiers in IFS, IPS, and OTC and then fit a mixed-effects logistic regression predicting these variables with the log-transformed number of trials that had elapsed since the last rule switch (Fig. 9). We restricted this analysis to a maximum of six repetitions because performing three consecutive miniblocks under the same dimension rule was rare and we correspondingly had very few data points for these values. We found no evidence supporting an interaction between region and repetition ($\chi_2^2 = 2.18$; $p = 0.34$), but we found a strong main effect of repetition ($\beta = -0.20$; SE $= 0.034$; $\chi_1^2 = 18.71$; $p < 0.0001$) such that classifier prediction accuracy decreased over repeated trials with the same rule. This effect was not explained by changes in mean signal amplitude across the ROIs. We also examined the effects of including the classifier evidence from IFS and IPS in a model predicting OTC accuracy with trial position. The IFS ($\beta = 0.085$; SE $= 0.014$; $\chi_1^2 = 21.19$; $p < 0.0001$) and IPS ($\beta = 0.17$; SE $= 0.019$; $\chi_1^2 = 28.83$; $p < 0.0001$) remained strong predictors of OTC classifier accuracy; importantly, they also mediated the effect of trial position, which dropped below significance ($\beta = -0.10$; SE $= 0.072$; $\chi_1^2 = 1.81$; $p = 0.18$). These analyses provide further evidence that frontoparietal areas adapt in the face of demands on cognitive control to implement a top-down bias on perceptual processes and support goal-directed attention.

Discussion

Cognitive control allows humans to interact flexibly with their environment in the service of diverse goals. Here, we present empirical measurements of frontoparietal representations that facilitate top-down control. Our data provide several novel insights about the mechanisms that align sensorimotor processing with shifting task demands. In the PFC, we found that task information was transiently represented during decision making within a region running along the inferior frontal sulcus, which was defined using an atlas of task-independent functional networks. We observed similar results within the IPS, a parietal region that is thought to be functionally coupled with the IFS. Patterns of activation in visual cortex contained discriminative

information about which stimulus dimension was being attended while visual input was held constant, a signature of top-down influence over perception. Within this network of regions, we found evidence that context representation was graded in strength with several consequences for control implementation. Classifier performance was modulated by control demands that emerged from the task structure such that context discriminability gradually decayed after a rule shift; correspondingly, the strength of context representation in both frontal and parietal cortex independently predicted the degree of attentional bias in visual cortex.

To understand how the PFC enables cognitive control, it is important to determine its functional organization. One prominent class of theories divides lateral PFC along a coarse dorsal–ventral division (Courtney, 2004; Petrides, 2005; O'Reilly, 2010). In contrast, measurements of task-independent functional connectivity across a range of analytical frameworks have identified an intermediate prefrontal region spanning the IFS; at rest and during task-directed behavior, this region couples functionally with the IPS (Vincent et al., 2008; Power et al., 2011; Yeo et al., 2011). The correspondence between these regions and focal activations in task-based experiments suggests that they comprise a control or executive network (Nelson et al., 2010; Power et al., 2011). Our data support this perspective. Decoding accuracy in the IFS exceeded the performance of every other prefrontal ROI, and information about task context appeared to be distributed throughout it. Elsewhere in PFC, average decoding performance generally rose above chance, but individual models tended to be unreliable. Because these other regions surround the IFS, this could reflect errors introduced by using macroanatomical structure to transform functional ROIs into native space. It is also important to note that the ROI definitions in the Yeo atlas (Yeo et al., 2011) are only an estimate of functional organization. The exact boundaries of the regions are sensitive to different analytical choices, and the true underlying organization of PFC may be dynamic and context dependent. Nevertheless, our data show that patterns of functional connectivity provide an informative definition of control regions in frontoparietal cortex. This allowed us to explore the relationship among frontoparietal context representation, attentional effects in visual cortex, and changes in task demands. Collectively, these results demonstrate that a combination of task-independent and task-based methods can elucidate the functions of human PFC.

