Published in final edited form as:
Neuroimage. 2012 November 15; 63(3): 1285–1294. doi:10.1016/j.neuroimage.2012.08.034.

Rostral-Caudal Gradients of Abstraction Revealed by Multi-Variate Pattern Analysis of Working Memory

Derek Evan Nee[1,2] and **Joshua W. Brown**[1]

[1]Dept. of Psychological & Brain Sciences, Indiana University, Bloomington, IN 47405

[2]Dept. of Psychology, University of Michigan, Ann Arbor, MI 48109

Abstract

The lateral frontal cortex (LFC) is thought to represent contextual and rule-based information that allows adaptive behavior according to circumstance. Recent progress has suggested that the representations of the LFC vary along its rostral-caudal axis with more abstract, higher level representations associated with rostral areas of the LFC and more concrete, lower level representations associated with caudal areas of the LFC. Here, we investigated this proposal. Subjects responded to stimuli based upon a nested series of contextual cues stored in working memory (WM) while being scanned with fMRI. Higher level context cues denoted an abstract rule set while lower level context cues provided more concrete information. Using multi-variate pattern analysis (MVPA), we found varying forms of representation along the rostral-caudal axis of the LFC depending on the type of information stored in WM. Rostral areas of frontal cortex in the lateral orbitofrontal cortex (OFC) represented the higher level context, but not more concrete information, and only when more concrete information was unavailable. Mid-level areas in the mid-dorsolateral prefrontal cortex (DLPFC) and inferior frontal junction (IFJ) represented more concrete rules, but only when the forthcoming response could not be anticipated. By contrast, the dorsal premotor cortex (PMd) and primary motor cortex (M1) represented contextual and response information when the forthcoming response could be anticipated on the basis of context. Collectively, these data indicate that representations dedicated to higher levels of abstraction become less discriminating when more concrete information becomes available. These patterns are consistent with rostral-caudal abstraction proposals of the LFC.

Keywords

prefrontal cortex; PFC; frontal lobe; cognitive control; executive function; hierarchical

Introduction

A distinguishing feature of intelligent behavior is the ability to act flexibly based upon internally stored contexts. When faced with identical overt stimuli, humans can use contextual information to select the most appropriate action. The ability to use context to guide performance is thought to depend upon the lateral frontal cortex (LFC; Miller and Cohen, 2001). Commensurately, a great deal of research has investigated the organizational

Address Correspondence To: Derek Evan Nee, 1101 E Tenth St., Bloomington, IN 47405, denee@indiana.edu.

and representational properties of the LFC that afford flexible behavior (Badre and D'Esposito, 2009; Koechlin and Summerfield, 2007; Miller and Cohen, 2001; O'Reilly, 2010). Despite intense interest, the organization of the LFC remains elusive.

Recent theories have attempted to describe organizing principles of the LFC that enable flexible behavior. At the heart of many of these theories is the idea that representations in the LFC vary along a rostral-caudal axis (Badre and D'Esposito, 2007; Fuster, 1997; Koechlin and Summerfield, 2007). Under these theories, rostral areas of the LFC are thought to represent more abstract, higher level content that influences distant actions. Conversely, caudal areas of the LFC are thought to represent more concrete, lower level content that influences proximate actions. These theories are supported by anatomical data that demonstrate that rostral areas of the LFC are primarily connected to caudal areas of the LFC (Petrides and Pandya, 2007), which are in turn connected to more posterior cortices (Cavada and Goldman-Rakic, 1989; Petrides and Pandya, 1984, 1999). This organizational structure places the rostral-most regions of the LFC in position to either exert top-down control over more posterior regions of the LFC, integrate information represented in posterior regions of the LFC, or both (Badre and D'Esposito, 2009).

Further support for rostral-caudal frontal gradients of abstraction comes from recent functional magnetic resonance imaging (fMRI) studies. In a landmark study, Koechlin and colleagues (2003) examined fMRI activation while subjects responded to color and letter stimuli. The authors varied whether responses could be determined on the basis of the stimuli (sensory control), on the basis of contextual cues (contextual control), or on the basis of a combination of cues (episodic control). While sensory control was associated with activation in caudal LFC (area 6), contextual control was associated with activation more rostrally (areas 44 and 45), and episodic control was associated with even more rostral activation (area 46). These results were interpreted in terms of a cascade model whereby more rostral regions of the LFC represent more temporally abstracted signals which bias processing in more caudal regions of the LFC (Koechlin and Summerfield, 2007; Koechlin et al., 2003). Badre and D'Esposito (2007) found similar activations in a series of experiments that varied whether cues indicated a particular action, a particular feature that specified an action, or a particular dimension that specified a feature that specified an action. They argued that the rostral-caudal axis of the LFC is organized by the abstractness of action representations such that rostral areas of the LFC represent more general action sets. These ideas were supported in a follow-up study that examined brain-damage patients. They found that impairments in action decisions depended upon the rostral-caudal location of LFC lesions such that rostral LFC lesions impaired performance on abstract, but not concrete tasks, while caudal LFC lesions impaired performance on concrete tasks, as well as abstract tasks (Badre et al., 2009). Together, these studies demonstrate that more caudal regions of the LFC are involved for action decisions that are closer in time and more concrete while more rostral regions are involved for action decisions that are more distant in time and more abstract (Fuster, 1997).

While extant fMRI data is largely supportive of abstraction proposals of LFC organization, there are a number of limitations that weaken the inferences that can be drawn. First, abstraction effects rarely selectively engage dissociable frontal regions. In most cases, abstraction effects that engage rostral areas of the LFC also engage caudal areas of the LFC. For example, while Koechlin and colleagues (2003) associated effects of episodic control with activations in mid-lateral prefrontal cortex (PFC), episodic control effects were also observed in the caudal inferior frontal sulcus/inferior frontal junction (IFJ) and dorsal premotor cortex (PMd). Similarly, Badre and D'Esposito (2007) associated effects of dimension competition with activations in mid-lateral PFC, but these effects were also present caudally in the IFJ. While these patterns might indicate that caudal regions

accumulate information from rostral areas (Koechlin et al., 2003), such patterns are also amenable to complexity/difficulty arguments. Second, prior designs have compared different levels of control/representation across experiments which differed heavily with regard to ancillary demands such as working memory (WM) load, vigilance, and task difficulty. Furthermore, many comparisons were done block-wise across experimental conditions, thereby averaging activations across a long time-scale. Such block-wise analyses have poor sensitivity to isolate component processes (Curtis and D'Esposito, 2003). Event-related analyses that separately isolate processes associated with different levels of representational abstraction/control would provide stronger evidence for rostral-caudal gradients of abstraction. Moreover, by separating out different sources of information (i.e. episodic cues, contextual cues, etc.) it may be possible to cleanly dissociate abstraction effects on a level-by-level basis. Hence, more sophisticated designs are needed to provide further insights into the organization of the LFC.

The present study was designed to investigate the organizational structure of the LFC. Here, subjects responded to stimuli based upon a nested series of cues held in WM (Figure 1). Under the '1' context, subjects made a target response to the letter 'X' if it was preceded by the letter 'A' and made a non-target response otherwise. Under the '2' context, subjects made a target response to the letter 'Y' if it was preceded by the letter 'B' and made a non-target response otherwise. Hence, subjects had to keep in WM both a higher level context ('1' or '2') and a lower level context ('A' or 'B') to determine how to respond to the letters 'X' and 'Y'. Cues were spaced in time affording the ability to isolate the information represented in WM at different task intervals. As a result, we could estimate the representations associated with different levels of abstract cues apart from effort associated with responding itself. Furthermore, we used multi-variate pattern analysis (MVPA) to train a machine learning algorithm the patterns associated with different combinations of cues. This method has been demonstrated to reveal distributed representations that may not necessarily be identified by uni-variate measures of fMRI signal (Norman et al., 2006; Pereira and Botvinick, 2011). A key aspect of this method is that it allows the comparison of conditions that are well-matched in ancillary demands such as attention and WM load since differences between conditions are assessed by dissociable *patterns* of activation rather than the gross magnitude of activation. Additionally, MVPA permits assessments of the degree to which different conditions are similar to one another, providing richer information to understand the frontal representations that underlie flexible cognitive control.

Here, we used MVPA to identify regions of the LFC that represent contextual information in WM. By examining different task intervals, we tracked the areas of the LFC that represented context as information about the rules governing the forthcoming response was accrued. According to rostral-caudal abstraction proposals of the LFC, rostral areas of the LFC should demonstrate representation of the higher level context when no other information was provided. However, as additional contextual information was presented, making action rules more concrete, more caudal areas of the LFC should be involved in contextual representation. Accordingly, we used MVPA to examine two intervals: 1) the delay period after the higher level context cue was presented, but before the lower level context cue was presented (higher level context, first delay); 2) the delay period after the lower level context cue was presented when information about both the higher and lower level context were known (higher + lower level context, second delay), but before a response was made. We anticipated that rostral areas of the LFC would represent the higher level context during the first delay since the higher level context denoted an abstract rule set. By contrast, we expected that more caudal areas of the LFC would represent the combination of the higher and lower level context during the second delay since these representations were more concrete. Such a pattern would provide further support for rostral-caudal abstraction proposals of the LFC.

