

Functional Magnetic Resonance Imaging Evidence for a Hierarchical Organization of the Prefrontal Cortex

David Badre and Mark D'Esposito

Abstract

■ The prefrontal cortex (PFC) is central to flexible and organized action. Recent theoretical and empirical results suggest that the rostro-caudal axis of the frontal lobes may reflect a hierarchical organization of control. Here, we test whether the rostro-caudal axis of the PFC is organized hierarchically, based on the level of abstraction at which multiple representations compete to guide selection of action. Four functional magnetic resonance imaging (fMRI) experiments parametrically manipulated the set of task-relevant (a) responses, (b) features, (c) dimensions, and (d) overlapping cue-to-dimension

mappings. A systematic posterior to anterior gradient was evident within the PFC depending on the manipulated level of representation. Furthermore, across four fMRI experiments, activation in PFC subregions was consistent with the sub- and superordinate relationships that define an abstract representational hierarchy. In addition to providing further support for a representational hierarchy account of the rostro-caudal gradient in the PFC, these data provide important empirical constraints on current theorizing about control hierarchies and the PFC. ■

INTRODUCTION

At any time, more actions are available to us than are appropriate for our current situation and goals. Although behavioral flexibility derives from this wealth of options, it also presents a fundamental indeterminacy problem to be solved by the nervous system. To overcome indeterminacy, we use our perceptions, knowledge, and goals to bias selection of relevant alternatives over competitors, termed cognitive control (Passingham & Rowe, 2002; Miller & Cohen, 2001; Desimone & Duncan, 1995; Shimamura, 1995). The prefrontal cortex (PFC) has an established role in such coordination between thought and action. However, the organization of the PFC and the architecture of cognitive control remain controversial (Petrides, 2005; Anderson, Qin, Stenger, & Carter, 2004; Badre & Wagner, 2004; Bunge, 2004; Buckner, 2003; Koechlin, Ody, & Kouneiher, 2003; Duncan & Owen, 2000; Fuster, 1997; Stuss & Benson, 1987).

One influential hypothesis is that the anterior-to-posterior axis of the PFC is organized hierarchically en route to action (Koechlin & Summerfield, 2007; Bunge & Zelazo, 2006; Koechlin & Jubault, 2006; O'Reilly & Frank, 2006; Petrides, 2006; Fuster, 2004; Koechlin et al., 2003; Sakai & Passingham, 2003; O'Reilly, Noelle, Braver, & Cohen, 2002; Christoff & Gabrieli, 2000; D'Esposito, Postle, & Rypma, 2000; Fuster, 1997). The hierarchy hypothesis derives from the central assumption that the frontal lobes are critical for the selection and

execution of action (Fuster, 1997). Broadly construed, this function entails specifying an abstract action goal, like driving to work, down to a concrete instantiation in the form of a particular sequence of neuromuscular outputs.

Structuring action problems of this kind hierarchically has a number of advantages. In particular, hierarchies partition alternatives, making lower-level choices more tractable and easing the “degrees of freedom problem” (Saltzman, 1979; Bernstein, 1967). In a related sense, hierarchies permit the representation of broader, more abstract action goals (i.e., “make coffee”) concurrently with information about more proximate subgoals (i.e., “add grounds”). These and other properties make hierarchical frameworks common in neural and information processing models of complex or sequential action (Cooper & Shallice, 2006; Newell, 1990; Estes, 1972; Miller, Galanter, & Pribram, 1960; Lashley, 1951; although see Botvinick, 2007; Botvinick & Plaut, 2004). To what extent, then, does the rostro-caudal organization of the PFC truly reflect a hierarchical architecture of cognitive control?

Initial support for the hypothesis of hierarchy in the PFC has come from (a) the pattern of connectivity between posterior neocortical regions supporting perceptual hierarchies and the frontal lobes (Fuster, 1997), and (b) the observation across a range of neuroimaging and neuropsychological studies of a functional gradient along the anterior-to-posterior axis of the PFC, whereby progressively anterior subregions of the PFC are associated with higher-order processing requirements of planning and selection of action (Hazy, Frank, & O'Reilly, 2006; O'Reilly & Frank, 2006; Petrides, 2005, 2006; Fuster,

2004; Bunge, Kahn, Wallis, Miller, & Wagner, 2003; Sakai & Passingham, 2003; Braver & Bongiolatti, 2002; Christoff & Gabrieli, 2000). The interpretation of these functional gradients has differed, to some extent, over the way in which the proposed control hierarchy is ranked, or what consistent principle distinguishes one level of the hierarchy from the next. Some theoretical perspectives have emphasized differences in the character of processing, such as the selection of motor actions in posterior PFC regions on the basis of content-specific conditions versus the domain general monitoring of working memory by more anterior regions (Petrides, 2006). Others have focused on whether information monitored and maintained by working memory is internally or externally derived (Christoff & Gabrieli, 2000), whether the system must maintain/implement goals or subgoals (Braver, Reynolds, & Donaldson, 2003; Braver & Bongiolatti, 2002), what level of complexity is entailed by an action rule (Bunge & Zelazo, 2006; Bunge et al., 2003), and whether anterior regions are necessary to actively organize domain-specific working memory processes in posterior regions (Sakai & Passingham, 2003).

The present article focuses on the specific hypothesis that the posterior to anterior axis of the PFC comprises a representational hierarchy, in that the character of processing at all levels of the hierarchy is the same (i.e., computationally uniform) and so processing differences do not necessarily distinguish subregions of the PFC. Rather, it is the nature of the representations with which control interacts that distinguishes levels of the hierarchy and so distinguishes subregions of the PFC along its rostral-caudal axis.

Representational Hierarchy Ranked by the Control Signal: The “Cascade Model”

To date, the strongest evidence in favor of a representational hierarchy in the PFC comes from a functional magnetic resonance imaging (fMRI) experiment that varied whether subjects selected a response on the basis of a sensory input (termed the “sensory level”), a task on the basis of a contextual cue (termed the “context level”), or a response/task on the basis of an episodic context (termed the “episodic level”; Koechlin et al., 2003). Importantly, these three manipulations obtained activation inclusive of the premotor cortex (all three levels), lateral PFC/anterior premotor cortex (sensory and context levels), and anterior lateral PFC (episodic level only). This expanding posterior to anterior spread of activation was interpreted within an elegant information theoretic framework called the “cascade model” (Koechlin & Summerfield, 2007; Koechlin et al., 2003).

The cascade model itself comprises a representational hierarchy in that all levels are computationally uniform and may be understood as resolving competition (entropy) among action representations (i.e., responses or task sets) on the basis of mutual information with var-

ious control signals (i.e., sensory, context, or episodic). Importantly, it is the type of control signal (sensory, context, or episodic) that determines the anterior-to-posterior locus of frontal lobe activity. In other words, the cascade model proposes that “the LPFC [lateral prefrontal cortex] is organized as a hierarchy of representations originating from premotor cortex and processing distinct signals involved in controlling the selection of appropriate stimulus–response associations” (Koechlin et al., 2003, p. 1181). Thus, as an example, if episodic information is necessary to resolve entropy about an action representation, whether a response or a task set, then the anterior lateral region of the PFC will be active in order to mediate this episodic control.

Although fundamental, these data and the associated formal model have left open some important questions regarding the general hypothesis of representational hierarchy and the organization of the PFC. First, the empirical results did not show selectivity to sensory, contextual, and episodic control across the anterior spreading subregions of the PFC. Rather, activation among posterior regions increased proportionally along with activation of more anterior regions (though, importantly, not vice versa). This provided strong empirical evidence of a single dissociation among these regions but not selectivity to a given control signal. Consequently, the interpretation of the functional data is subject to the usual limitations associated with single dissociations.

It is important to note, however, that the single dissociation pattern actually conforms to an important property of the cascade model, namely, that control signals at higher levels are passed to lower-level controllers directly and proportionally. Consequently, information accumulates at lower levels as a result of changes in higher-level control demands. This predicts that as control demands change at higher levels and anterior regions of the PFC become more active, posterior regions should become correspondingly active. Thus, from the strong version of this perspective, it may not be possible to find selective effects in subregions along the anterior-to-posterior axis of the PFC even over larger parameter ranges than were used in this initial study.

Second, although the cascade model differentiates levels on the basis of differences between control signals (episodic, context, and sensory), the operationalization of these differences in the fMRI experiment and how one generally assigns a given control signal to a given level of hierarchy is not always concrete. In the experiment, the episodic level was clearly distinguished from the lower two levels on the basis of whether control signals emanated from perceptual features of the immediate environment (context/sensory) or from knowledge of the “temporal episode.” However, a similar distinction did not differentiate a “sensory” and a “contextual” signal, as both arose from the same aspect of the sensory environment in the actual experiment (a color cue). Recently, Koechlin and Summerfield (2007) have suggested

that the context and sensory levels are distinguished by whether control operates over simple sensory–motor associations versus sets of sensory–motor associations. Hence, from this perspective, upper levels are distinguished by the form of the control signal (temporal distinction) and lower levels are distinguished by the type of representation over which control operates (also see Koechlin & Jubault, 2006). A third alternative is to differentiate all levels on the basis of the abstractness of the selected representation using a classing rule. This idea is tested in the present experiment.

Representational Hierarchy Ranked by Abstraction: The Current Study

The present study sought to test whether the functional gradient along the anterior-to-posterior axis of the PFC derives from a representational hierarchy ranked by the abstractness of the representation to be selected, rather than by the control signal.

First, we sought to apply a consistent principle for ranking representations at all levels of the hierarchy. Specifically, our hierarchical levels are distinguished based on the abstractness of the action representation to be selected. In general, a representation may be defined as more abstract to the extent that it generalizes over specific instances. In other words, a more abstract or superordinate representation comprises a category or class of subordinate representations. Hence, our concept of abstractness derives from a classing rule, implicit in the tree structure of a hierarchy. Consider, as an example, a hierarchy of semantic representations. The concept “dog” is more abstract and is superordinate to the concept “spaniel” because the former is a class that includes spaniels and other dogs. Similarly, the concept “spaniel” is abstract relative to content-specific properties of a spaniel, such as its form, color, sound, and so forth.