Our experiment's event-related design further allowed us to explore the temporal dynamics of context representation. PFC neurons exhibit sustained, distractor-resistant activity in the absence of sensory input (Miller et al., 1996), which is often thought to critically underlie the PFC's role in cognitive control (Rougier et al., 2005). Contrary to this emphasis on persistent neuronal firing, we did not find evidence that context representations were sustained between decision events within a miniblock. These results partially diverge from the model of PFC organization advanced by Dosenbach et al. (2007), in which the frontoparietal
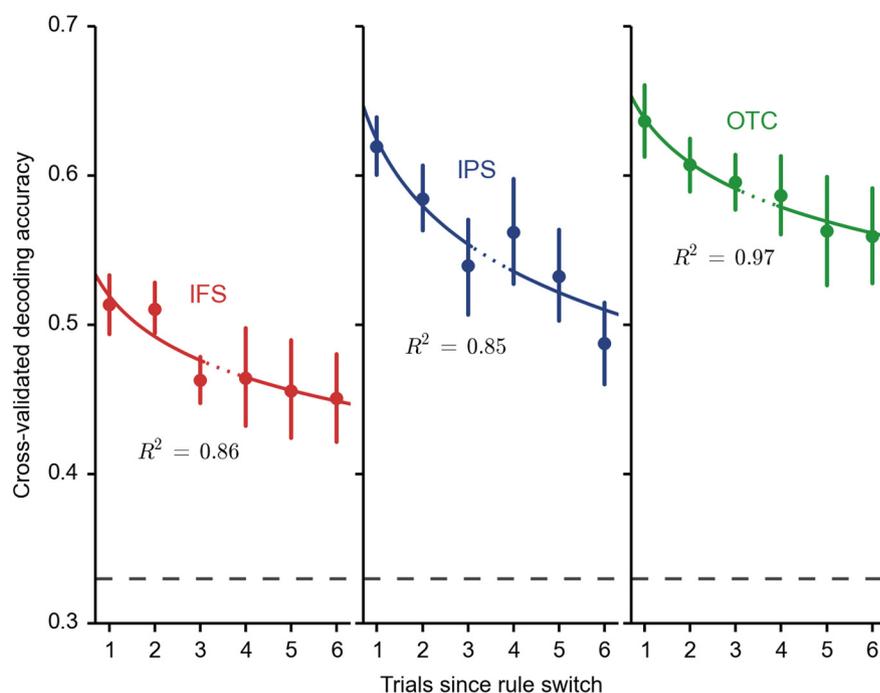

Figure 9. Relationship between control demands and dimension rule decoding. Points and error bars show means and bootstrapped SEs across subjects, respectively. Solid traces show the predictions of a log-linear fit to all data points. Traces are dashed between trials 3 and 4 to indicate passage through a miniblock boundary. Horizontal dashed line shows chance performance.

and cinguloopercular networks support independent control processes with, respectively, transient and sustained dynamics. Although the transient frontoparietal response is consistent with this model, the context information predictably spanned miniblocks and should thus correspond with stable task control. However, information about context was only weakly present in the aIns and aMFG, which are part of the cinguloopercular system. Intriguingly, there is evidence that different contexts can remain separated in neuronal state space even while average population activity settles to baseline between trials (Stokes et al., 2013). It is thus possible that subtle differences in baseline firing maintain a memory trace for the current context but do not drive the hemodynamic response powerfully enough to support reliable decoding from BOLD data in between stimulus events.