During analysis, an additional consideration became evident: certain context combinations (i.e. '1B' and '2A') always led to a non-target response, regardless of the forthcoming stimulus. By contrast, other context combinations left the forthcoming response fully undetermined (i.e. '1A' and '2B'). We refer to context combinations that inform the forthcoming response as "response certain" while those that leave the forthcoming response undetermined "response uncertain". Given that response certain contexts provide complete and concrete information, these contexts would be expected to be represented in the caudal-most regions of the LFC (i.e. motor cortex). Response uncertain contexts would be expected to form an intermediate level of abstraction in-between the higher level context and response certain contexts. Hence, when examining the combination of higher and lower level contexts during the second delay, we further distinguished response certain and uncertain contexts to investigate more detailed levels of abstraction along the LFC.

Material and Methods

Participants

Twenty-one (11 female) right-handed native English speakers with normal or corrected-to-normal vision participated in the experiment (mean age 23.7 years; range 21–32). Informed consent was obtained in accordance with the Institutional Review Board at Indiana University. Subjects were compensated at a rate of $20/hr for participation plus a performance based bonus (mean $2.43; range $1.24–$3.76).

Procedure

The task is depicted in Figure 1. Subjects performed a variant of the AX-CPT (Barch et al., 2009; MacDonald, 2008; Servan-Schreiber et al., 1996) referred to as the 1-2-AX-CPT (Frank et al., 2001; O'Reilly and Frank, 2006). The 1-2-AX-CPT requires subjects to hold two levels of contexts in mind in order to make responses. These contexts are hierarchical, forming higher and lower level action rules (sometimes referred to as an outer and inner loop). In this task, subjects observed a series of visually presented digits and letters and made responses to the letters 'X' and 'Y'. Responses to these letters were based on a hierarchical digit-letter sequence. Under the '1' context, subjects made a target response to the letter 'X' if it was preceded by the letter 'A' and made a non-target response otherwise. Under the '2' context, subjects made a target response to the letter 'Y' if it was preceded by the letter 'B' and made a non-target response otherwise. Hence, subjects had to keep in WM both a higher level context ('1' or '2') and a lower level context ('A' or 'B') to determine how to respond to the letters 'X' and 'Y'. Additional stimuli were presented that subjects were told to ignore ('3','C', and 'Z') for the purposes of uni-variate analysis described elsewhere (Nee and Brown, in press).

Responses were made with the index finger of either hand with the target hand counter-balanced between subjects. All relevant digits ('1'/'2') and letters ('A'/'B', 'X'/'Y') appeared in equal proportions throughout the experiment. Each stimulus was presented for 1 second. Jittered 4–6 second intervals separated successive letter stimuli and each digit stimulus was preceded and followed by a 10 second interval. Subjects completed 6 runs of 18 trials each while being scanned. Within a week prior to scanning, subjects completed a full session outside the scanner in order to minimize learning effects during scanning. These data were not analyzed and were used merely for practice purposes and to limit potentially confounding effects of learning during scanning.

Imaging Acquisition and Preprocessing

Images were acquired on a 3T Siemens Trio. Stimuli were presented to the subject via a projector at the rear of the scanner, reflected off a mirror mounted to the headcoil.

Experimental tasks were presented using E-Prime software (Psychology Software Tools, Inc., Pittsburgh, PA).

Functional T2*-weighted images were acquired using an EPI sequence with 35 contiguous slices and 3.44 × 3.44 × 3.75 mm voxels (TR = 2000 ms; echo time = 25 ms; flip angle = 70; field of view = 220). Phase and magnitude images were collected to estimate the magnetic inhomogeneity. T1-weighted MPRAGE images were collected for spatial normalization (256 × 256 × 192 matrix of 1 × 1 × 1 mm^3 voxels; TR = 1800 ms; echo time = 2.67 ms; flip angle = 9).

Functional data were spike-corrected to reduce the impact of artifacts using AFNI's 3dDespike (http://afni.nimh.nih.gov/afni). Subsequent preprocessing was done using SPM5 (http://www.fil.ion.ucl.ac.uk/spm/). Functional images were corrected for differences in slice timing using sinc-interpolation and head movement using least-squares approach and a 6 parameter rigid body spatial transformation. These data were then analyzed using the general linear model implemented in SPM5 and then submitted to MVPA (described in more detail below).

Imaging Analysis

Multi-variate analysis was performed using a parameter estimate approach. First, separate parameter estimates were calculated for each event using the general linear model implemented in SPM5. The model included impulse regressors for each stimulus event as well as epoch regressors to capture delay periods associated with WM maintenance. These delay period regressors were used in subsequent analyses. Regressors were convolved with a canonical hemodynamic response function. This approach to using individual event parameter estimates is similar in logic to the beta series approach for functional connectivity (Rissman et al., 2004). The model also included a temporal high-pass filter (128 s) and correction for temporal autocorrelation using an autoregressive AR(1) model. Parameter estimates for events associated with error trials were modeled separately and were not included in MVPA analyses.

Six different categories of delay periods were distinguished on the basis of the cues subjects held in WM. Two categories corresponded to delay periods after the higher level context cue was presented, but before the lower level context cue was presented (first delay). We refer to these categories as the '1' context and the '2' context, respectively. The remaining four categories corresponded to delay periods following the lower level context cue, but before the response (second delay). These four categories corresponded to the four combinations of the higher level context ('1'/'2') and lower level context ('A'/'B'): i.e. '1A', '1B', '2A', '2B'. Notably, because of the use of irrelevant stimuli ('3','C', and 'Z'), these delay period categories did not necessarily correspond to the preceding stimulus. For example, the '1' context followed the stimuli '1' and '3' equally as often. Similarly, the '2' context followed the stimuli '2' and '3' equally as often.

All MVPA analyses were performed using a leave-1-run-out procedure. In this procedure, data from all but one run were used to train a machine learning classifier the patterns associated with a given category. Then, data from the remaining run was tested. This procedure was rotated so that all 6 runs were tested. Performance was assessed by the proportion of correctly classified test examples. For each MVPA analysis, equal numbers of examples of each category were used to train and test the machine learning classifier. When different numbers of examples of each category were present (due to more errors in one condition than another), the larger category was randomly sampled to equate the number of examples. This procedure is necessary since unequal numbers of category examples can bias classifier algorithms. For ROI analyses, random sampling was repeated 20 times and

performance was averaged across each sampling to ensure that results were not driven by a particular random sample.

Whole-brain MVPA was performed using a searchlight procedure with Searchmight software implementing a Gaussian Naïve Bayesian classifer (http://minerva.csbmb.princeton.edu/searchmight/; Pereira & Botvinick, 2011). For each voxel, a neighborhood of surrounding voxels was established consisting of all voxels that shared a vertex (i.e. 3 × 3 × 3 cube of 27 voxels per neighborhood). The pattern of activation across these voxels was used to train and test a machine learning classifier. Classification performance was associated with the center voxel and this process was repeated for all voxels in the brain. Two different whole-brain analyses were performed. The first looked for areas that distinguished the higher level context in isolation (i.e. '1' context vs '2' context) during the first delay. The second looked for areas that distinguished the combination of the higher level and lower level contexts (i.e. '1A', '1B', '2A', '2B') during the second delay. Whole-brain searchlight analyses were performed separately for each subject in their native space. The analyses resulted in voxel-wise accuracy maps reflecting the proportion of correct classifications. These accuracy maps were transformed into z-maps and normalized to the MNI template. Normalized maps were subsequently smoothed with an 8-mm FWHM Gaussian kernel and subjected to a second level random-effects group analysis. The group analysis revealed voxels that demonstrated classification significantly greater than chance. Group whole-brain maps were thresholded at $p < 0.005$ uncorrected at the voxel level with a 171 voxel extent providing correction for multiple comparisons ($p < 0.05$, cluster corrected) according to simulations using AlphaSim.

To assess regions showing selectively greater classification of the higher level context in isolation than the combination of higher and lower level contexts and vice versa, we directly contrasted the whole-brain searchlight maps of each classification. Contrast whole-brain maps were thresholded at $p < 0.005$ uncorrected at the voxel level and restricted to regions demonstrating significantly above-chance classification in the analyses above. A cluster extent of 66 voxels provided correction for multiple comparisons ($p < 0.05$, cluster corrected) according to simulations using AlphaSim.