Based on this hierarchical definition of abstraction and starting at the lowest level from the concrete motor response, increasingly abstract levels of representation were defined explicitly such that each more abstract level of representation defined a class of representations at the immediately subordinate level. Applying this rule permitted including an additional intermediate level, termed the feature level, that bridged the selection of responses (analogous to “sensory control” in cascade) and the selection of “dimensions” (analogous to “context control” in cascade). Specifically, from concrete to abstract, four fMRI experiments manipulated competition (a) among manual *responses*, (b) among sets of perceptual *feature*-to-response mappings, (c) among perceptual *dimensions* that comprise a set of relevant perceptual features, and (d) among sets of *contextual* cue-to-dimension mappings. We note that because the distinguishing feature of the present framework is that it defines its levels based on the abstractness of the rep-

resentations to be selected, our levels are named based on these selected representations. Hence, our context level does not refer to the use of a context signal for control in the same sense as the cascade model. Rather, in our study, context refers to competition between multiple context-to-dimension mappings (and if anything, corresponds to the episodic control level of the cascade model). As competing representations became more abstract, activation was predicted to progress systematically along the rostral-caudal axis of the PFC, starting from the premotor cortex and progressing to anterior regions of the PFC.

A second novel feature of the present experiment was that, within each level, control demands were varied by parametrically increasing the number of competing alternatives at a particular hierarchical level of representation. Parametric values ranged from no competition, to a mid-level of competition, to a high level of competition. As competition increases, so should control also be increasingly required to adjudicate among alternatives, and the level of activation at a given locus in the PFC should increase accordingly. Thus, in contrast to the hierarchical level that was indexed by *locus of activation* along the rostral-caudal axis, competition was indexed by the *level of activation* at a given locus in the PFC. Furthermore, as processors at different levels of the abstract representational hierarchy are only sensitive to competition at that level, they may exhibit different activation profiles than anterior regions. Consequently, it may be possible to distinguish selective effects across the wider parameter range used in this experiment.

METHODS

Subjects

Nineteen (9 women) right-handed, native English speakers, aged 18–31 years, with normal or corrected-to-normal vision, were recruited for participation in four mini-experiments across two fMRI scanning sessions. Data from an additional two subjects were excluded due to an inability to complete both scanning sessions and significant MRI signal artifacts, respectively. All subjects underwent prescreening for neurological or psychological disorders, as well as contraindications for MRI. Normal color vision was verified for all subjects as assessed by the Ishihara test for color deficiency. Informed consent was obtained from subjects in accordance with procedures approved by the Committee for Protection of Human Subjects at the University of California, Berkeley. Subjects received payment for their participation in both MRI sessions.

Behavioral Tasks

Each of four mini-experiments parametrically varied competition at one of the four hierarchical levels of

representation while maintaining competition at subordinate levels of representation constant and superordinate levels of representation minimal.

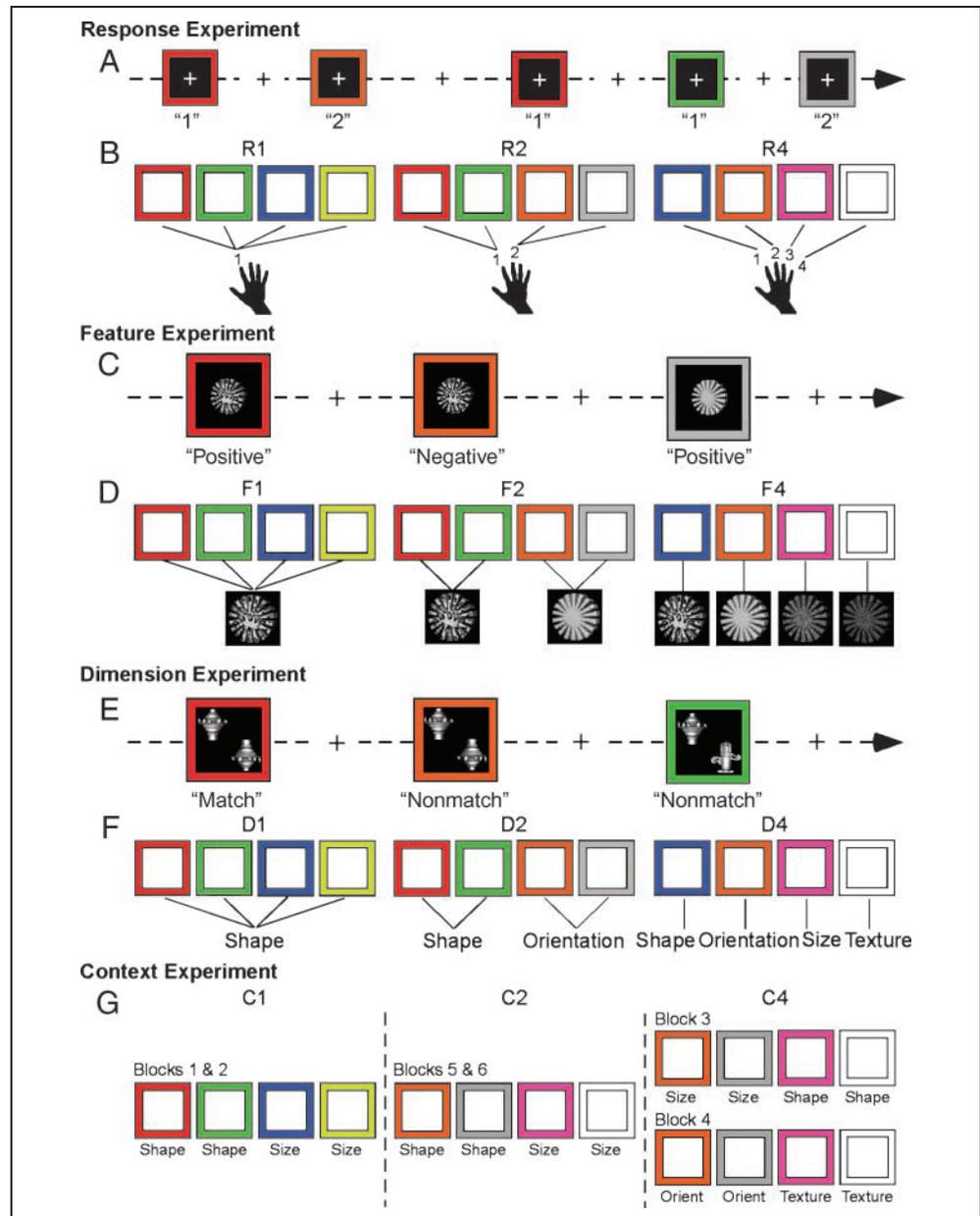
Response Experiment

The response experiment manipulated response competition by varying the number of eligible manual responses on a given block of trials. Blocks consisted of a series of colored squares presented one at a time

(1900 msec followed immediately by a brief noise mask 100 msec; Figure 1A). Trials within a block were separated by jittered fixation-null events (0–6 sec). Based on a learned color-to-response mapping, subjects made a response on a keypad under their right hand depending on the color of the presented square.

As is depicted in Figure 1B, only four color-to-response mappings were relevant for each block. An instruction screen cued subjects which four color-to-response mappings were relevant at the outset of each block (10 sec

Figure 1. Schematic depiction of example trial sequences and rule sets across four mini-experiments. (A) In the response experiment, subjects were presented with a series of colored squares presented one at a time (dashed line represents time), and made a button response (quoted number below square) depending on the color of the square. (B) Four color-to-response mappings were relevant for each block of trials, and on a given block colors mapped to one (R1), two (R2), or four (R4) responses; this increased response competition. (C) In the feature experiment, subjects were presented with a series of colored squares containing a single object that varied along one dimension such as texture in the figure. Subject responded “positive” to a particular feature and “negative” to other features depending on the color of the surrounding square. (D) Four color-to-feature mappings were relevant for each block of trials, and on a given block one (F1), two (F2), or four (F4) features mapped to a positive response, increasing feature competition. (E) In the dimension experiment, subjects were presented with a series of colored squares containing two objects. Subjects decided whether the objects matched or did not match along one dimension. The relevant dimension was cued by the color of the square. (F) Four color-to-dimension mappings were relevant for each block of trials, and on a given block one (D1), two (D2), or four (D4) dimensions were relevant to a match decision, increasing dimension competition. (G) In the context experiment, the task was identical to D2 blocks of the dimension experiment. Context competition was manipulated by varying the frequency with which a given color-to-dimension mapping occurred across all six blocks. This frequency varied from 1 (C1), to 0.5 (C2), to 0.25 (C4). For example, in this figure, gray maps to shape on blocks 5 and 6 but to other dimensions on other blocks. Thus, the frequency of gray to shape is 0.5. Subjects knew which mapping was currently relevant based on the instruction provided at the beginning of their current block.



followed by an 8-sec fixation period). Response competition was manipulated based on the number of candidate responses that the subject could make on a given block (Figure 1B). During 1-Response blocks (R1), all four colors mapped to a single response. Hence, in R1, there was no competition about what response to select. During 2-Response blocks (R2), there were two candidate responses (two of the four colors mapped to each response; Figure 1B), producing response competition. Finally, in 4-Response blocks (R4), each of the four colors uniquely cued one of the four responses (Figure 1B), entailing the greatest response competition. As the same cue dimension (color) and cue features (individual colors) were always relevant to a given response, competition at superordinate levels (i.e., feature, dimension, and context) was minimal.

Feature Experiment

The feature experiment manipulated feature competition by varying the number of perceptual features (i.e., the number of specific textures of an object or the number of specific orientations of an object) that could map to a given response. Subjects were presented with a series of colored squares one at a time (3900 msec, followed immediately by a noise mask of 100 msec), and each square contained a single object (Figure 1C). Trials were separated by a jittered null fixation interval (0–4 sec).

From trial to trial, the object varied along one dimension (either texture or orientation between subjects) and all other dimensions were held constant (Figure 1C). This variable dimension could have four particular instances, termed features. For example, if the variable dimension was orientation, the four features were up, down, left, and right. On each trial, a particular feature was relevant (e.g., a mottled texture). If the relevant feature was presented, the subject made a positive response by pressing a key on the keypad (Figure 1C). Otherwise, the subject pressed a second key indicating a negative response (e.g., for any other texture). The subject knew which feature was relevant on a particular trial on the basis of the color of the square that surrounded the object. Hence, there was not enough information in just the object feature or the color alone to determine the correct response. The object feature had to be considered in conjunction with the color to make a response. Analogous to the response experiment, four color-to-feature mappings were relevant for each block of trials. The only difference from the response experiment was that the colors now mapped to relevant features that cued a correct response, rather than mapping directly to the correct response (Figure 1D).

Feature competition was manipulated in a blocked fashion by varying the number of feature-to-response sets potentially relevant on a given block (Figure 1D). On alternate blocks, one (F1), two (F2), or four (F4)

feature-to-response sets could be relevant with equal frequency, thereby parametrically increasing feature competition. Again, which feature set is relevant is determined by the color, so once feature competition is present (F2 and F4 blocks), the particular color or object feature by itself is insufficient information to determine the correct response.