In our task, the context for each decision was specified by explicit rules presented at the start of each block. Neurons in the macaque PFC show rule-selective responses in similar tasks (Wallis et al., 2001; Buschman et al., 2012) and we suspect that such selectivity is a major component of the PFC representation underlying our decoding results. PFC neurons also exhibit selective responses to specific stimuli and task contexts that are not explicitly instructed, often with diverse responses mixed within individual cells (Mante et al., 2013; Rigotti et al., 2013; Stokes et al., 2013). Although rules provide efficient structure for goal-directed behavior, control over decision making both in the real world and in the laboratory can rely on diverse forms of information. We thus do not think of the frontoparietal control network as being specialized for representing rules per se. Instead, the frontoparietal representation separates task variables to the extent that they effectively parameterize different behavioral strategies given environmental demands and internal goals. These representations are likely constructed and tuned by reinforcement learning mechanisms that gate information into and out of the context representation based on whether the use of that in-
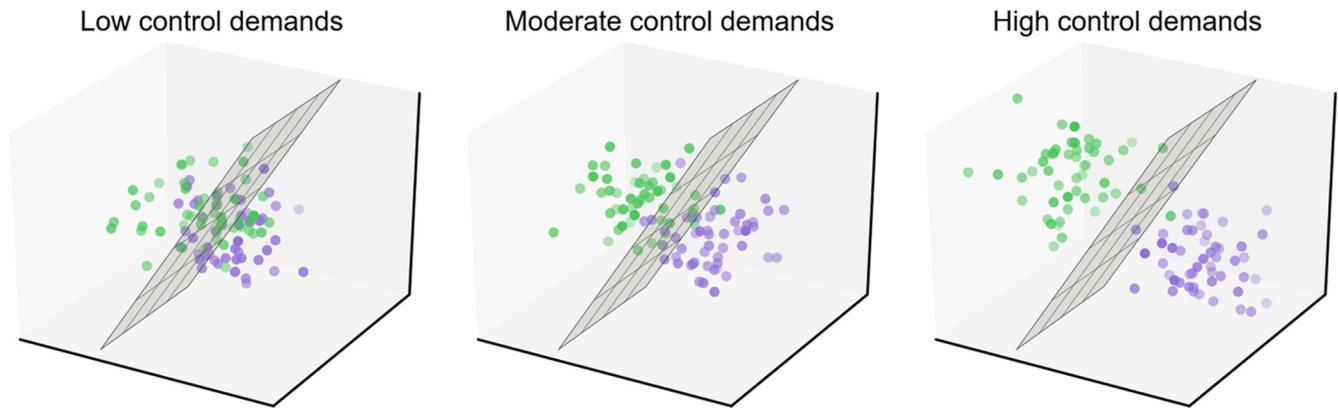
Figure 10. Schematic diagram of the relationship between control demands and context representation. Points illustrate high-dimensional patterns of activity across a population of voxels (or neurons) on each trial sorted into different contexts by color. The plane implements a linear classifier, which portrays both our analytical approach and the hypothesized readout mechanism of downstream areas (Rigotti et al., 2013). As control demands increase, patterns associated with different contexts grow more separated in the neural state space, which leads to increased classification accuracy and greater top-down control.

formation leads to rewarding outcomes (Braver and Cohen, 2000; Badre et al., 2010).

It is surprising, then, that we found no evidence that populations of voxels in PFC discriminated the decision rules. To ensure that our results would not be driven by differences in rule difficulty or complexity, we corrected all analyses for RT (Todd et al., 2013). We can thus conclude only that no residual activation, after removing the effects of RT, contained information about the decision rules. Although previous work has conceptualized task rules as providing a conditional input–output map that links sensory input to action selection (Bunge, 2004), the striking interaction in our data between decision rule and feature matching suggests that the subjects did not implement the set of decision rules in this manner. Nevertheless, because the decision rules decoupled the processing of stimulus information from response selection, our dimension-rule-decoding results demonstrate that frontoparietal representations underlying control can be independent of the action plan that emerges from controlled decision making.

Both spatial and temporal characteristics dissociated the IFS and IPS models from those in OTC. It is important to note that these results may reflect only differences in the vasculature and hemodynamic coupling across these regions (Handwerker et al., 2004; Logothetis and Wandell, 2004). Alternatively, the temporal dissociation could be accounted for by our slow event-related design and the anticipatory cue, which allowed for preparatory set in advance of stimulus processing. Spatially, single-unit neurophysiology experiments typically find that neurons with selectivity to opposing rules are intermingled throughout prefrontal areas (Buschman et al., 2012), whereas attentional bias in sensory cortex likely increases activation in subregions that preferentially represent different stimulus dimensions and thus yield the coarser organization we observed in OTC. These observations could be strengthened by future experiments that provide for more targeted analyses in visual cortex.