To elucidate the representations that lead to significant classification in whole-brain analyses, follow-up region-of-interest (ROI) analyses were performed. 10 mm radius spheres were centered on frontal peaks uncovered by the whole-brain analyses. Since regions that demonstrated significant classification were predominantly left-lateralized, we chose to focus on left hemisphere regions including the lateral frontopolar cortex (FPC; area 10), mid-dorsolateral PFC (DLPFC; area 9/46), inferior frontal junction (IFJ; area 8), dorsal premotor cortex (PMd; area 6), and primary motor cortex (area 4). For the mid-DLPFC, whole-brain analysis revealed a significant right, but not left hemisphere peak. For consistency, a left hemisphere mid-DLPFC ROI was created by reversing the sign of the x-coordinate. Notably, results in the left and right mid-DLPFC were qualitatively similar and no effect of hemisphere was found. For each ROI, MVPA was performed using the Princeton MVPA Toolbox (http://code.google.com/p/princeton-mvpa-toolbox/). Classification was performed using L2 penalized logistic regression with an optimal penalty search (Rissman et al., 2010). Of particular interest in these analyses was classifier estimates of category evidence. For each test example, the classifier estimated the degree to which the test pattern matched the learned pattern of each category (i.e. category evidence). By comparing classifier evidence for the correct category versus incorrect categories, the regions' sensitivity to a particular category or class of categories can be determined. For example, suppose that when testing examples of the category '1A' the classifier found strong evidence for both categories '1A' and '1B', but not '2A' and '2B'. That '1A' and '1B' are highly confusable suggests that what is represented is the higher level context ('1')

irrespective of the lower level context. Hence, examination of the amounts of evidence for each category can further elucidate the type of representation that lead to above-chance classification. This approach is similar in logic to representational similarity analysis (Connolly et al., 2012; Kriegeskorte et al., 2008), but uses classifier evidence as a similarity metric rather than Pearson correlations.

Results

Behavioral Results

Full details of the behavioral results are reported elsewhere (Nee and Brown, in press). Two aspects of the behavioral data are relevant for present purposes. First, subjects performed the task highly accurately (mean error-rate < 5%) indicating that they understood the instructions and appropriately maintained contextual information in WM. Second, the data were symmetrical with respect to the higher level context (effect of higher level context $F(1,20) < 1$ for both error-rate and reaction time; see Supplemental Figure 1). In other words, performance in the '1A' condition was identical to performance in the '2B' condition and performance in the '1B' condition was identical to performance in the '2A' condition. This symmetry indicates appropriate use of contextual cues to guide performance.

Whole-brain MVPA Results

Higher Level Context (First Delay)—We began by looking for regions that represented the higher level context. To do so, a whole-brain searchlight analysis was performed that searched for regions that correctly distinguished context '1' from context '2' in WM during the first delay. Significantly above-chance classification was found in bilateral rostral PFC extending from the ventral frontal pole into the lateral orbitofrontal cortex (OFC; area 10/11; left hemisphere: −30 54 −18; right hemisphere peak: 32 66 0). For clarity, we will refer to this region as the lateral OFC to distinguish it from more dorsolateral frontopolar regions discussed below (i.e. lateral FPC). Significant classification also extended medially into the bilateral ventromedial PFC (VMPFC; area 11; left hemisphere peak: −14 60 −8; right hemisphere peak: 10 40 −22). No other regions demonstrated significant classification. Hence, these data demonstrate that higher level contexts are represented in the rostral-ventral PFC (Table 1; Figure 2).

Higher and Lower Level Context Combination (Second Delay)—Next, we looked for context representation during the second delay when both higher and lower level context information was maintained in WM. In this analysis, a whole-brain searchlight was performed to find voxels that correctly distinguished the combination of contexts '1A', '1B', '2A', and '2B' in WM during the second delay. Notably, this 4-way classification could be driven by the classification of the higher level context (i.e. '1A' and '1B' are confused but distinguished from '2A' and '2B'), the lower level context (i.e. '1A' and '2A' are confused but distinguished from '1B' and '2B'), or any combination of the two. Hence, this analysis was designed to identify regions that represented any combination of the higher and lower level context. We also report 2-way classifications that collapse across higher and lower level context combinations in the Supplemental Material (see Supplemental Figure 5).

In contrast to the focal areas that represented the higher level context, a broad network of regions correctly distinguished the combination of higher and lower level contexts (Table 1; Figure 2). Classification was strongest in the bilateral sensorimotor cortex (areas 4 and 3, left hemisphere peak: −40 −30 50; right hemisphere peak 40 −24 56). Peaks were also found in numerous regions previously associated with hierarchical control including the PMd (area 6; −30 −12 60), IFJ (area 8; −36 2 42), mid-DLPFC (area 9/46; 40 28 34) and lateral FPC (area 10; −36 54 8). Significant classification was also found in the mid-cingulate and

supplementary motor area, as well as in the precuneus, intra-parietal sulcus, and temporal-parietal junction. We unpack the representations that drove significant classification below (Results: Classifier Evidence within Regions-of-Interest). Before doing so, we explicitly compare representation of the higher level context in the first delay and representation of the combination of higher and lower level contexts in the second delay.

Contrasting Levels of Context Representation—Interestingly, there was no overlap in the regions that represented the higher level context in isolation during the first delay and regions that represented the combination of higher and lower level context during the second delay. To examine whether this was a thresholding issue, we further interrogated classifier performance within the left and right lateral OFC during the second delay when subjects held the combination of higher and lower level contexts in WM. ROIs within the left (mean accuracy 24.9%, n.s.) and right (mean accuracy 25.7%, $t(20) = 0.95$, $p > 0.35$) OFC did not demonstrate significantly greater than chance (25%) classification of the four combinations of the higher and lower level context. A second classification analysis was then performed which collapsed the combinations ‘1A’ and ‘1B’ (context ‘1’) and the combinations ‘2A’ and ‘2B’ (context ‘2’). This afforded the examination of whether the higher level context continued to be represented in the lateral OFC during the second delay independently from the lower level context. The results did not demonstrate significantly above chance (50%) classification in the left (mean accuracy 49.1%, n.s.) and right (mean accuracy 50.1%, $t(20) = 0.07$, $p > 0.9$) lateral OFC. Finally, to examine whether any region represented the higher level context during the second delay, we performed a separate whole-brain classification analysis again collapsing combinations ‘1A’ and ‘1B’ and contrasting them with combinations ‘2A’ and ‘2B’. This analysis did not reveal any LFC regions, but there were clusters in the left posterior insula (−40 −20 6, 171 voxels) and preCuneus (0 −64 46, 244 voxels; see Supplemental Figure 5). Hence, the LFC did not appear to sustain representation of the higher level context in isolation after the lower level context was presented. These results suggest that when more concrete information is presented, more abstract information is discarded.

To further assess whether distinct regions represent the higher level context in isolation and the combination of the higher and lower level contexts, we directly contrasted the whole-brain searchlight maps for each classification. The results demonstrated that the left lateral OFC (peak: −30 54 −14) demonstrated significantly stronger classification of the higher level context in isolation during the first delay than the combination of higher and lower level contexts during the second delay. A reverse pattern was found in bilateral sensorimotor cortex (left peak: −44 −26 50; right peak: 36 −24 56), left PMd (−30 −10 68), right mid-DLPFC (40 24 30), left lateral FPC (−34 52 12), precuneus (2 −56 50), and right intra-parietal sulcus (28 −64 52). These results demonstrate that distinct regions represent varying forms of abstracted contexts in WM at different points in time (Table 2; Figure 3).

Classifier Evidence within Regions of Interest

The whole-brain searchlight results revealed several lateral frontal regions that distinguished the combination of higher and lower level contexts during the second delay. These regions corresponded quite closely to previous investigations of rostral-caudal gradients of control/representation (Badre and D’Esposito, 2007; Koechlin et al., 2003). To further understand the representations/processes that led to significant classification, we examined patterns of classifier evidence within ROIs centered around peaks uncovered in the whole-brain analyses. Specifically, we examined the left lateral FPC (area 10; −36 54 8), mid-DLPFC (area 9/46; −40 28 34), IFJ (area 8, −36 2 42), PMd (area 6; −30 −12 60), and primary motor cortex (area 4; −30 −30 62). For each combination of higher and lower level contexts (e.g. ‘1A’), we examined the amount of evidence the classifier found for each context

combination (i.e. ‘1A’, ‘1B’, ‘2A’, and ‘2B’). We then tested whether there was significantly more evidence for the correct context combination compared to each incorrect context combination. This approach is similar to representational similarity analysis (Connolly et al., 2012; Kriegeskorte et al., 2008) and seeks to identify representational structure through patterns of neural similarity and dissimilarity.

As alluded to earlier, the context combinations ‘1B’ and ‘2A’ always led to a non-target response (response certain), regardless of the forthcoming stimulus. Hence, subjects could prepare a response in advance. Self-report data indicated that 18 of the 21 subjects used a proactive strategy in which they used the observed cues to prepare for forthcoming stimuli. This contrasts with reactive strategies where cues are passively stored and response sets are not prepared in advance (Braver, 2012). Consistent with the self-report data, reaction times were faster when the response could be prepared in advance relative to when the forthcoming response was uncertain (i.e. relative to ‘1A’ and ‘2B’ combinations; mean difference 47.1 ms, $t(20) = 4.92$, $p < 0.0001$). As a result, it is likely that the combination of ‘1B’/’2A’ cues led to the representation of the forthcoming response, which may explain the strong involvement of motor regions in classifying the combination of higher and lower level context cues. Hence, in the forthcoming analyses, we distinguish between “response certain” (‘1B’/’2A’) and “response uncertain” (‘1A’/’2B’) context combinations to investigate potential proactive representations of motor responses in addition to representations of context more generally.