As the relevant dimension (such as orientation or texture) and the color-to-feature mappings were invariant across all conditions of the feature experiment, competition at superordinate levels (i.e., dimension and context) was minimal. Likewise, once a decision about the relevant set of feature-to-response mappings had been made based on the color, two responses (positive and negative) were always equally relevant, so competition at the subordinate response level was maintained theoretically equivalent to R2 blocks of the response experiment in which there is also competition between two responses. Hence, as F1 blocks contain no feature-level competition but include competition between two responses, these blocks and R2 are theoretically analogous despite superficial differences between the experiments.

Dimension Experiment

The dimension experiment manipulated dimension competition by varying the number of perceptual dimensions (i.e., classes of features such as size, orientation, shape, and texture) that determine the action relevant features of an object. Subjects were presented with a series of colored squares one at a time containing two objects (3.9 sec followed by a 100-msec noise mask; Figure 1E). Individual trials were separated by a jittered null fixation interval (0–4 sec). Subjects were required to compare the two objects along one of four dimensions (texture, orientation, shape, or size) and indicate with a keypress whether the objects matched or mismatched along only the relevant dimension. On a given trial, the subject knew which of the four dimensions was relevant to the match decision on the basis of the color of the square bounding the objects. Again, four color-to-dimension mappings were relevant for each block. Hence, the design for the mappings was identical to the feature and response experiments, except that color mapped to dimension rather than feature or response.

Competition at the dimension level was manipulated in a blocked fashion on the basis of the number of candidate dimensions that could be relevant to the match decision for a given block. On alternate blocks, one (D1), two (D2) or four (D4) dimensions could be relevant with equal frequency, thereby parametrically increasing dimension competition (Figure 1F).

The color-to-dimension mappings were invariant across conditions, thus competition at the superordinate context level was constant and minimal. There were

always two responses (match/nonmatch), and the match decision entailed two sets of feature–response relationships distinguished by a contingent relationship between two features (i.e., two competing feature–response mapping sets). Hence, subordinate response and feature competition were maintained theoretically equivalent to F2 blocks of the feature experiment. D1 conditions do not have competition at the dimension level and so are theoretically analogous to F2 conditions of the feature experiment, despite other superficial differences between these experiments.

Context Experiment

The context experiment manipulated context competition among overlapping mappings between a particular contextual cue (i.e., a particular colored square) and different dimensions during the match/nonmatch task. Competition was manipulated by varying the frequency with which a given color mapped to a particular dimension, such that less frequent mappings suffer from greater competition from more frequent mappings. The subjects' task in the context experiment involved making a match/nonmatch decision about two objects in a colored square, and was identical in its details to that described for the Dimension experiment (Figure 1E), except that two dimensions were always relevant to the match decision across all competition conditions of the context experiment. Hence, dimension competition was held constant at two dimensions.

Competition about the context was varied based on the frequency with which a given colored square mapped to a particular dimension across all blocks of the experiment. The four colors presented during 1-Context blocks (C1), always cued their respective dimensions (e.g., red always cued “shape” across all blocks of the experiment). In other words, there were no blocks during which a color found in a C1 block cued a different dimension. Hence, during these blocks, there was no competition for the mapping between a given color and a given dimension from mappings of that color to other dimensions. By contrast, during 2-Context blocks (C2), a given color mapped to its corresponding dimension 50% of the time across all blocks of the experiment (e.g., orange maps to shape during 50% of blocks but to other dimensions during other blocks). This arrangement results in competition between a given color-to-dimension and other color-to-dimension mappings of the same color. For example, conflict arose if on some blocks gray mapped to shape, whereas on other blocks gray mapped to other dimensions. In these cases, mere presentation of the color (e.g., gray) was insufficient to completely specify the relevant dimension. Rather the subject had to rely on additional information, such as an episodic context like a current block (Koechlin et al., 2003), to determine which mapping is currently relevant.

Finally, on 4-Context blocks (C4), the colors mapped to their corresponding dimension on only 25% of blocks. Thus, color-to-dimension mappings on C4 blocks entailed the greatest competition from other mappings.

In the context experiment, only the frequency of color-to-dimension mappings varied. Two dimensions, features, and responses were always relevant. Consequently, competition at all subordinate levels was held constant across context competition conditions equivalent to the D2 condition of the dimension experiment. As there is no context-level competition during C1 conditions and yet subjects make a match/nonmatch decision based on one of two candidate dimensions, C1 blocks are identical to D2 blocks of the dimension experiment.

Nested Design

Figure 2 depicts schematically how the four experiments described here derive from a single classing rule. Representations at higher levels represent a set of representations one level down. The tree thus continues all the way down to its termination at the level of response.

Based on this structure, a nested design was employed such that the lowest parametric level of conflict in each experiment (F1, D1, and C1) was theoretically equivalent (in terms of conflict at all hierarchical levels) to the mid-conflict condition of the immediately subordinate experiment (R2, F2, and D2, respectively). For example, on F1 blocks, one feature maps to one response 50% of the time and the other features map to the other response 50% of the time. Consequently, although there is no feature conflict, there are two responses equivalent to R2. This same pattern is replicated all the way up to the context experiment in which C1 blocks have no context uncertainty but require a choice between two dimensions.

For mid- and high-level conflict conditions (e.g., F2 and F4, D2 and D4, C2 and C4), competition at lower levels (response; response and feature; and response, feature, and dimension, respectively) was maintained constant, assuming that competition at higher levels is fully resolved. To illustrate this idea, consider the F4 condition of the feature experiment. Once the color decision is made and the currently relevant feature-to-response mapping set is selected, the task reduces to the F1 condition, in that one feature maps to one response and the remaining features map to the other response. Thus, the nested design, with its explicit definition of conflict at each level, permitted the specification of a minimum degree of conflict expected at each level regardless of conflict at higher levels. Critically, under this controlled design, deviations from this minimum in subordinate regions must be due to the conflict at the highest, parametrically varying level in a given experiment, and so this provides insight into the interaction

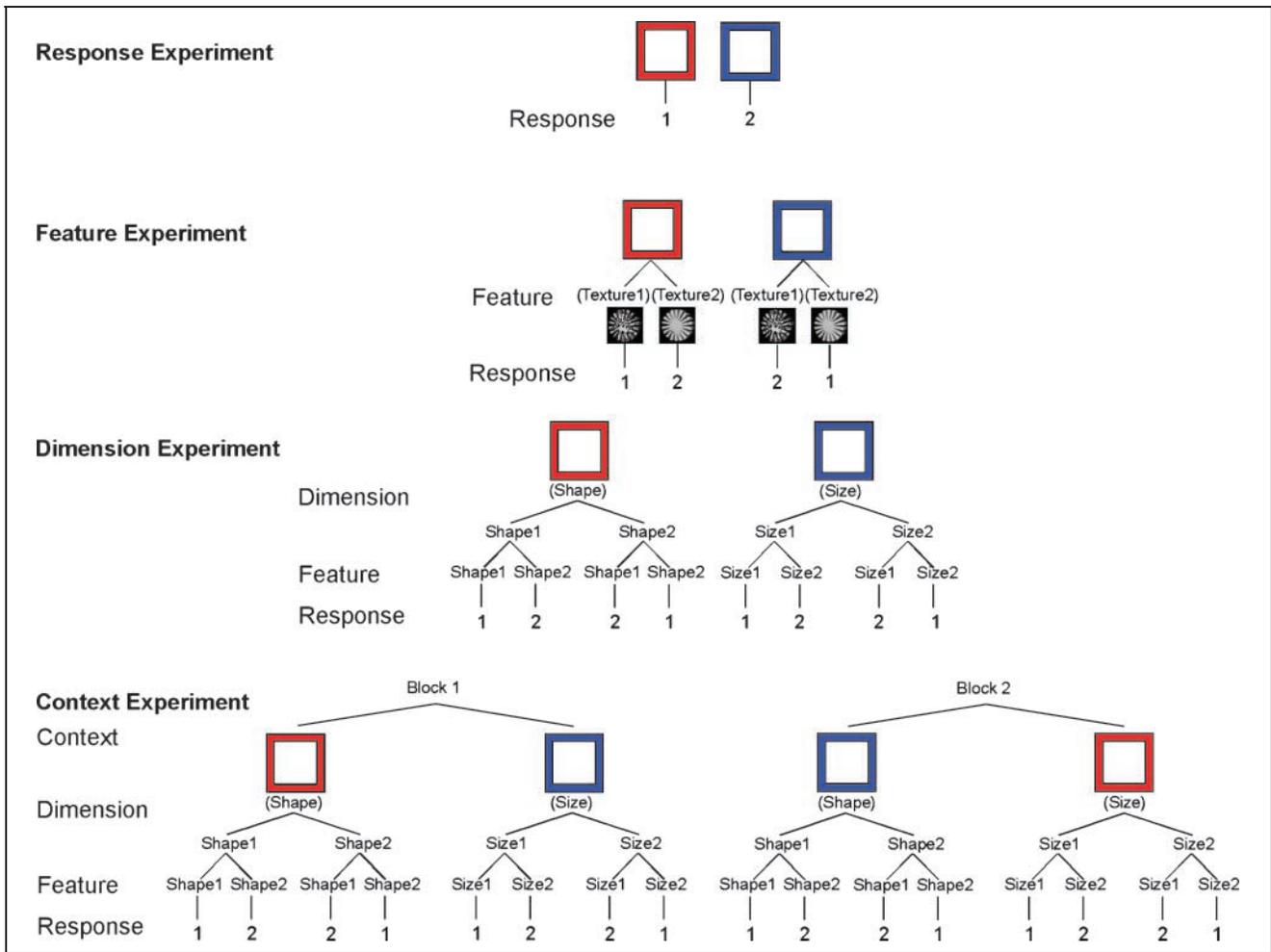


Figure 2. Schematic examples illustrating the hierarchical structures underlying the four fMRI experiments. Representations were defined as more abstract to the extent that they classed representations at the next subordinate level. Each experiment, response, feature, dimension, and context, included control demands at an additional level of abstraction. The examples schematized here are based on two alternatives at the highest level of each experiment (R2, F2, D2, and C2 conditions) and show specific examples of dimensions and features. In the actual experiment, alternatives could increase to four (see Figure 1), and the specific dimensions and features were counterbalanced within and between subjects. The level of abstraction at which control operates was predicted to determine the locus of activation along the posterior-to-anterior gradient in the PFC.

between levels, including accumulation effects, such as those predicted by cascade.

Stimuli

Across all experiments, cues consisted of centrally presented colored squares (subtending $\sim 6^\circ$ of visual angle). For a given experiment, one of two distinct color sets of eight colors was used twice across the four experiments, and the order was counterbalanced across subjects. Piloting determined that colors within a color set were easily distinguishable.