We found that task switching influenced context representation such that decoding performance declined over the trials immediately after a rule switch. It is well established that the PFC makes essential contributions to behavior in shifting environments, but its exact role has remained unclear (Keele and Rafal, 2000; Buchsbaum et al., 2005). According to one view, increased prefrontal activation on switch trials reveals a process of reconfiguration (Konishi et al., 1998) that has been proposed as a source of behavioral switching costs (Monsell, 2003). In our ex-

periment, we did not observe a behavioral switching effect, likely due to the long delay between the task cue and first stimulus event of each block. Instead, we argue that, independently of reconfiguration, proximity to a switch places demands on cognitive control; our results indicate that humans can adjust their attentional processes on a relatively short time scale after a switch such that there is less need for a separation between rules in the frontoparietal context representation (Fig. 10). These data clarify the mechanisms underlying findings on the effects of PFC lesions in humans (Keele and Rafal, 2000) and macaques (Rossi et al., 2007), along with characteristic patterns of BOLD activation in task-switching experiments (Braver et al., 2003; Brass and Cramon, 2004; Badre and Wagner, 2006; De Baene et al., 2012).

More broadly, this result supports the perspective that human cognition exists on a continuum of automaticity (Cohen et al., 1990), which stands in contrast to models that posit a strong distinction between automatic and controlled behavior. In our account, increased engagement of control sharpens the abstract representations of variables that are critical for performance of the current task. This flexible discriminability leads to enhanced lower-level processing of relevant environmental information and controls the selective gating of that information into decision making and action selection. Through these mechanisms, human cognition adapts to a broad variety of natural environments and produces intelligent behavior.

Notes

Supplemental material for this article is available at https://github.com/WagnerLabPapers/Waskom_JNeurosci_2014. This repository contains all code that was used to produce the results and figures in the paper. This material has not been peer reviewed.

References

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. CrossRef Medline

Badre D, Wagner AD (2006) Computational and neurobiological mechanisms underlying cognitive flexibility. Proc Natl Acad Sci U S A 103:7186–7191. CrossRef Medline

Badre D, Kayser AS, D'Esposito M (2010) Frontal cortex and the discovery of abstract action rules. Neuron 66:315–326. CrossRef Medline

Barr DJ, Levy R, Scheepers C, Tily HJ (2013) Random effects structure for confirmatory hypothesis testing: keep it maximal. Journal of Memory and Language 68:255–278. CrossRef

Brass M, von Cramon DY (2004) Selection for cognitive control: a func-

tional magnetic resonance imaging study on the selection of task-relevant information. J Neurosci 24:8847–8852. CrossRef Medline

Braver TS, Cohen JD (2000) On the control of control: the role of dopamine in regulating prefrontal function and working memory. In: Attention and performance XVIII: control of cognitive processes, Ed 18 (Monsell S, Driver JS, eds), pp 713–737. Cambridge, MA: MIT.

Braver TS, Reynolds JR, Donaldson DI (2003) Neural mechanisms of transient and sustained cognitive control during task switching. Neuron 39: 713–726. CrossRef Medline

Buchsbaum BR, Greer S, Chang WL, Berman KF (2005) Meta-analysis of neuroimaging studies of the Wisconsin card-sorting task and component processes. Hum Brain Mapp 25:35–45. CrossRef Medline

Buckley MJ, Mansouri FA, Hoda H, Mahboubi M, Browning PG, Kwok SC, Phillips A, Tanaka K (2009) Dissociable components of rule-guided behavior depend on distinct medial and prefrontal regions. Science 325: 52–58. CrossRef Medline

Bunge SA (2004) How we use rules to select actions: a review of evidence from cognitive neuroscience. Cogn Affect Behav Neurosci 4:564–579. CrossRef Medline

Buschman TJ, Denovellis EL, Diogo C, Bullock D, Miller EK (2012) Synchronous oscillatory neural ensembles for rules in the prefrontal cortex. Neuron 76:838–846. CrossRef Medline

Chiu YC, Esterman M, Han Y, Rosen H, Yantis S (2011) Decoding task-based attentional modulation during face categorization. J Cogn Neurosci 23:1198–1204. CrossRef Medline

Cohen JD, Dunbar K, McClelland JL (1990) On the control of automatic processes: a parallel distributed processing account of the Stroop effect. Psychol Rev 97:332–361. CrossRef Medline