Full results are depicted in Supplementary Figure 2. As can be seen, the results were highly symmetrical indicating that classifier patterns for ‘1A’ and ‘2B’ were mirror inverses, as were classifier patterns for ‘1B’ and ‘2A’. This is the same symmetrical pattern demonstrated by the behavioral data (Supplemental Figure 1). Hence, for simplicity, we collapsed the data by folding ‘1A’ and ‘2B’ together and ‘1B’ and ‘2A’ together. Context combination were categorized as “Target” (evidence for correct category), “SameHi” (evidence for the incorrect category that shared the same higher level context), “SameLo” (evidence for the incorrect category that shared the same lower level context), and “Inverse” (evidence for the mirror-reverse category). For example, if the correct context combination was ‘1A’, ‘1A’ would be the “Target”, ‘1B’ would be the “SameHi”, ‘2A’ would be the “SameLo”, and ‘2B’ would be the “Inverse”.

Data are summarized in Table 3 and Figure 4.

Lateral Frontopolar Cortex—Starting rostrally, the data indicated that in the lateral FPC different combinations of contexts were highly confusable. The only significant distinction was between ‘1A’/’2B’ and the “SameLo” ($t(20) = 2.90$, $p < 0.01$). This pattern became clearer through examination of the confusion matrix (Supplemental Figure 3). When normalizing by the classifier’s overall rate of guessing a particular category, the confusion matrix indicated that the classifier guessed the correct category most frequently for all context combinations. For 3 of the 4 context combinations, the “SameLo” was guessed least frequently. These data demonstrate that the lateral FPC distinguishes cases where the same lower level context takes on a different meaning due to different higher level contexts.

Mid-Dorsolateral PFC and Inferior Frontal Junction—Distinctions were stronger in the mid-DLPFC. Here, when shown examples of ‘1A’/’2B’, the classifier found more evidence for the “Target” category than all other categories (all $t(20) > 2.55$, $p < 0.05$). These data indicate representation of the contexts ‘1A’ and ‘2B’ in the mid-DLPFC. By contrast, the classifier did not distinguish ‘1B’/’2A’ from incorrect categories. As described above, context combinations ‘1B’/’2A’ always led to a non-target response (response certain). So, these context combinations provided complete and concrete response

information. That the classifier correctly distinguished ‘1A’/’2B’ (response uncertain) from incorrect categories, but not ‘1B’/’2A’ (response certain) suggests that the mid-DLPFC represents the combination of higher and lower level contexts only when the response is yet to be determined. A nearly identical pattern was found in the IFJ.

PreMotor and Motor Cortex—A very different set of results was found caudally in the PMd and primary motor cortex. In these areas, ‘1B’/’2A’ (response certain) could be distinguished from all incorrect categories. Hence, when the combination of contexts determined the forthcoming response with certainty, motor areas represented the combination of contexts. Notably, this representation cannot be reduced to the forthcoming response itself irrespective of context. If this were the case, ‘1B’ and ‘2A’ would be confused with each other (i.e. the correct category would be confused with the “Inverse”). This is because both context combinations lead to an identical response (i.e. non-target). The significant difference between ‘1B’/’2A’ and the “Inverse” indicates the representation of context over-and-above the forthcoming response itself.

A slightly different pattern was found for context combinations that did not predetermine the forthcoming response. While ‘1A/’2B’ (response uncertain) could be distinguished from both the “SameHi” and “SameLo”, there was equivalent evidence for the “Inverse” category. This pattern suggests a binary distinction: cases where the response could be not prepared in advance (‘1A’/’2B’) were distinct from cases where the response was already known (‘1B’/’2A’). However, cases where the response could not be prepared in advance could not be distinguished from each other (i.e. contexts ‘1A’ and ‘2B’ were confused with each other).

The data above suggest dissociable representational structures in caudal (PMd, motor cortex) and mid-lateral (mid-DLPFC, IFJ) frontal regions. Mid-lateral regions appeared to represent contexts in which the response was yet to be determined. However, when complete response information was available, contextual representation in mid-lateral regions was reduced and contextual represented in caudal motor regions was increased. To formally contrast the representational patterns in mid-lateral and caudal areas of the frontal lobes, we calculated a selectivity index separately for cue combinations that indicated the forthcoming response with certainty (‘1B’ and ‘2A’: “response certain”) and those that did not (‘1A’ and ‘2B’: “response uncertain”). This index was calculated by summing the difference between the classifier evidence for the correct category and each incorrect category. These data were submitted to a 2 × 4 ANOVA with factors of Response Certainty (response certain, response uncertain) and Region (mid-DLPFC, IFJ, PMd, motor cortex). This analysis revealed a significant Response Certainty × Region interaction ($F(3,60) = 3.15$, $p < 0.05$) with no main effects (ME Response Certainty: $F(1,20) = 0.04$, $p > 0.8$; ME Region: $F(3,60) = 2.18$, $p = 0.1$). As depicted in Figure 5, this interaction was driven by enhanced selectivity in motor regions (PMd, motor cortex) when the response was certain relative to uncertain (averaged across PMd and motor cortex: $t(20) = 2.46$, $p < 0.05$), with a trend in the opposite direction in mid-lateral regions (averaged across mid-DLPFC and IFJ: $t(20) = -1.99$, $p = 0.06$). Collectively, these data indicate that as information becomes more concrete, representation of context is reduced in mid-lateral regions, but increased in caudal motor areas.

Discussion

The present study sought to examine the representational structure of different rostral-caudal regions along the LFC. The results demonstrated distinct representational patterns between the lateral OFC, lateral FPC, mid-DLPFC and IFJ, and motor regions including PMd and primary motor cortex. The lateral OFC was exclusively sensitive to higher level contexts that denoted abstracted sets without information about particular future responses. This

sensitivity in the lateral OFC was abolished following lower level contextual cues that provided more concrete stimulus-response contingencies. In cases where the combination of higher and lower level context left the forthcoming response undetermined, mid-lateral regions in the mid-DLPFC and IFJ demonstrated representation of context combinations. However, when contexts already dictated the forthcoming response with certainty, only motor regions in the PMd and primary motor cortex demonstrated representation of contextual information. Collectively, these results demonstrate gradients of representation related to abstractions of action sets along the rostral-caudal axis of the LFC. As more concrete information regarding the necessary action was available, the representation of context was reduced in rostral areas of the LFC, but increased in more caudal areas of the LFC.

These results complement previous analyses on this same dataset that focused on processes involved in updating WM with different levels of contextual information (Nee and Brown, in press). There, we demonstrated that rostral areas of the LFC (area 46) were involved when updating WM with higher level contextual cues while caudal areas of the LFC (area 6) were involved when updating WM with lower level contextual cues. Together, both the uni-variate and multi-variate results support rostral-caudal abstraction proposals of the LFC (see also Supplemental Results: Relationship Between Uni-variate and Multi-variate Analyses).

Orbitofrontal and Frontal Polar Cortex

Our searchlight analyses demonstrated two rostral regions that demonstrated different forms of representation at different points in the task. The first, which we have referred to as the lateral OFC, extended from the ventral frontal pole (area 10) through the lateral orbital surface (area 11). This region demonstrated significant classification of the higher level context in isolation, which was abolished following the lower level context cue. The second, which we have referred to as the lateral FPC, was located dorsal to the lateral OFC cluster and was restricted to the lateral surface. This region demonstrated significant classification of the combination of higher and lower level contexts. These regions were in close proximity and previous research has demonstrated activations that extend across these regions during processes of sub-goaling (Braver and Bongiolatti, 2002; Charron and Koechlin, 2010) and reward-related uncertainty (Badre et al., 2012). In the present data, despite their close spatial proximity, the OFC and lateral FPC demonstrated distinct representational roles.

A considerable amount of evidence in humans, monkeys, and other species has demonstrated that the OFC is critical for flexible behavior (see Rolls, 2004; Schoenbaum, Roesch, Stalnaker, & Takahashi, 2009 for reviews). Lesions to the OFC produce impairments in the ability to learn changing stimulus-reward contingencies in cases such as reversal learning tasks. In the present study, the lateral OFC was selectively involved in representing the higher level context in isolation. However, when lower level context cues were presented, these representations could no longer be distinguished. This pattern suggests that the OFC is important to establish a contextual set, but not necessarily to maintain it across successive stimuli. The establishment of a contextual set and associated reward contingencies is likely to be critical to reversal learning and other forms of flexible behavior.