Object stimuli used in the feature, dimension, and context experiments consisted of gray-scale, 3-D shapes. Objects were constructed using the Strata 3D™ software and were designed to be unfamiliar, difficult to name, and without real-world counterparts. Piloting deter-

mined that these feature variations in these objects were easily distinguishable.

General Experimental Protocol

The four mini-experiments (response, feature, dimension, context) were tested in two fMRI scanning sessions. Two experiments were conducted per scanning session. Experiments were counterbalanced for order across subjects, with the exception that the context experiment was always performed first. Performing context first ensured that there was no overlap between the color mappings manipulated in this mini-experiment and those in the other mini-experiments. The largest potential for contamination of the context manipulation comes from the dimension experiment. Hence, the dimension experiment was always performed in the session after

the context experiment, and using a different color set. Twenty-four hours to one week was permitted to elapse between scanning sessions.

Subjects performed 192 trials of each experiment divided into eight 8-trial blocks of each competition condition. Blocks were separated by a 12-sec null fixation period and were fully counterbalanced for order for the response, feature, and dimension experiments. In the context experiment, the block condition order was fixed within each of the four fMRI scanning runs. Specifically, subjects engaged first in two C1 blocks followed by two C4 blocks and concluded with two C2 blocks. Following the logic of Koechlin et al. (2003), the fixed order allowed subjects to use the block sequence as an additional cue to resolve context competition. The C1, C4, C2 ordering uncoupled practice or fatigue effects from the predicted parametric context competition function.

Across all blocks of the response, feature, and dimension experiments, a total of eight specific color mappings were combined with repetition of some mappings into three sets of four mappings, with each set of four mappings being used for all blocks of a given competition condition. The repetition of a given color mapping followed a specific algorithm, such that the first two colors from the low conflict conditions (e.g., R1, F1, D1) were repeated in the mid-conflict condition (e.g., R2, F2, D2), whereas the third color from the low conflict condition and the third color from the mid-conflict condition were repeated in the high conflict condition (e.g., R4, F4, D4). Figure 1B, D, and F illustrate an example using specific color mappings. Color mappings were counterbalanced between subjects.

Where applicable in each experiment, cue switches (from one color to another), response switches (from one response to another), feature switches, and dimension switches were controlled for frequency and repetition across blocks of each condition. Within block trial orders were generated using an algorithm that optimized the efficiency of the design matrix to permit event-related analysis, and efficiency was equated across competition conditions. All combinations of colors and features in the feature experiment and colors and shapes in the dimension and context experiments were controlled across competition and switching conditions.

fMRI Procedures

At the beginning of each scanning session, subjects trained and practiced each of the two tasks to be tested during that session. During initial training outside the scanner, subjects were taught the mappings they would encounter that session and then practiced both outside and inside the scanner.

Whole-brain imaging for all four experiments was performed on a 4-T Varian/Inova (Palo Alto, CA) MRI system using a standard radio-frequency head coil. Functional

data were acquired using a two-shot gradient-echo, echo-planar pulse sequence (TR = 2 sec, TE = 28 msec, 18 axial slices, $3.5 \times 3.5 \times 5$ mm, 0.5 mm interslice gap, 204/228 volume acquisitions for the response/feature, dimension, and context experiments). Prior to collection of functional data, T1-weighted anatomical images in the same plane as the functional images were acquired using a gradient-echo multislice sequence (GEMS). High-resolution T1-weighted (MP-FLASH 3-D) anatomical images were also collected at the conclusion of each MRI session for anatomical visualization. Head motion was restricted using firm padding that surrounded the head. Visual stimuli were displayed using Psychtoolbox (www.psychtoolbox.org) on a Macintosh computer running MATLAB, and were projected onto a screen that was viewed through a mirror attached to the head coil.

fMRI Analysis

Functional images were phase map corrected and differences in slice timing were corrected using sinc interpolation. Functional data were interpolated to 1 sec resolution through bilinear interpolation of each shot (half of k -space) with its two flanking shots.

Following slice timing correction, global signal was assessed for outliers or artifacts in mean signal, global variance, and slice-to-slice variance. Large outliers in global signal were evident in three subjects during the feature experiment. These outlier volumes were replaced by the global mean signal, effectively removing these volumes from the time series. Trimming was done in an unbiased fashion based on the global signal and blind to experimental conditions.

All subsequent preprocessing was conducted using SPM2 (Wellcome Department of Cognitive Neurology, London). Motion correction was applied across all runs using sinc interpolation. Following registration to the anatomical GEMS images, the MP-FLASH T1 image was normalized to a T1 template on the basis of MNI stereotaxic space using a 12-parameter affine transformation along with a nonlinear transformation using cosine basis functions. Functional data were then normalized based on parameters derived from the normalization of the MP-FLASH T1 image. Images were resampled into 3-mm cubic voxels and then spatially smoothed with an 8-mm full-width half-maximum isotropic Gaussian kernel.

As each mini-experiment constituted an independent dataset, separate statistical models were constructed for each of the four experiments (i.e., response, feature, dimension, and context) under the assumptions of the general linear model. Region-of-interest (ROI) analysis (see below) was used for comparison between experiments. Epochs corresponding to each block of trials within a session were included in the statistical model along with regressors for the instruction periods at the beginning of each block. Furthermore, blocks of each

competition condition were weighted as a parametric regressor in the model.

Statistical effects were estimated using a subject-specific fixed-effects model, with session-specific effects and low-frequency signal components (<0.01 Hz) treated as confounds. Linear contrasts were used to obtain subject-specific estimates for each effect. These estimates were entered into a second-level analysis treating subjects as a random effect, using a one-sample *t* test against a contrast value of zero at each voxel. Voxel-based group effects were considered reliable to the extent that they consisted of at least 5 contiguous voxels that exceeded an uncorrected threshold of $p < .001$. For the purpose of additional anatomical precision, group contrasts were also rendered on an MNI canonical brain that underwent cortical “inflation” using FreeSurfer (CorTechs Labs) (Dale, Fischl, & Sereno, 1999; Fischl, Sereno, & Dale, 1999).

Blocked parametric effects obtained in the group-level voxel-based contrasts were supplemented with ROI analyses. ROI analysis was used (1) to verify the characteristic event-related signal change in voxels identified by a block-level contrast, (2) to confirm parametric effects when analysis is restricted to correct trials, (3) to independently test parametric effects on switch and repeat trials, and (4) to test predictions about each level of competition across independent experiments. ROIs were thus defined based on significant voxels within a 6-mm radius of a chosen maximum from the blocked parametric contrast at each level of representation. Selective averaging with respect to peristimulus time was conducted using the Marsbars toolbox (Brett, Anton, Valabregue, & Poline, 2002), and allowed assessment of the signal change associated with each condition. Peak percent signal change was computed based on the integral of the peak across conditions, plus and minus one timepoint. All ROI and behavioral data were subjected to repeated-measures analyses of variance.

RESULTS

Dorsal Premotor Cortex is Sensitive to Response Competition

Increasing competition at the response level was associated with a parametric increase in median RT [$F(2, 36) = 181.5, p < .0001$; Figure 3A]. The proportion of trials on which a response switches from the previous trial necessarily increases with the number of candidate responses. Switching of responses or tasks can result in an increase in RT relative to repeating responses or tasks (e.g., Monsell, 2003). However, increases in the proportion of response switching across competition conditions did not account for the increase in RT, as a parametric increase in RT was still evident when analysis was restricted to trials on which the cue and response repeated [$F(2, 36) = 73.0, p < .0001$; Figure 4B]. Over-

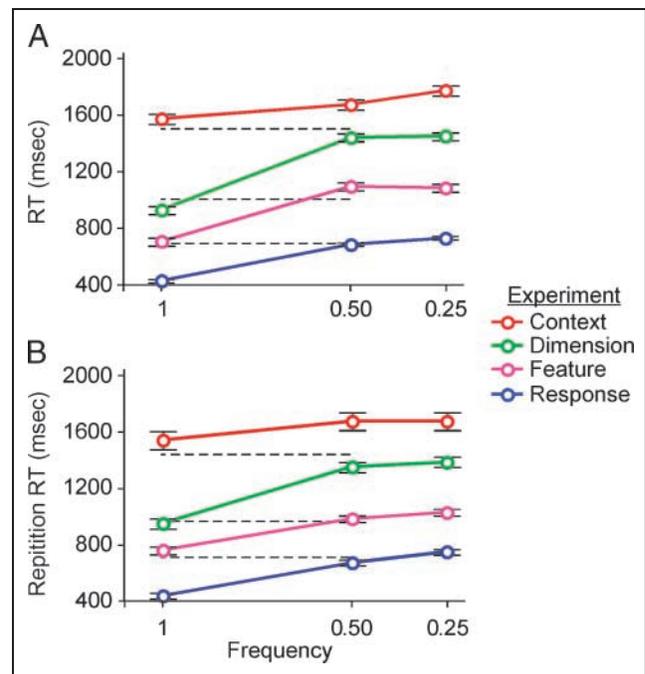


Figure 3. Median RT from the four fMRI experiments. Notably, although increases at higher competition conditions are evident across all correct trials (A), the parametric effect is also evident when analysis is restricted to correct repeat trials (B). Dashed lines connect theoretically equivalent conditions across experiments.

all, subjects made few errors in the response experiment (4% errors, $SD = 3\%$), and errors nonsignificantly increased from R1 to R4 ($F = 1.9$).

Increased response competition across blocks was associated with a parametric activation increase in the left dorsal premotor cortex (PMd; $-30 -10 68; -24 -6 58$; Figure 4A). Activation also extended posteriorly to the left primary motor cortex (M1; $-36, -24, 56$) and the superior portion of the intraparietal sulcus (IPS; $-28 -50 42$), and medially along the cingulate sulcus between supplementary motor area (SMA) and the anterior cingulate cortex (ACC; $-8 -6 54; -6 12 46$).

ROI analyses confirmed that blocked parametric effects in the PMd were evident in the event-related signal [$F(2, 36) = 21.0, p < .0001$; Figure 4A]. Furthermore, there was no reliable effect of switching the response cue on activity in the PMd ($F = 2.5$), nor did the parametric effect of response competition interact with switching ($F = 0.25$).