Cole MW, Etzel JA, Zacks JM, Schneider W, Braver TS (2011) Rapid transfer of abstract rules to novel contexts in human lateral prefrontal cortex. Front Hum Neurosci 5:142. CrossRef Medline

Cole MW, Reynolds JR, Power JD, Repovs G, Anticevic A, Braver TS (2013) Multi-task connectivity reveals flexible hubs for adaptive task control. Nat Neurosci 16:1348–1355. CrossRef Medline

Courtney SM (2004) Attention and cognitive control as emergent properties of information representation in working memory. Cogn Affect Behav Neurosci 4:501–516. CrossRef Medline

Dale AM, Fischl B, Sereno MI (1999) Cortical surface-based analysis I–segmentation and surface reconstruction. Neuroimage 9:179–194. CrossRef Medline

De Baene W, Kühn S, Brass M (2012) Challenging a decade of brain research on task switching: brain activation in the task-switching paradigm reflects adaptation rather than reconfiguration of task sets. Hum Brain Mapp 33:639–651. CrossRef Medline

Desimone R, Duncan J (1995) Neural mechanisms of selective visual attention. Annu Rev Neurosci 18:193–222. CrossRef Medline

Dosenbach NU, Fair DA, Miezin FM, Cohen AL, Wenger KK, Dosenbach RAT, Fox MD, Snyder AZ, Vincent JL, Raichle ME, Schlaggar BL, Petersen SE (2007) Distinct brain networks for adaptive and stable task control in humans. Proc Natl Acad Sci U S A 104:11073–11078. CrossRef Medline

Egner T, Hirsch J (2005) Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. Nat Neurosci 8:1784–1790. CrossRef Medline

Fan R-E, Chang K-W, Hsieh C-J, Wang X-R, Lin C-J (2008) LIBLINEAR: a library for large linear classification. The Journal of Machine Learning Research 9:1871–1874.

Fischl B, Sereno MI, Tootell RB, Dale AM (1999) High-resolution intersubject averaging and a coordinate system for the cortical surface. Hum Brain Mapp 8:272–284. CrossRef Medline

Goldman-Rakic PS (1987) Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In Handbook of physiology, Section 1: The nervous system, Vol 5: Higher functions of the brain (Plum F, ed), pp 373–416. Bethesda, MD: American Physiological Society.

Gorgolewski K, Burns CD, Madison C, Clark D, Halchenko YO, Waskom ML, Ghosh SS (2011) Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. Front Neuroinform 5:13. CrossRef Medline

Greve DN, Fischl B (2009) Accurate and robust brain image alignment using boundary-based registration. Neuroimage 48:63–72. CrossRef Medline

Handwerker DA, Ollinger JM, D'Esposito M (2004) Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. Neuroimage 21:1639–1651. CrossRef Medline

Keele SW, Rafal RD (2000) Deficits of task set in patients with left prefrontal cortex lesions. In: Control of cognitive processes: attention 60 and performance XVIII, 18 Ed (Monsell S, Driver JS, eds), pp 627–651. Cambridge, MA: MIT.

Konishi S, Nakajima K, Uchida I, Kameyama M, Nakahara K, Sekihara K, Miyashita Y (1998) Transient activation of inferior prefrontal cortex during cognitive set shifting. Nat Neurosci 1:80–84. CrossRef Medline

Logothetis NK, Wandell BA (2004) Interpreting the BOLD signal. Annu Rev Physiol 66:735–769. CrossRef Medline

MacDonald AW 3rd, Cohen JD, Stenger VA, Carter CS (2000) Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. Science 288:1835–1838. CrossRef Medline

Mante V, Sussillo D, Shenoy KV, Newsome WT (2013) Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature 503:78–84. CrossRef Medline

Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. Annu Rev Neurosci 24:167–202. CrossRef Medline

Miller EK, Erickson CA, Desimone R (1996) Neural mechanisms of visual working memory in prefrontal cortex of the macaque. J Neurosci 16: 5154–5167. Medline

Milner B (1963) Effects of different brain lesions on card sorting: the role of the frontal lobes. Archives of Neurology 9:90–100. CrossRef