Dorsal to the OFC, the lateral FPC is thought to be involved in branching: the process of completing one task while holding another in a pending state (Koechlin and Hyafil, 2007). Compared to situations when subjects abandon one task in favor of another, the lateral FPC is activated when subjects switch to a new task, but return to the original task at a later time (Charron and Koechlin, 2010; Koechlin et al., 1999). Other data have demonstrated that the lateral FPC is sensitive to the reward value of unchosen options (Boorman et al., 2009, 2011). Together, these data indicate that the lateral FPC is important in maintaining and

evaluating alternative choices or task-sets (Rushworth et al., 2011). In the present data, the whole-brain searchlight identified the lateral FPC as a region that demonstrated above-chance classification of the combination of higher and lower order contexts. More detailed analysis suggested that the lateral FPC distinguished a very specific case of instances: cases where the lower level context was the same (e.g. 'A'), but the higher level context differed (e.g. '1' vs '2'). This pattern is seen most clearly in the confusion matrix (Supplemental Figure 3). As depicted in the figure, the pattern classifier correctly guessed the correct category more frequently than any incorrect category. In 3 of the 4 cases, the category matching the correct category in lower level context was guessed least frequently. These data indicate a distinction between the currently relevant higher level context and the irrelevant higher level context when processing the same lower level context. This distinction may be related to the maintenance and evaluation of alternative choices/task-sets.

Mid-Dorsolateral Prefrontal Cortex and Inferior Frontal Junction

Our whole-brain searchlight analysis uncovered a region in the mid-DLPFC that was sensitive to the combination of higher and lower level context cues. In non-human primates, single neurons in the mid-lateral PFC are sensitive to abstract action rules and strategies (Genovesio et al., 2005; Wallis and Miller, 2003; Wallis et al., 2001). Human neuroimaging research has demonstrated the importance of the mid-lateral PFC when rules are switched (Bunge, 2004) or when a less prepotent rule has to be selected over a more natural rule mapping (MacDonald et al., 2000). Together, there is strong evidence that the mid-lateral PFC represents abstract rules (Bunge and Zelazo, 2006; Bunge, 2004; Bunge et al., 2005; Crone et al., 2006). In the present data, the mid-lateral PFC discriminated the combination of higher and lower level contexts only when these combinations did not inform the forthcoming response. That is, in these cases what was likely maintained in WM was a rule (i.e. 'X'-target, 'Y'-non-target). By contrast, when the forthcoming response was already known based on the higher and lower level contexts, the mid-lateral PFC did not show discrimination of any category. In these cases, it is likely that non-rule information was held in mind such as the response representation (i.e. non-target). Hence, these data are consistent with the notion that the mid-lateral PFC represents rules rather than actions themselves, whereas prospective actions are coded posteriorly in premotor and motor regions (Wallis and Miller, 2003).

Representational patterns in the IFJ were nearly identical to the mid-lateral PFC. Like the mid-lateral PFC, the IFJ has also been implicated in rule representation and early neuroimaging studies may have inappropriately attributed activations in the IFJ to the mid-lateral PFC (Brass et al., 2005; Derrfuss et al., 2005). While the mid-lateral PFC did not show any distinctions when the forthcoming response was known, the IFJ did demonstrate a significant difference between context combinations that indicated the upcoming response with certainty and the inverse combination of contexts. Hence, compared to the mid-lateral PFC, the IFJ did appear to differentiate some aspects of concrete contexts. This is consistent with the idea that the IFJ is an intermediate region between the mid-lateral PFC and motor regions.

Dorsal PreMotor and Motor Cortex

The present data indicated that activity in the PMd and primary motor cortex demonstrated highly significant classification of the combination of higher and lower level contexts. Classification was particularly accurate when the forthcoming response could be determined by the combination of contexts stored in WM. Notably, that cue combinations could be distinguished from their inverses demonstrates that these areas coded contextual information and not just the forthcoming response itself. That is, if the PMd and primary motor cortex simply represented "non-target" or "left" response, the context combinations '1B' and '2A'

would be indistinguishable since both require identical motor activity. That such context combinations could, in fact, be discriminated indicates the maintenance of context in addition to the forthcoming motor action. These data are compatible with demonstrations that motoric regions encode both concrete actions and more abstract rules (Wallis and Miller, 2003).

When the response could not be prepared in advance (i.e. '1A' and '2B'), both the PMd and primary motor cortex confused representations of a given context combination with its inverse. In other words, '1A' was confused with '2B' and vice versa. This pattern is in contrast to activations in the mid-DLPFC and IFJ that could distinguish between such cases. Together, these data indicate that the portion of cortex that represents context combinations varies depending upon uncertainty. When responses are uncertain, mid-level regions of the PFC appear to represent contextual information. However, when full response information is available, context is represented in motor regions. Hence, the particular frontal regions involved in representing context in WM vary as a function of information certainty (Koechlin and Summerfield, 2007).

Summary

MVPA during distinct phases of hierarchical WM task revealed a gradient of context representation along the rostral-caudal axis of the LFC. The rostral-most activations in the lateral OFC were involved during the establishment of a higher level contextual set. These activations did not persist after successive stimuli. Instead, other areas of frontal cortex were important in integrating aspects of higher and lower level contexts to inform action decisions. The lateral FPC represented a distinct relationship between the lower and higher level contexts. The mid-DLPFC and IFJ represented action rules when the combination of higher and lower level contexts left the forthcoming response uncertain. When the response could be anticipated, the PMd and primary motor cortex represented the forthcoming motor response, while also maintaining contextual information. Together, the rostral-most areas of the LFC were involved in representations that least resembled concrete motor acts, while the caudal areas of the LFC demonstrated the opposite pattern. Collectively, these data are consistent with rostral-caudal abstraction gradient proposals (Badre and D'Esposito, 2009; Badre, 2008; Koechlin and Summerfield, 2007; Koechlin et al., 2003). Contextual information that is abstract and distal from actual motor acts is represented in rostral areas of the LFC, while information that is concrete and proximate to actual motor acts is represented in caudal areas of the LFC.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported in part by AFOSR FA9550-07-1-0454 (JB), R03 DA023462 (JB), R01 DA026457 (JB), and the Indiana METACyt Initiative of Indiana University, funded in part through a major grant from the Lilly Endowment, Inc. Supported in part by the Intelligence Advanced Research Projects Activity (IARPA) via Department of the Interior (DOI) contract number D10PC20023. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DOI or the U.S. Government. The authors thank C. Chung, B. Pruce, and R. Fukunaga for help with data collection and F. Pereira for technical help with Searchmight analysis software.

References

Badre D. Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. Trends Cogn Sci. 2008; 12:193–200. [PubMed: 18403252]

Badre D, Doll BB, Long NM, Frank MJ. Rostrolateral prefrontal cortex and individual differences in uncertainty-driven exploration. Neuron. 2012; 73:595–607. [PubMed: 22325209]

Badre D, D'Esposito M. Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. J Cogn Neurosci. 2007; 19:2082–99. [PubMed: 17892391]

Badre D, D'Esposito M. Is the rostro-caudal axis of the frontal lobe hierarchical? Nat Rev Neurosci. 2009; 10:659–69. [PubMed: 19672274]

Badre D, Hoffman J, Cooney JW, D'Esposito M. Hierarchical cognitive control deficits following damage to the human frontal lobe. Nat Neurosci. 2009; 12:515–22. [PubMed: 19252496]

Barch DM, Berman MG, Engle R, Jones JH, Jonides J, Macdonald A, Nee DE, Redick TS, Sponheim SR. CNTRICS final task selection: working memory. Schizophr Bull. 2009; 35:136–52. [PubMed: 18990711]

Boorman ED, Behrens TE, Rushworth MF. Counterfactual choice and learning in a neural network centered on human lateral frontopolar cortex. PLos Biol. 2011; 9:e1001093. [PubMed: 21738446]

Boorman ED, Behrens TEJ, Woolrich MW, Rushworth MFS. How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. Neuron. 2009; 62:733–43. [PubMed: 19524531]

Brass M, Derrfuss J, Forstmann B, von Cramon DY. The role of the inferior frontal junction area in cognitive control. Trends Cogn Sci. 2005; 9:314–6. [PubMed: 15927520]

Braver TS. The variable nature of cognitive control: a dual mechanisms framework. Trends Cogn Sci. 2012; 16:106–13. [PubMed: 22245618]

Braver TS, Bongiolatti SR. The role of frontopolar cortex in subgoal processing during working memory. NeuroImage. 2002; 15:523–36. [PubMed: 11848695]

Bunge SA. How we use rules to select actions: a review of evidence from cognitive neuroscience. Cogn Affect Behav Neurosci. 2004; 4:564–79. [PubMed: 15849898]

Bunge SA, Wallis JD, Parker A, Brass M, Crone EA, Hoshi E, Sakai K. Neural circuitry underlying rule use in humans and nonhuman primates. J Neurosci. 2005; 25:10347–50. [PubMed: 16280570]

Bunge SA, Zelazo PD. Brain-Based Account of the Development of Rule Use in Childhood. Curr Dir Psychol Sci. 2006; 15:118–121.