Response competition should be present, but constant, across competition conditions of the feature, dimension, and context experiments as each require selection between two responses. Consistent with this hypothesis, PMd activation across the feature, dimension, and context experiments was greater than in the R1 condition of the response experiment ($F_s > 11.4, p_s < .0001$; Figure 4A). Moreover, there were no differences in PMd activation across the competition conditions of these superordinate experiments ($F_s < 1.3$), nor was

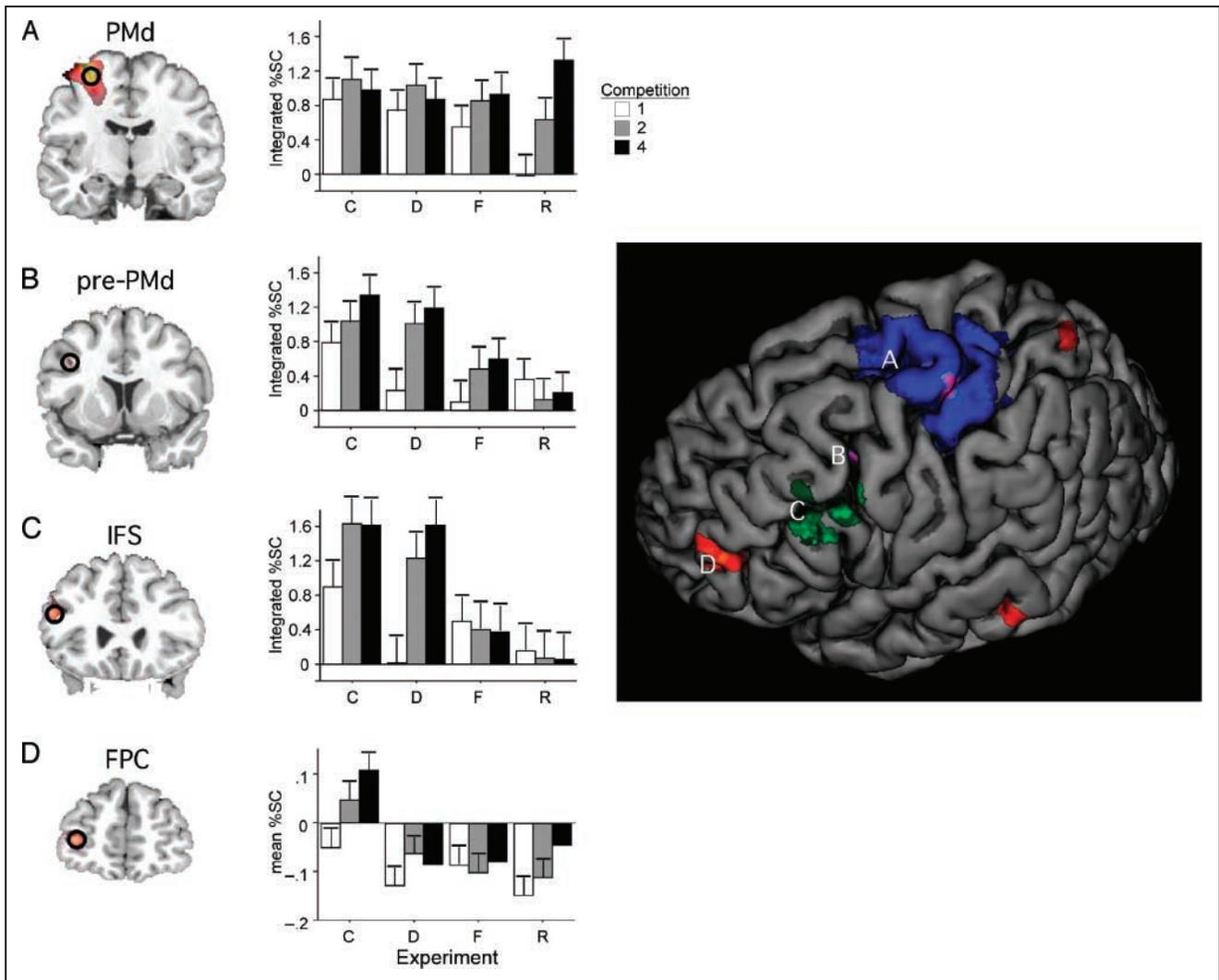


Figure 4. Whole-brain analysis of parametric effect across the four experiments. A rostro-caudal gradient was evident as the level of abstraction increased. Specifically, response competition activated the PMd (A), feature competition activated the pre-PMd (B), dimension competition activated the IFS (C), and context competition activated the FPC (D). Graphs from ROI analyses for the PMd, pre-PMd, and IFS depict the integrated peak signal for each parametric level in each experiment (C = Context, D = Dimension, F = Feature, and R = Response). Bars for FPC depict the mean percent signal change over the entire block.

response switching reliable ($F_s < 2.4$). The between-experiment difference in PMd sensitivity to competition conditions was reliable [Experiment \times Competition: $F(6, 108) = 3.7, p < .005$].

Anterior Dorsal Premotor Cortex is Sensitive to Feature Competition

RT increased with competition at the feature level [$F(2, 36) = 61.1, p < .0001$; Figure 3A]. However, there was little difference between F2 and F4 blocks, with RT actually decreasing nonsignificantly (~ 14 msec). Notably, F2 blocks contain trials on which the cue switches but the response does not, an arrangement that can elicit a substantial behavioral switch cost (Logan & Bundesen, 2004). This could have inflated F2 RT. In-

deed, when RT analysis is restricted to trials on which the cue and feature repeat from the previous trial, RT demonstrates the predicted quantitative parametric increase from F1 to F2 to F4 [$F(2, 36) = 36.0, p < .0001$; Figure 3B]. Error rates were again low overall (3%), but reliably mirrored RT effects across competition conditions [$F(2, 36) = 9.1, p < .001$]. Finally, the F1 condition has no feature competition, but two responses, and so is theoretically equivalent to the R2 condition of the response experiment. Consistent with this hypothesis, there was no reliable difference in RT between F1 (714 msec) and R2 [685 msec; $F(1, 18) = 1.2, p = .28$].

As predicted, the parametric effect of increased feature competition across blocks was associated with increased activation in the left anterior dorsal premotor cortex (pre-PMd; $-38\ 10\ 34$; Figure 4B). Also associated

with the parametric increase in feature competition were the left M1 (-44 -24 58), left ventral IPS (-28 -68 26), right fusiform (28 -52 -4), and left parahippocampus (-38 -16 -14).

The association of the pre-PMd with the parametric increase in feature competition was confirmed in the event-related response [$F(2, 36) = 3.3, p < .05$; Figure 4B]. The pre-PMd was not sensitive to switching the feature cue ($F = 0.04$), and there was no interaction of feature competition with switching ($F = 2.3$).

Feature competition was hypothesized to be minimal and constant during the subordinate response experiment. Consistent with this hypothesis (Figure 4B): (A) activation in the pre-PMd during the response experiment did not reliably differ from F1 blocks ($F = 0.66$) or across response competition conditions ($F = 0.68$); (B) F2 and F4 activation in the pre-PMd was reliably greater than at any point during the response experiment [$F(1, 18) = 5.0, p < .05$].

The superordinate dimension and context experiments were hypothesized to include constant feature competition. Consistent with this hypothesis, activity in the pre-PMd during the dimension and context experiments was reliably greater than the F1 condition of the feature experiment ($F_s > 17.3, p_s < .0001$). However, activation in the pre-PMd was not constant during the dimension experiment. Pre-PMd activation was greater during D2 and D4 conditions than during D1 conditions [$F(2, 36) = 10.9, p < .0005$]. There was a reliable effect of cue switching in pre-PMd during the dimension experiment when analysis was restricted to D2 and D1 conditions [$F(1, 18) = 7.5, p < .01$], and the effect of D2 and D4 over D1 was highly reliable on switch trials [$F(1, 18) = 26.8, p < .0001$], but was only a trend on repeat trials [$F(1, 18) = 3.9, p = .055$]. This difference resulted in a borderline interaction of switching with dimension competition [$F(2, 36) = 2.98, p = .06$]. Finally, the signal in pre-PMd trended parametric in the context experiment [$F(2, 36) = 3.04, p = .06$].

Overall, however, the difference between the parametric effect obtained in the feature experiment and the patterns of pre-PMd activation from the other experiments was reliable [Experiment \times Competition: $F(6, 108) = 3.0, p < .01$]. Furthermore, the selective responses of the PMd and pre-PMd to response and feature competition dissociated these regions [Region \times Experiment \times Competition: $F(2, 36) = 8.5, p < .001$].

The Inferior Frontal Sulcus is Sensitive to Dimension Competition

Increasing competition at the dimension level was reflected in a corresponding increase in median RT [$F(2, 36) = 101.7, p < .0001$; Figure 3A]. Furthermore, this effect was not dependent on switching the relevant dimension, as the effect of RT was reliable for cue and dimension repeat trials [$F(2, 36) = 44.4, p < .0001$;

Figure 3B]. Errors were minimal overall (4%), and the nonsignificant quantitative pattern mirrored RT ($F = 2.2$). D1 conditions have no dimension competition, but do have feature competition roughly equivalent to the F2 condition of the feature experiment. However, F2 RT (1097 msec) was reliably longer than D1 RT [934 msec; $F(1, 18) = 17.3, p < .0005$]. This difference could be due to the cue switching effect on F2 trials. Indeed, when the analysis is restricted to cue repeat trials from the two experiments, there is no difference between F2 (965 msec) and D1 [936 msec; $F(1, 18) = 0.42, p = .52$].

The blocked parametric effect of dimension competition (Figure 4C) was evident along the left inferior frontal sulcus (IFS; -50 26 24; -52 28 38; -32 18 24), extending ventrally to the triangularis portion of the inferior frontal gyrus (mid-VLPFC; -52 18 14). Activation was also present in the left pre-PMd (-48 2 26; -58 6 42; -56 12 32), left IPS (-24 -60 32), and right fusiform (44 -56 -10).

The parametric response to dimension competition in the left IFS was also evident in the event-related responses across the IFS foci ($F = 16.7, p < .0001$). Furthermore, ROI analysis revealed that activation in the IFS was not reliably affected by switching versus repeating the dimension cue ($F = 1.2$), nor was the parametric effect of dimension competition impacted by cue switching ($F = 2.5$).

The subordinate response and feature experiments were hypothesized to entail minimal dimension competition as the relevant dimension(s) was constant. Consistent with this hypothesis: (a) Activation in the dorsal IFS was stable across conditions of the feature and response experiments ($F_s < 0.2$; Figure 4C); (b) Activation in the IFS during the response experiment was never greater than during baseline blocks (D1) of the dimension experiment ($F = 0.13$) and only trended greater across conditions of the feature experiment [Figure 4C; $F(1, 18) = 3.7, p = .06$]; (c) Activation in D2 and D4 conditions of the dimension experiment was significantly greater than at any point in the feature or response experiment (Figure 4C; $F_s > 34.1, p_s < .0001$).