Monsell S (2003) Task switching. Trends Cogn Sci 7:134–140. CrossRef Medline

Nelissen N, Stokes M, Nobre AC, Rushworth MF (2013) Frontal and parietal cortical interactions with distributed visual representations during selective attention and action selection. J Neurosci 33:16443–16458. CrossRef Medline

Nelson SM, Cohen AL, Power JD, Wig GS, Miezin FM, Wheeler ME, Velanova K, Donaldson DI, Phillips JS, Schlaggar BL, Petersen SE (2010) A parcellation scheme for human left lateral parietal cortex. Neuron 67: 156–170. CrossRef Medline

Nichols TE, Holmes AP (2002) Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum Brain Mapp 15:1–25. CrossRef Medline

Ojala M, Garriga GC (2010) Permutation tests for studying classifier performance. The Journal of Machine Learning Research 11:1833–1863.

O'Reilly RC (2010) The what and how of prefrontal cortical organization. Trends Neurosci 33:355–361. CrossRef Medline

Pedregosa F et al. (2011) Scikit-learn: Machine learning in Python. The Journal of Machine Learning Research 12:2825–2830.

Perez F, Granger BE (2007) IPython: a system for interactive scientific computing. Computing in Science and Engineering 9:21–29. CrossRef

Petersen SE, Posner MI (2012) The attention system of the human brain: 20 years after. Annu Rev Neurosci 35:73–89. CrossRef Medline

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

Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, Vogel AC, Laumann TO, Miezin FM, Schlaggar BL, Petersen SE (2011) Functional network organization of the human brain. Neuron 72:665–678. CrossRef Medline

Reverberi C, Görgen K, Haynes JD (2012) Distributed representations of rule identity and rule order in human frontal cortex and striatum. J Neurosci 32:17420–17430. CrossRef Medline

Rigotti M, Barak O, Warden MR, Wang XJ, Daw ND, Miller EK, Fusi S (2013) The importance of mixed selectivity in complex cognitive tasks. Nature 497:585–590. CrossRef Medline

Rossi AF, Bichot NP, Desimone R, Ungerleider LG (2007) Top down attentional deficits in macaques with lesions of lateral prefrontal cortex. J Neurosci 27:11306–11314. CrossRef Medline

Rougier NP, Noelle DC, Braver TS, Cohen JD, O'Reilly RC (2005) Prefrontal cortex and flexible cognitive control: Rules without symbols. Proc Natl Acad Sci U S A 102:7338–7343. CrossRef Medline

Sakai K, Passingham RE (2003) Prefrontal interactions reflect future task operations. Nat Neurosci 6:75–81. CrossRef Medline

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Mat-

thews PM (2004) Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 23:S208–S219. CrossRef Medline

Stokes MG, Kusunoki M, Sigala N, Nili H, Gaffan D, Duncan J (2013) Dynamic coding for cognitive control in prefrontal cortex. Neuron 78:364–375. CrossRef Medline

Todd MT, Nystrom LE, Cohen JD (2013) Confounds in multivariate pattern analysis: theory and rule representation case study. Neuroimage 77:157–165. CrossRef Medline

Vincent JL, Kahn I, Snyder AZ, Raichle ME, Buckner RL (2008) Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. J Neurophysiol 100:3328–3342. CrossRef Medline

Wallis JD, Anderson KC, Miller EK (2001) Single neurons in prefrontal cortex encode abstract rules. Nature 411:953–956. CrossRef Medline

Woolgar A, Hampshire A, Thompson R, Duncan J (2011) Adaptive coding of task-relevant information in human frontoparietal cortex. J Neurosci 31:14592–14599. CrossRef Medline

Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zöllei L, Polimeni JR, Fischl B, Liu H, Buckner RL (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J Neurophysiol 106:1125–1165. CrossRef Medline

Zanto TP, Rubens MT, Thangavel A, Gazzaley A (2011) Causal role of the prefrontal cortex in top-down modulation of visual processing and working memory. Nat Neurosci 14:656–661. CrossRef Medline

Zhang J, Kriegeskorte N, Carlin JD, Rowe JB (2013) Choosing the rules: distinct and overlapping frontoparietal representations of task rules for perceptual decisions. J Neurosci 33:11852–11862. CrossRef Medline