Cavada C, Goldman-Rakic PS. Posterior parietal cortex in rhesus monkey: II. Evidence for segregated corticocortical networks linking sensory and limbic areas with the frontal lobe. J Comp Neurol. 1989; 287:422–445. [PubMed: 2477406]

Charron S, Koechlin E. Divided representation of concurrent goals in the human frontal lobes. Science. 2010; 328:360–363. [PubMed: 20395509]

Connolly AC, Guntupalli JS, Gors J, Hanke M, Halchenko YO, Wu YC, Abdi H, Haxby JV. The representation of biological classes in the human brain. J Neurosci. 2012; 32:2608–18. [PubMed: 22357845]

Crone EA, Wendelken C, Donohue SE, Bunge SA. Neural evidence for dissociable components of task-switching. Cereb Cortex. 2006; 16:475–86. [PubMed: 16000652]

Curtis CE, D'Esposito M. Persistent activity in the prefrontal cortex during working memory. Trends Cogn Sci. 2003; 7:415–423. [PubMed: 12963473]

Derrfuss J, Brass M, Neumann J, von Cramon DY. Involvement of the inferior frontal junction in cognitive control: meta-analyses of switching and Stroop studies. Hum Brain Mapp. 2005; 25:22–34. [PubMed: 15846824]

Frank MJ, Loughry B, O'Reilly RC. Interactions between frontal cortex and basal ganglia in working memory: a computational model. Cogn Affect Behav Neurosci. 2001; 1:137–60. [PubMed: 12467110]

Fuster, J. The prefrontal cortex: Anatomy, physiology, and neuropsychology of the frontal lobe. 3. Lippincott-Raven; Philadelphia: 1997.

Genovesio A, Brasted PJ, Mitz AR, Wise SP. Prefrontal cortex activity related to abstract response strategies. Neuron. 2005; 47:307–20. [PubMed: 16039571]

Koechlin E, Basso G, Pietrini P, Panzer S, Grafman J. The role of the anterior prefrontal cortex in human cognition. Nature. 1999; 399:148–51. [PubMed: 10335843]

Koechlin E, Hyafil A. Anterior prefrontal function and the limits of human decision-making. Science. 2007; 318:594–8. [PubMed: 17962551]

Koechlin E, Ody C, Kouneiher F. The architecture of cognitive control in the human prefrontal cortex. Science. 2003; 302:1181–5. [PubMed: 14615530]

Koechlin E, Summerfield C. An information theoretical approach to prefrontal executive function. Trends Cogn Sci. 2007; 11:229–35. [PubMed: 17475536]

Kriegeskorte N, Mur M, Ruff DA, Kiani R, Bodurka J, Esteky H, Tanaka K, Bandettini PA. Matching categorical object representations in inferior temporal cortex of man and monkey. Neuron. 2008; 60:1126–1141. [PubMed: 19109916]

MacDonald AW. Building a Clinically Relevant Cognitive Task: Case Study of the AX Paradigm. Schizophr Bull. 2008; 34:619–628. [PubMed: 18487225]

MacDonald AW, Cohen JD, Stenger VA, Carter CS. Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. Science. 2000; 288:1835–1838. [PubMed: 10846167]

Miller EK, Cohen JD. An integrative theory of prefrontal cortex function. Annu Rev Neurosci. 2001; 24:167–202. [PubMed: 11283309]

Nee DE, Brown JW. Dissociable frontal-striatal and frontal-parietal networks involved in updating hierarchical contexts in working memory. Cereb Cortex. in press.

Norman, Ka; Polyn, SM.; Detre, GJ.; Haxby, JV. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. Trends Cogn Sci. 2006; 10:424–30. [PubMed: 16899397]

O'Reilly RC. The What and How of prefrontal cortical organization. Trends Neurosci. 2010; 33:355–61. [PubMed: 20573407]

O'Reilly RC, Frank MJ. Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. Neural Comput. 2006; 18:283–328. [PubMed: 16378516]

Pereira F, Botvinick M. Information mapping with pattern classifiers: a comparative study. NeuroImage. 2011; 56:476–96. [PubMed: 20488249]

Petrides M, Pandya DN. Projections to the frontal cortex from the posterior parietal region in the rhesus monkey. J Comp Neurol. 1984; 228:105–16. [PubMed: 6480903]

Petrides M, Pandya DN. Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. Eur J Neurosci. 1999; 11:1011–36. [PubMed: 10103094]

Petrides M, Pandya DN. Efferent association pathways from the rostral prefrontal cortex in the macaque monkey. J Neurosci. 2007; 27:11573–86. [PubMed: 17959800]

Rissman J, Gazzaley A, Esposito MD. Measuring functional connectivity during distinct stages of a cognitive task. NeuroImage. 2004:1–12.

Rissman J, Greely HT, Wagner AD. Detecting individual memories through the neural decoding of memory states and past experience. Proc Natl Acad Sci USA. 2010; 107:9849–54. [PubMed: 20457911]

Rolls ET. The functions of the orbitofrontal cortex. Brain Cogn. 2004; 55:11–29. [PubMed: 15134840]

Rushworth MFS, Noonan MP, Boorman ED, Walton ME, Behrens TE. Frontal cortex and reward-guided learning and decision-making. Neuron. 2011; 70:1054–69. [PubMed: 21689594]

Schoenbaum G, Roesch MR, Stalnaker Ta, Takahashi YK. A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. Nat Rev Neurosci. 2009; 10:885–92. [PubMed: 19904278]

Servan-Schreiber D, Cohen JD, Steingard S. Schizophrenic deficits in the processing of context. A test of a theoretical model. Arch Gen Psychiatry. 1996; 53:1105–1112. [PubMed: 8956676]

Wallis JD, Anderson KC, Miller EK. Single neurons in prefrontal cortex encode abstract rules. Nature. 2001; 411:953–6. [PubMed: 11418860]

Wallis JD, Miller EK. From rule to response: neuronal processes in the premotor and prefrontal cortex. J Neurophys. 2003; 90:1790–806.

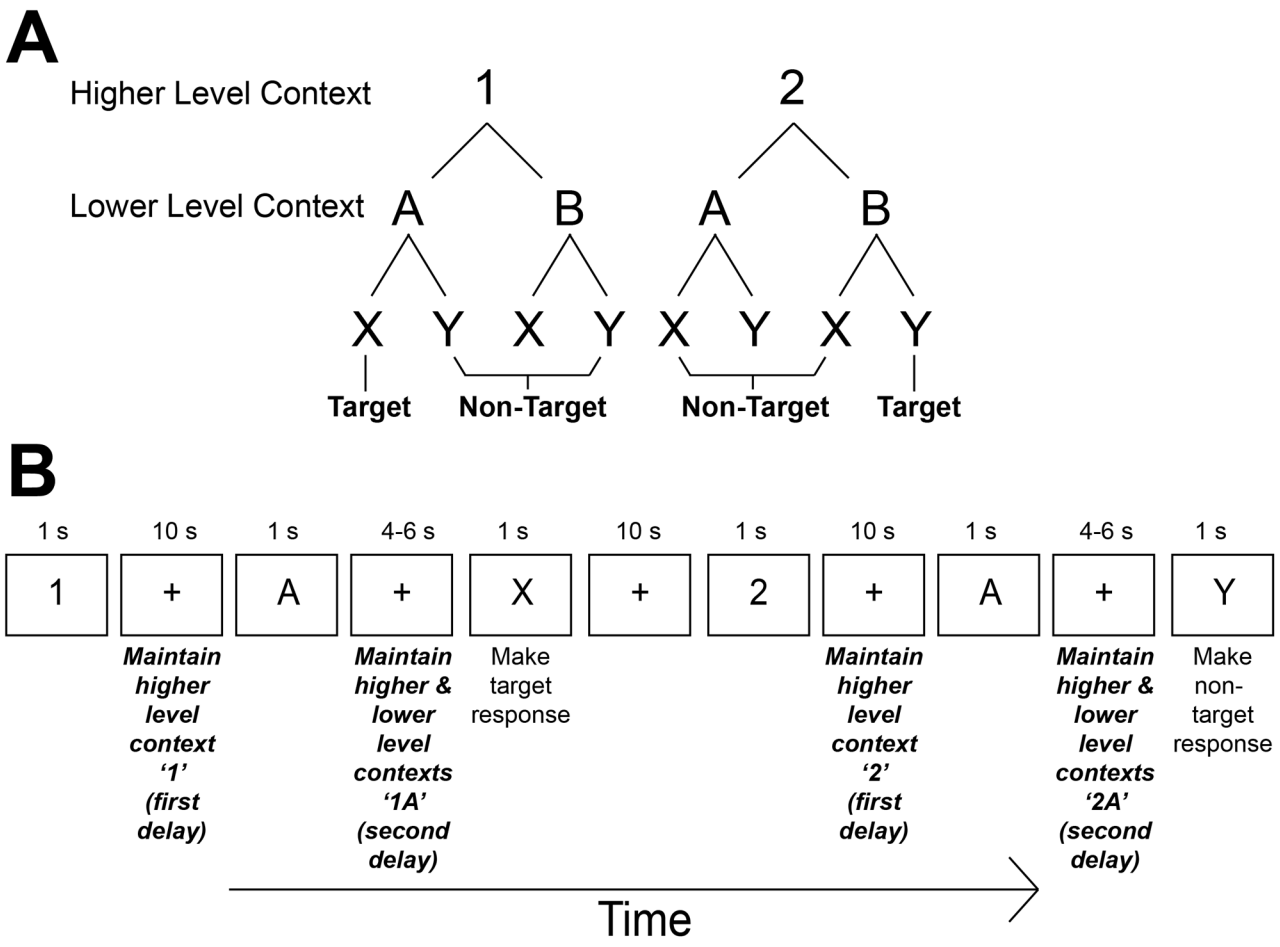

Figure 1.