All conditions of the superordinate context experiment were hypothesized to elicit constant dimension competition. Consistent with this hypothesis, activation in the IFS during the context experiment was consistently greater than during D1 blocks [Figure 4C; $F(1, 18) = 40.6, p < .0001$]. Notably, the activation was not entirely constant across context competition conditions, as activation in the IFS was greater on C2 and C4 blocks than C1 blocks [Figure 4C; $F(2, 36) = 5.7, p < .01$]. Interestingly, there was also a reliable main effect of cue switching in the IFS during the context experiment [$F(1, 18) = 6.4, p < .05$].

Critically, however, the selective parametric pattern in IFS during the dimension experiment resulted in a reliable Experiment \times Competition interaction [$F(6, 108) = 5.4, p < .0001$]. Indeed, the differential responses to re-

sponse and dimension competition dissociated the PMd and IFS [Region \times Experiment \times Competition: $F(2, 36) = 29.5, p < .0001$], and the differential responses to feature and dimension competition dissociated the pre-PMd and IFS [Region \times Experiment \times Competition: $F(2, 36) = 9.5, p < .0005$].

Frontal Polar Cortex is Sensitive to Context Competition

RT increased with context competition [$F(2, 36) = 17.3, p < .0001$; Figure 3A]. Unlike the subordinate experiments, the proportion of switches did not increase along with context competition. Hence, switching cannot account for the context competition effect on RT. Overall, subjects performed the test accurately (6% errors). Notably, there was an effect of competition on errors [$F(2, 36) = 13.4, p < .0001$], reflecting higher errors on C2 blocks (10% errors) than on C1 [4% errors; $F(1, 18) = 23.3, p < .0001$] or C4 blocks [5% errors; $F(1, 18) = 16.3, p < .0005$]. Finally, despite being structurally identical tasks, there was a reliable difference between C1 (1582 msec) and D2 [1443 msec; $F(1, 18) = 11.9, p < .005$], although the faster RT on D2 likely stems from the benefit of greater practice by the time subjects perform the dimension experiment.

The blocked parametric effect of context competition was evident in the left frontopolar cortex (FPC; $-36\ 50\ 6$; Figure 4D). Activation was also observed in the left posterior IFS/pre-PMd ($-36\ 14\ 28$), right DLPFC ($40\ 22\ 44$), bilateral posterior parietal cortex ($44\ -62\ 42, -40\ -56\ 42$), left inferior temporal cortex ($-66\ -38\ -12$), and left fusiform gyrus ($-34\ -70\ -10$).

Although activation in the FPC was highly reliable in the blocked analysis, the event-related analysis revealed the parametric effect in the FPC to be much weaker at the item level ($F = 1.2$). By contrast, the signal from the FPC averaged over the entire block revealed a reliable and robust parametric effect of context competition [$F(2, 36) = 6.2, p < .005$; Figure 4D].

The contexts that determined the relevant dimension, feature, or response were held constant across all blocks of the dimension, feature, and response experiments, respectively, thus context competition was minimal. Consistent with this hypothesis: (a) There were no differences in sustained FPC activation across competition conditions of the subordinate experiments ($F_s < 1.4$); (b) No competition condition of a subordinate experiment had greater sustained activation in the FPC than the C1 condition of the context experiment ($F_s < 3.6$); (c) All subordinate conditions had less activation than the C2 and C4 conditions of the context experiment ($F_s > 63.5, p_s < .0001$; Figure 4D). Notably, these results are not dependent on the sustained response. The event-related responses during the subordinate experiments were either equivalent to the C1 signal (dimension: $F = 1.8$) or reliably less than C1 (response and

feature: $F_s > 4.2, p_s < .05$), and all were reliably less than C2 and C4 blocks ($F_s > 6.1, p_s < .05$).

The sustained parametric pattern in the FPC was selective to the context experiment [Experiment \times Competition: $F(6, 108) = 2.7, p < .05$]. Furthermore, the sensitivity of the FPC to context competition dissociated this region from the PMd [Region \times Experiment \times Competition: $F(2, 36) = 8.8, p < .001$] and from the IFS [Region \times Experiment \times Competition: $F(2, 36) = 3.3, p < .05$]. The feature and context experiments also dissociated between the pre-PMd and FPC in overall activation [Region \times Experiment: $F(1, 18) = 11.2, p < .005$]. But, as the pre-PMd showed a trend for a parametric pattern during the context experiment, the three-way interaction with competition was not reliable [Region \times Experiment \times Competition: $F(2, 36) = 0.16$].

Primary Motor Cortex, Intraparietal Sulcus, and Fusiform Gyrus

In addition to the predicted activation observed along the posterior to anterior PFC gradient, the parametric analyses reported above also revealed activation in regions of primary motor, parietal, and inferior temporal cortices. ROI analysis further interrogated activation in these regions across experiments.

Activation from the parametric regressor of the response experiment identified activation in M1 ($-36\ -24\ 56$). This focus corresponds closely to the estimated location of M1 activity due to finger movement from both a meta-analysis ($-38.2\ -22.2\ 56.9$; Picard & Strick, 2001) and high resolution fMRI ($-36.6\ -23.4\ 56.8$; Moore et al., 2000). M1 activity likely reflects the general manual motor requirements of the tasks, and should be the ultimate termination of the motor hierarchy. In particular, M1 revealed a parametric increase in activation across conditions of the response experiment as more fingers were required for the task [$F(2, 36) = 36.0, p < .0001$; Figure 6A]. However, the PMd did not differ reliably from baseline on R1 blocks [$t(18) = -0.06, p = .95$] when there was only one response and so no response competition. By contrast, M1 was reliably active relative to baseline when there was no response competition [$t(18) = 4.2, p < .001$; Figure 6A]. This difference between the PMd and M1 is likely because making a single response regardless of the cue type does not entail any response competition, and so does not elicit activation in the PMd. However, it does require the execution of a voluntary movement, and so M1 is active under these conditions. This difference between M1 and PMd under R1 conditions dissociated these regions [Region \times Competition: $F(2, 36) = 3.9, p < .05$].

Activation along the left IPS was observed in the parametric competition contrast in the response ($-28\ -50\ 42$), feature ($-28\ -68\ 26$), and dimension ($-24\ -60\ 32$) experiments, and more laterally in both hemispheres in the context experiment ($44\ -62\ 42, -40\ -56$

42; Figure 6B). This suggests a general sensitivity to the increase in selection demands without being selective to competition at a particular representational level. Indeed, an ROI in the IPS defined from the response experiment confirmed reliable parametric effects (response and feature experiments) or differences between high competition conditions and low competition conditions (dimension and context experiments) for all four experiments ($F_s > 3.8$, $p_s < .05$), confirming a general sensitivity to increases in selectivity demands across experiments (Figure 5B). Although the significance of this pattern is unclear from the present data, one intriguing possibility is that the parietal cortex serves to temporarily store or represent the relevant contextual mappings, regardless of their level of representation. This interpretation is roughly consistent with a range of perspectives on the contribution of the parietal cortex to maintenance and updating of action rules or stimulus/problem representations (Cavina-Pratesi et al., 2006; Sohn, Goode, Stenger, Carter, & Anderson, 2003; Bunge, Hazeltine, Scanlon, Rosen, & Gabrieli, 2002).

Finally, fusiform gyrus activation was observed in the dimension (44 -56 -10) and context (-34 -70 -10) experiments. An ROI defined from the focus identified in the dimension experiment revealed reliable effects of high versus low competition conditions in the context ($C4 > C2 > C1$) and dimension ($D4$ and $D2 > D1$) experiments ($F_s > 4.5$, $p_s < .05$). There was no reliable effect of competition in the feature or response experiments ($F_s < 2.0$), although both showed activation above baseline. Whether the changes in activation in the fusiform gyrus reflect the impact of top-down control at higher hierar-

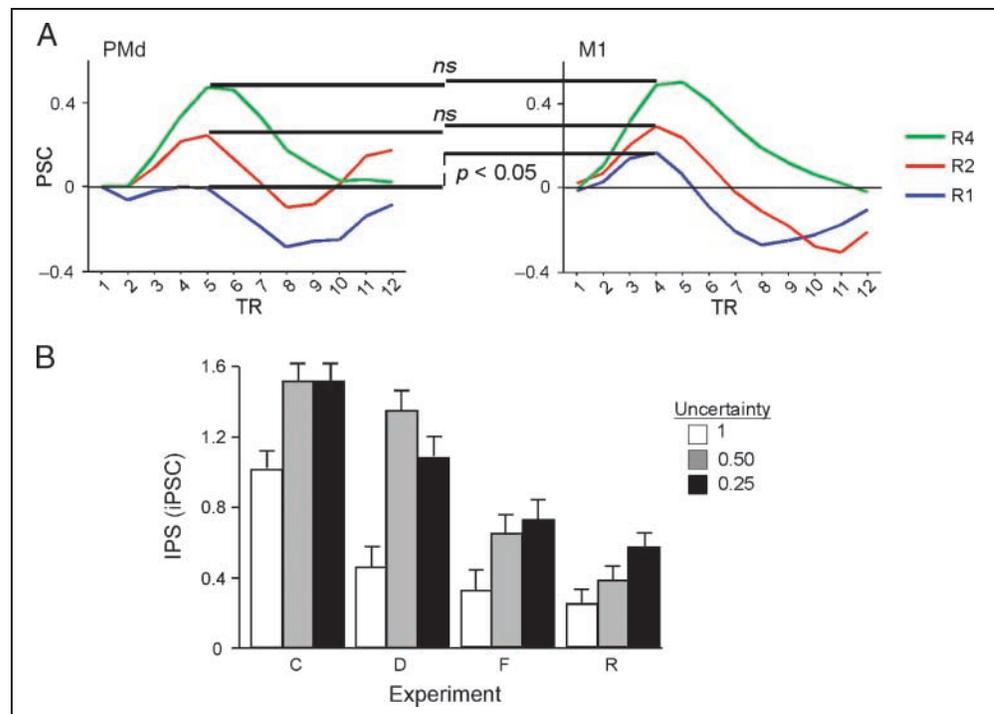
chical levels in the face of constant perceptual demands remains to be tested in future work.

DISCUSSION

These results provide strong empirical support for the hypothesis that cognitive control is organized in a representational hierarchy along the rostro-caudal axis of the frontal lobes. Furthermore, these results suggest that levels of the representational hierarchy, and so subregions of the PFC, may be differentiated by the level of abstraction at which the representations that guide action must be selected over competition. Abstraction is defined based on a classing rule whereby more abstract representations generalize across a class or set of representations at subordinate levels.

The design of this experiment varied control demands along two axes of difficulty, (1) the abstractness of representations to be selected determined the locus of activation along the anterior-to-posterior gradient of the frontal lobes and (2) the difficulty of selection via competition determined the level of activation at any given locus. Varying competition parametrically permitted the selective identification of regions in the PFC involved in control at each hierarchical level. Specifically, from posterior to anterior, the PMd resolved competition about responses, the pre-PMd resolved competition about sets of feature-to-response mappings, the IFS resolved competition between relevant dimensions that classed feature sets, and the FPC resolved competition about sets of contextual cue-to-dimension mappings (see Figure 5).

Figure 5. ROI analyses of M1 and IPS. (A) Although sensitive to an increase in response competition, event-related signal change in M1 (-36, -24, 56; right plot) was distinguished from the PMd in that it was also active, above baseline, during R1 blocks, when there was a motor response but no response-level competition. (B) Integrated percent signal change (iPSC) from an ROI in the IPS (-28 -50 42) demonstrated reliable sensitivity to increases in competition at all hierarchical levels, although this change was quantitatively more parametric for the response and feature experiments.



Critically, only these four regions along the rostro-caudal axis of the frontal lobes demonstrated all three of the following properties, consistent with a representational hierarchy: (1) a parametric increase with competition that started from baseline under minimal competition conditions and that was independent of an increase in the proportion of switches, (2) baseline activation levels for experiments in which competition was minimal, (3) stable and greater than baseline, but nonparametric, activation in posterior regions when activation in anterior regions varied parametrically. We discuss these points in more detail below.

It is important to consider the location of these regions in some detail, as anatomical precision is critical for comparison of these results to others in the literature. In particular, the PMd region located in the present experiment is very close to the definition of Picard and Strick (2001) based on a meta analysis (Picard & Strick, 2001: $x\ y\ z = \pm -37\ -14\ 60$; Current study $x\ y\ z = -30\ -10\ 68$). Similarly, what is termed the pre-PMd in the present study ($-38\ 10\ 34$) appears close to the anterior portion of the PMd that Picard and Strick (2001) term as the pre-PMd. Although we do note that this present focus may be at the more ventral and anterior extent of what Picard and Strick (2001) refer to as the pre-PMd, and indeed, also may be close to the dorsal and anterior extent of what has been termed the inferior frontal junction (IFJ; Brass, Derrfuss, Forstmann, & von Cramon, 2005). However, the pre-PMd in the present study is on or dorsal to the upper bound (only a 5–25% likelihood) of the pars opercularis subdivision of the inferior frontal gyrus (BA 44) based on a probability map computed from structural MRI in 108 individuals (Tomaiuolo et al., 1999). Hence, we view it as highly unlikely that the pre-PMd corresponds to BA 44. Similarly, the IFS activation is mostly dorsal to the pars triangularis portion of the inferior frontal gyrus, centering in the IFS and extending into middle frontal gyrus (~BA 46). Finally, activation in the FPC (~BA 10) appears highly consistent with the definition of lateral frontal pole used across a range of neuroimaging experiments (cf., Gilbert et al., 2006). Also consistent with prior findings (Braver et al., 2003; Donaldson, Petersen, Ollinger, & Buckner, 2001), the FPC was marked by a sustained signal that distinguished it from the subordinate, posterior regions. However, in the present experiment, as context mappings themselves were updated at the block level, it is unclear whether these data reflect something fundamental about the temporal dynamics of the FPC or whether selecting context level representations on a trial-by-trial basis would have resulted in an event-related FPC response.

The nested factorial design and parametric manipulation of conflict at each level rule out a number of alternative explanations of these results. First, any superficial differences between the individual experiments—such as the presence of a match/nonmatch decision during the dimension experiment or perceptual differences in

the type of object feature used—cannot account for the posterior to anterior progression of activation. This is because the association of various regions of the PFC with competition at particular levels of representation was not derived from simple comparisons between the experiments, but rather from within-experiment parametric increases in activation. Importantly, in each experiment, the parametric function increased from a baseline condition (R1, F1, D1, or C1) that included no competition at the parametrically varying level but was otherwise identical in all other superficial experimental details to the other mid- and high-competition conditions of that experiment. Thus, parametric differences between these mid- and high-competition conditions are only due to changes in the parametrically varying factor (i.e., competition).

Second, the nested parametric design also means that the anterior-to-posterior progression of activation cannot be explained in terms of simple task difficulty. As already noted, difficulty in this design varies along two axes: the abstract level of representation to be selected and the degree of competition among representations at a given level of abstraction. These axes were separately indexed by the locus and the level of activation, respectively. For example, from R2, the task could become harder in one of two ways: there were either more responses to choose from (R4), or there was more than one feature-to-response set to choose from (F2). The former was associated with an increase in the PMd but no reliable increase in the pre-PMd, and the latter with an increase in the pre-PMd but no reliable increase in the PMd. This logic was reproduced at each level of the hierarchy.

Broadly, the present data and explicit hierarchical framework support and build on past studies demonstrating a PFC gradient and proposing a representational hierarchy (Koechlin & Jubault, 2006; Koechlin et al., 2003; O'Reilly et al., 2002). The distinctions between hierarchical levels did not derive from differences in the nature of the control mechanisms, *per se*. Rather, the parametric manipulation that was tested at all levels of representation provided strong evidence that control demands are computationally similar across hierarchical levels and covary with competition, which we conceptualize in the same way as entropy (Koechlin et al., 2003) or representational conflict (Botvinick, Braver, Barch, Carter, & Cohen, 2001). Furthermore, the present study controlled for differences in the proportion of task switches across competition conditions and determined that hierarchical effects were evident even during repeat trials. At the lowest level, the present study replicates the location of the “sensory control” effect in the PMd reported by Koechlin et al. (2003) and verifies its operation over a larger parameter range than was previously demonstrated. Of course, in the present framework, we interpret the locus of this activation as deriving from a parametric increase in competition among distinct response representations, as opposed to the sensory

control signal or the number of sensory-to-response mappings, which do not necessarily change over parametric levels. Based on this finding alone, however, models emphasizing the control signal versus the representation to be selected appear indistinguishable.

There are, however, aspects of the present findings that appear to diverge in important ways from a representational hierarchy that is ranked based on the control signal, such as the cascade model. First, from cascade, the form of the control signal (sensory, context, and episodic control) should predict what regions of the PFC are engaged in resolving competition. For example, when episodic information is used for control, Koechlin et al. (2003) located activation in the anterior lateral PFC (Figure 6F). The context experiment of the present study was derived directly from the logic of episodic control in Koechlin et al., and so uses an episodic control signal. Consequently, the context experiment should also result in parametric activation in the anterior lateral PFC. However, parametric activation in the context experiment was in the FPC and clearly anterior to the Koechlin et al. episodic control region (Figure 6G vs. Figure 6F). On the other hand, the more posterior “episodic control” region is highly convergent with the IFS region associated with dimension competition in the present experiment (Figure 6F vs. Figure 6E). The dimension experiment contained no “episodic control.” Hence, to summarize, when the same control signal (i.e., episodic) was used

across two studies, different regions of the PFC were active; and, when different control signals were used (episodic versus sensory/context), the same region of the PFC was active. Thus, it seems difficult to interpret these results in terms of the form of the control signal, at least not in any obvious sense.

Furthermore, looking at the position of the blue and red spheres depicted in Figure 6, it is clear that the three sites of activation from Koechlin et al. (2003) map to the lowest three levels of the present study, despite identical logic between the highest and lowest levels regarding the form of the control signal. In other words, the insertion of the feature task—which is only meaningfully different under the classing rule of the present abstract representational hierarchy—appears to have shifted the upper two levels forward.

One potential account for this systematic divergence between these studies is that the rank of the representation being selected determines the site of activation along the posterior to anterior axis, and in the experiment by Koechlin et al. (2003), responses could be grouped into task sets and task sets could be grouped into blocks (three levels). In the present study, responses were grouped by features that were grouped by dimensions which were grouped by blocks (four levels). Thus, the rank of the representation derived from the classing rule may account for the overlap of the bottom three levels, despite differences in control signal, and the addition of a fourth level in the FPC. Of course, there is an obvious issue that this account introduces. Taken to the extreme, one could indefinitely add more superordinate levels, but one could not continually add additional anterior regions in the PFC. Hence, there must be some constraint, either in capacity or processing, that restricts the levels of control defined through this classing rule structure. Specifying the constraints of such a system is thus an important direction of further research.

A second potential point of divergence between the present dataset and the predictions of the cascade model is that we do not find strong evidence that information accumulates at lower levels of the hierarchy directly proportional to control demands at higher levels. Consider the PMd (Figure 4A). This region is at the lowest level of the hierarchy and so should accumulate the most information over the course of the four experiments. However, when control demands increased parametrically at higher levels, such as during the context or dimension experiments, the PMd did not show a corresponding parametric increase. Similarly, the IFS did not show a parametric increase as context competition increased from C2 to C4 conditions. Of the three posterior regions, only the pre-PMd showed any sensitivity to parametric manipulations at higher levels, and even this quantitative pattern was not always reliable (context competition). Indeed, the robust parametric patterns that were evident on repeat trials, that rose

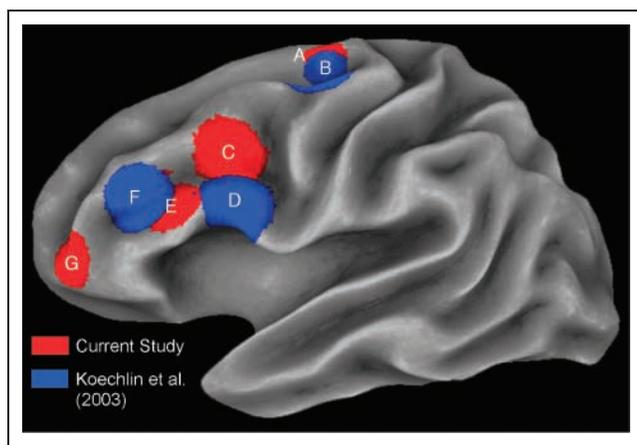


Figure 6. Comparison of peak activations from the present experiment and those obtained by Koechlin et al. (2003). Color patches do not represent extent of actual activations. Rather, in order to convey proximity of contiguous foci concretely, spheres (8 mm radius) were constructed centered on the peak foci from Koechlin et al. (blue) and from the current study (red). Specifically, spheres represent foci for (A) the response level (current experiment: Talairach $x, y, z = -30 -7 63$), (B) “sensory control” (Koechlin et al., 2003: $-32 -8 52$), (C) the feature level (current experiment: $-37 11 31$), (D) “context control” (Koechlin et al., 2003: $-44 8 20$), (E) the dimension level (current experiment: $-50 26 21$), (F) “episodic control” (Koechlin et al., 2003: $-40 32 20$), and (G) the context level (current experiment: $-36 50 6$). Spheres were then rendered on an inflated Talairach surface using AFNI.

from baseline, and that were unique to each highest level permitted selective identification of functionally specialized subregions. Thus, we do not find strong evidence for the direct accumulation of control signals at lower levels as higher levels varied over a large parametric range.