Depiction of the task. Subjects responded to 'X' and 'Y' stimuli based upon a nested series of cues. A) Nesting rule structure. Under the '1' higher level context, subjects made a target response to the letter 'X' if within the 'A' lower level context. They made a non-target response otherwise. Under the '2' higher level context, subjects made a target response to the letter 'Y' if within the 'B' lower level context. They made a non-target response otherwise. B) Ordering of stimuli and periods of interest. Data were drawn from delay periods in-between stimuli in order to assess contextual representations within working memory (bold and italic). Following a higher level context cue, subjects maintained the higher level context in working memory (first delay). Following the lower level context cue, subjects maintained both the higher and lower level context in working memory (second delay). fMRI data from these maintenance periods were analyzed using multi-variate pattern analysis (MVPA). The design also included irrelevant stimuli (i.e. '3', 'C', 'Z') that subjects were told to ignore (not depicted). Periods following these irrelevant stimuli were also included in MVPA and these periods were categorized with respect to the information held in working memory. For instance, if a '1' cue was followed by a 'C' cue, both the interval following the '1' cue and the interval following the 'C' cue were included as examples of higher level context '1'. 10 second intervals followed number stimuli while 4–6 second intervals followed letter stimuli (hence, following a 'C' cue, the higher level context was estimated in a 4–6 second interval rather than a 10 second interval).

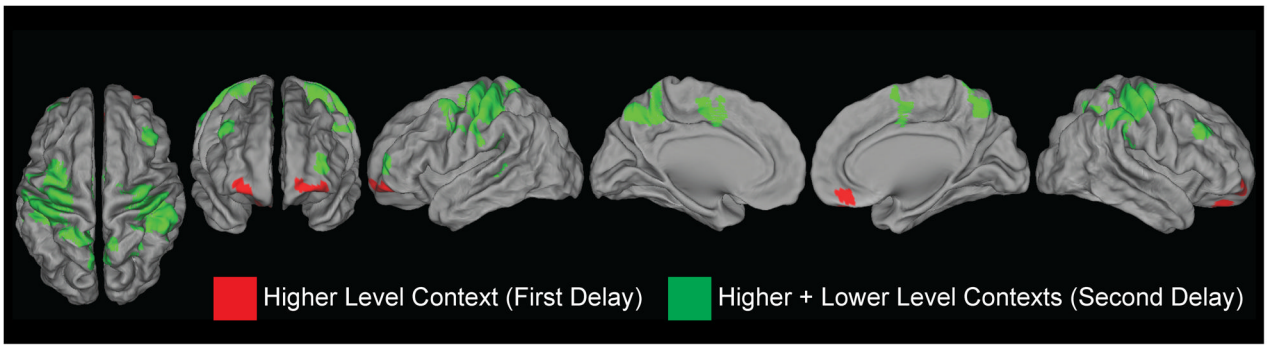

Figure 2.
Whole-brain searchlight results. Red: voxels demonstrating significant classification of the higher level context (i.e. discrimination of the '1' context from the '2' context). Green: voxels demonstrating significant classification of the combination of higher and lower level contexts (i.e. discrimination of the context combinations '1A', '1B', '2A', and '2B'). Results are thresholded at $p < 0.005$ at the voxel-level, with a 171 cluster extent providing a correct p-value of 0.05.

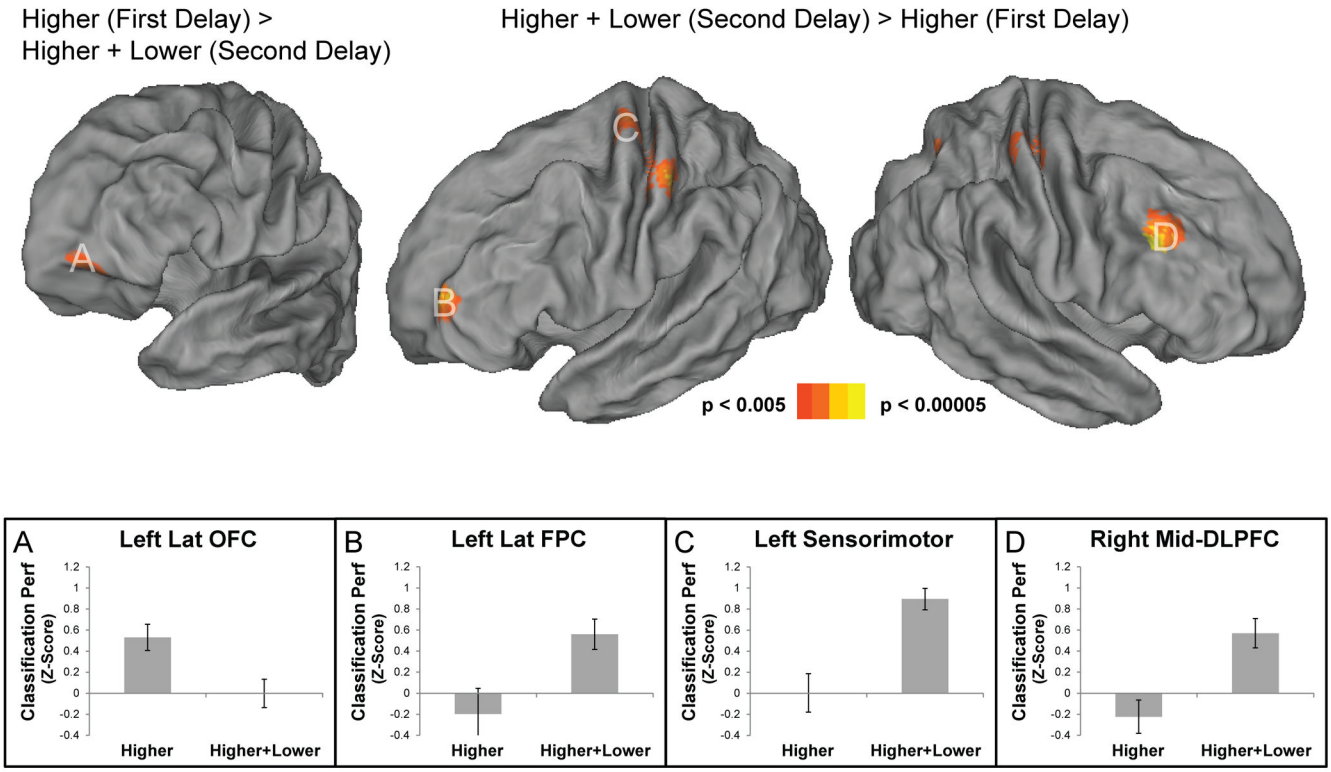

Figure 3.

Whole-brain contrasts. A) The left lateral orbitofrontal cortex (OFC) demonstrated significantly stronger classification of the higher level context than the combination of higher and lower level contexts. B–D) The left lateral frontopolar cortex (FPC), sensorimotor cortex, and mid-dorsolateral prefrontal cortex (DLPFC) demonstrated the opposite pattern.

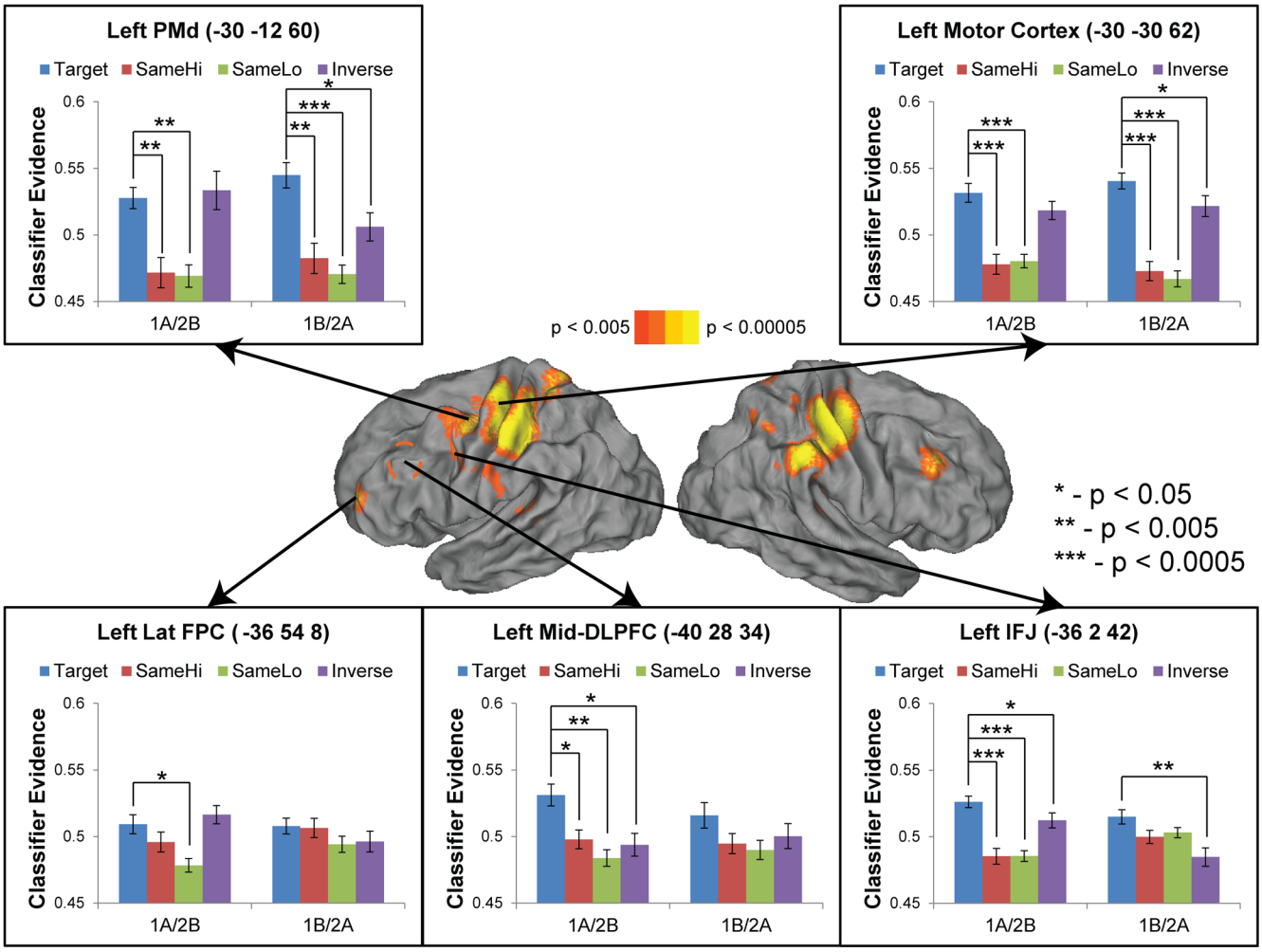

Figure 4.

Representational discrimination analyses. For each test pattern, the machine learning classifier generated a metric of evidence (classifier evidence) for each category (context combination). Significantly greater evidence for the correct category (target) than incorrect categories demonstrates distinct representational patterns. The degree to which evidence is strong for a given incorrect category demonstrates that patterns are confusable. The dashed circle represents the left mid-DLPFC region interrogated by flipping observed right mid-DLPFC activations about the x-axis. Target – correct category; SameHi – category matching the target in the higher level context; SameLo – category matching the target in the lower level context; Inverse – category opposite the target in both the higher and lower level context.

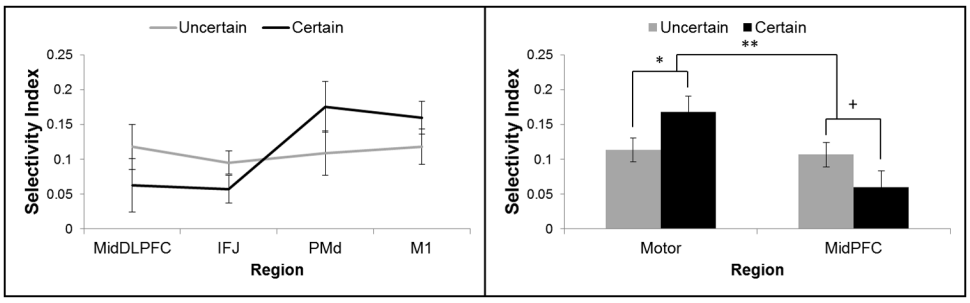

Figure 5.

Representational selectivity by response certainty and region. Context combinations ‘1B’ and ‘2A’ indicated a non-target response regardless of the forthcoming stimulus. In these cases, the response is certain. Context combinations ‘1A’ and ‘2B’ provide no information regarding the forthcoming response (50% target, 50% non-target). In these cases, the response is uncertain. Representational selectivity was stronger in motor areas (dorsal premotor – PMd; primary motor – M1) when the response was certain relative to when it was uncertain. Mid-lateral PFC areas (mid-dorsolateral PFC – DLPFC; inferior frontal junction – IFJ) showed the opposite pattern.

Table 1

Whole-brain MVPA searchlight results

	x	y	z	extent	z	Area	Region
Higher Level Context	−30	54	−18	372	4.24	10,11	left lat OFC
	−14	60	−8		3.13	10,11	left FPC/VMPFC
	32	66	0	293	3.82	10,11	right lat FPC/OFC
	26	60	−16		2.89	11	right lat OFC
	10	40	−22	172	3.57	14	right VMPFC
Higher + Lower Level Context	−40	−30	50	6714	5.05	3	left postCG
	−30	−30	62		4.30	3,4	left central sulcus
	−30	−24	70		4.18	4	left preCG
	−24	−44	70		4.09	1,2,7	left postCG/SPL
	−30	−12	60		3.90	6	left PMd
	−8	−58	44		3.84	7	left preCuneus
	−36	2	42		3.70	6,8	left IFJ
	40	−24	56	4095	4.95	4	right preCG
	54	−36	44		4.49	40	right SMG/IPL
	46	−36	40		4.17	40	right SMG/IPL
	2	−2	44	650	4.14	24,32	mid-cingulate
	−8	−12	50		3.56	23	mid-cingulate
	6	4	60		2.90	6	SMA
	40	28	34	284	3.96	9/46	right mid-DLPFC
	−36	54	8	175	3.58	10	left lat FPC

Abbreviations: DLPFC – dorsolateral prefrontal cortex; FPC – frontopolar cortex; IFJ – inferior frontal junction; IPL – inferior parietal lobule; OFC – orbitofrontal cortex; PMd – dorsal preMotor cortex; postCG – postCentral gyrus; preCG – preCentral gyrus; SMG – supramarginal gyrus; SPL – superior parietal lobule; VMPFC – ventromedial prefrontal cortex

Table 2

Whole-brain MVPA contrasts

	x	y	z	extent	z	Area	Region
Higher > Higher + Lower	−30	54	−14	82	3.26	10,11	left lat OFC
Higher + Lower > Higher	40	24	30	196	4.10	9/46	right mid-DLPFC
	−44	−26	50	723	4.05	4	left central sulcus/postCG
	−28	−20	72		3.17	4	left preCG
	−30	−28	58		2.95	3,4	left central sulcus
	−30	−10	68		2.88	6	left PMd
	2	−56	50	505	3.60	7	preCuneus
	28	−64	52		3.49	7	right intra-parietal sulcus
	16	−56	52		2.78	7	right preCuneus/SPL
	−34	52	12	90	3.48	10	left lat FPC
	36	−24	56	243	3.12	4	right preCG

Neuroimage. Author manuscript; available in PMC 2013 November 15.

Table 3

Representational Discrimination Analyses

Region	Context Combination	Discrimination	t-stat	p-val
lat FPC	1A/2B (Uncertain)	SameHi	1.528	0.142
		SameLo	***2.896***	***0.009***
		Inverse	−0.602	0.554
	1B/2A (Certain)	SameHi	0.136	0.893
		SameLo	0.852	0.404
		Inverse	0.913	0.372
mid-DLPFC	1A/2B (Uncertain)	SameHi	***2.970***	***0.008***
		SameLo	***3.873***	***<0.001***
		Inverse	***2.560***	***0.019***
	1B/2A (Certain)	SameHi	1.478	0.155
		SameLo	1.711	0.103
		Inverse	1.056	0.304
IFJ	1A/2B (Uncertain)	SameHi	***4.354***	***<0.001***
		SameLo	***6.578***	***<0.001***
		Inverse	***2.180***	***0.041***
	1B/2A (Certain)	SameHi	1.784	0.090
		SameLo	1.620	0.121
		Inverse	***3.172***	***0.005***
PMd	1A/2B (Uncertain)	SameHi	***3.788***	***0.001***
		SameLo	***4.138***	***<0.001***
		Inverse	−0.330	0.744
	1B/2A (Certain)	SameHi	***3.640***	***0.002***
		SameLo	***5.337***	***<0.001***
		Inverse	***2.595***	***0.017***
M1	1A/2B (Uncertain)	SameHi	***4.053***	***<0.001***
		SameLo	***5.232***	***<0.001***
		Inverse	1.669	0.111
	1B/2A (Certain)	SameHi	***5.866***	***<0.001***
		SameLo	***6.900***	***<0.001***
		Inverse	***2.131***	***0.046***