Although the activation in posterior regions did not consistently increase parametrically along with anterior regions, as might be predicted by cascade, activation in subordinate regions was not stable across superordinate conflict conditions either. First, baseline activity in posterior regions generally increased as more regions of the PFC became active. This is consistent with the idea that the rostro-caudal array of control processors forms a functional network rather than acting as fully autonomous units or stages; an idea consonant with the cascade model, as well. Second, although not always parametric, the PMd, pre-PMd, and IFS all showed at least a quantitative increase between baseline and mid- and high-level competition conditions of superordinate experiments. As noted above, this pattern may suggest that control is not entirely completed at higher levels before posterior regions are engaged in resolving conflict at lower levels. Moreover, direct or indirect interactions may exist between the levels of control. Such interactions could include the signaling of conflict at higher levels by lower levels and/or the direct modulation of lower levels by upper levels. Alternatively, even without direct interactions between levels, it may be that lower levels begin to engage in control before control at upper levels is complete. Hence, without the benefit of superordinate control, more alternative representations compete at lower levels, and so lower-level processors are placed under greater demands. Understanding the mechanisms of hierarchical control still awaits elucidation of the network dynamics between levels of representation within the PFC.

An alternate way of framing the present abstract representational hierarchy is in terms of the difficulty or complexity of the action rule (Bunge & Zelazo, 2006). Indeed, in the present experiment, as representations became more abstract, so did the complexity of the action rule (e.g., the number of contingencies traversed to determine a response), and this determined the locus of activation in the PFC. Consider that the response experiment tested a simple rule relating a cue and a response. The feature experiment tested the conjunction of a cue and an object to a response. The dimension experiment tested the conjunction of a cue with a feature conditional on a second feature to determine a response, and the context experiment tested the conjunction of a temporal context with the complex relation just expressed for the dimension experiment. Although this is a plausible progression, and follows from the classing rule, the need to maintain additional contingencies does not account for the parametric effects within each level. The absolute number of “rules” is equivalent, at four color mappings, across the mid- and high-competition condi-

tions. It is only the number of alternate competing representations that changes. A rule-complexity framework thus may describe the locus of control along the posterior to anterior axis, but it is competition among the representations that necessitates control and so determines the continuous level of activity in each region.

Although the focus of the present experiment sought to specify the organizing principle governing the rostro-caudal axis of the PFC, it is important to note the position of these activations with respect to their dorsal/ventral and lateral position. Peak activations across the four mini-experiments were mostly dorsal to the inferior frontal gyrus, either in the IFS or in the middle frontal gyrus. There is some evidence that the inferior frontal gyrus or the ventrolateral PFC (VLPFC) may comprise a separate functional gradient and hierarchy, with distinctions in abstraction and hierarchical level being found between the anterior and mid-VLPFC (e.g., Badre & Wagner, 2007; Gold et al., 2006; Badre, Poldrack, Pare-Blagoev, Insler, & Wagner, 2005) and between the anterior or mid- and posterior VLPFC/premotor cortex (e.g., Koechlin & Jubault, 2006; Poldrack et al., 1999). Although this dorsal/ventral distinction remains to be tested directly, the ventral gradient might be distinguished from the more dorsal gradient based on its involvement in retrieval and initiation of sequences of actions (Koechlin & Jubault, 2006) and the retrieval and organization of lexical, semantic, and phonological codes required for verbal behavior (Badre & Wagner, 2007). Activation in the present experiment was also left lateralized. However, this lateralization should not be interpreted too strongly. The PFC was activated bilaterally in a number of individual subjects, and prior studies using bimanual responses have reported bilateral activation (Koechlin et al., 2003). Hence, unilateral activation in the group maps may simply reflect differences in individual variability or the use of only a right-hand response in the currently study. Additional research will be required to prospectively test what principles govern these other axes of organization, as well as continuing to test hypotheses regarding the rostro-caudal organization of the PFC.

Acknowledgments

Supported by the National Institutes of Health (MH63901 and NS40813), Veterans Administration Research Service, and a National Research Service Award (F32 NS053337). We thank M. Paquette and F. Krienen for assistance with data collection.

Reprint requests should be sent to David Badre, Helen Wills Neuroscience Institute, 132 Barker Hall, MC3190, Berkeley, CA 94720-3190, or via e-mail: dbadre@berkeley.edu.

REFERENCES

- Anderson, J. R., Qin, Y., Stenger, V. A., & Carter, C. S. (2004). The relationship of three cortical regions to an information-processing model. *Journal of Cognitive Neuroscience*, *16*, 637–653.

- Badre, D., Poldrack, R. A., Pare-Blagoev, E. J., Insler, R. Z., & Wagner, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron*, *47*, 907–918.
- Badre, D., & Wagner, A. D. (2004). Selection, integration, and conflict monitoring: Assessing the nature and generality of prefrontal cognitive control mechanisms. *Neuron*, *41*, 473–487.
- Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia*, *45*, 2883–2901.
- Bernstein, N. (1967). *The coordination and regulation of movements*. London: Pergamon.
- Botvinick, M., & Plaut, D. C. (2004). Doing without schema hierarchies: A recurrent connectionist approach to normal and impaired routine sequential action. *Psychological Review*, *111*, 395–429.
- Botvinick, M. M. (2007). Multilevel structure in behaviour and in the brain: A model of Fuster's hierarchy. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*.
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*, 624–652.
- Brass, M., Derrfuss, J., Forstmann, B., & von Cramon, D. Y. (2005). The role of the inferior frontal junction area in cognitive control. *Trends in Cognitive Sciences*, *9*, 314–316.
- Braver, T. S., & Bongiolatti, S. R. (2002). The role of frontopolar cortex in subgoal processing during working memory. *Neuroimage*, *15*, 523–536.
- Braver, T. S., Reynolds, J. R., & Donaldson, D. I. (2003). Neural mechanisms of transient and sustained cognitive control during task switching. *Neuron*, *39*, 713–726.
- Brett, M., Anton, J.-L., Valabregue, R., & Poline, J.-B. (2002). *Region of interest analysis using an SPM toolbox*. In: 8th International Conference on Functional Mapping of the Human Brain, Sendai, Japan.
- Buckner, R. L. (2003). Functional–anatomic correlates of control processes in memory. *Journal of Neuroscience*, *23*, 3999–4004.
- Bunge, S. A. (2004). How we use rules to select actions: A review of evidence from cognitive neuroscience. *Cognitive, Affective & Behavioral Neuroscience*, *4*, 564–579.
- Bunge, S. A., Hazeltine, E., Scanlon, M. D., Rosen, A. C., & Gabrieli, J. D. (2002). Dissociable contributions of prefrontal and parietal cortices to response selection. *Neuroimage*, *17*, 1562–1571.
- Bunge, S. A., Kahn, I., Wallis, J. D., Miller, E. K., & Wagner, A. D. (2003). Neural circuits subserving the retrieval and maintenance of abstract rules. *Journal of Neurophysiology*, *90*, 3419–3428.
- Bunge, S. A., & Zelazo, P. D. (2006). A brain-based account of the development of rule use in childhood. *Current Directions in Psychological Science*, *15*, 118–121.
- Cavina-Pratesi, C., Valyear, K. F., Culham, J. C., Kohler, S., Obhi, S. S., Marzi, C. A., et al. (2006). Dissociating arbitrary stimulus–response mapping from movement planning during preparatory period: Evidence from event-related functional magnetic resonance imaging. *Journal of Neuroscience*, *26*, 2704–2713.
- Christoff, K., & Gabrieli, J. D. E. (2000). The frontopolar cortex and human cognition: Evidence for a rostrocaudal hierarchical organization within the human prefrontal cortex. *Psychobiology*, *28*, 168–186.
- Cooper, R. P., & Shallice, T. (2006). Hierarchical schemas and goals in the control of sequential behavior. *Psychological Review*, *113*, 887–916.
- Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage*, *9*, 179–194.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222.
- D'Esposito, M., Postle, B. R., & Rypma, B. (2000). Prefrontal cortical contributions to working memory: Evidence from event-related fMRI studies. *Experimental Brain Research*, *133*, 3–11.
- Donaldson, D. I., Petersen, S. E., Ollinger, J. M., & Buckner, R. L. (2001). Dissociating state and item components of recognition memory using fMRI. *Neuroimage*, *13*, 129–142.
- Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, *23*, 475–483.
- Estes, W. K. (1972). An associative basis for coding and organization in memory. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 161–190). Washington, DC: V. H. Winston & Sons.
- Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis: II. Inflation, flattening, and a surface-based coordinate system. *Neuroimage*, *9*, 195–207.
- Fuster, J. M. (1997). *The prefrontal cortex: Anatomy, physiology, and neuropsychology of the frontal lobe* (3rd ed.). Philadelphia, PA: Lippincott-Raven Publishers.
- Fuster, J. M. (2004). Upper processing stages of the perception–action cycle. *Trends in Cognitive Sciences*, *8*, 143–145.
- Gilbert, S. J., Spengler, S., Simons, J. S., Steele, J. D., Lawrie, S. M., Frith, C. D., et al. (2006). Functional specialization within rostral prefrontal cortex (area 10): A meta-analysis. *Journal of Cognitive Neuroscience*, *18*, 932–948.
- Gold, B. T., Balota, D. A., Jones, S. J., Powell, D. K., Smith, C. D., & Andersen, A. H. (2006). Dissociation of automatic and strategic lexical–semantics: Functional magnetic resonance imaging evidence for differing roles of multiple frontotemporal regions. *Journal of Neuroscience*, *26*, 6523–6532.
- Hazy, T. E., Frank, M. J., & O'Reilly, R. C. (2006). Banishing the homunculus: Making working memory work. *Neuroscience*, *139*, 105–118.
- Koechlin, E., & Jubault, T. (2006). Broca's area and the hierarchical organization of human behavior. *Neuron*, *50*, 963–974.
- Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science*, *302*, 1181–1185.
- Koechlin, E., & Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends in Cognitive Sciences*, *11*, 229–235.
- Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112–136). New York: Wiley.
- Logan, G. D., & Bundesen, C. (2004). Very clever homunculus: Compound stimulus strategies for the explicit task-cuing procedure. *Psychonomic Bulletin & Review*, *11*, 832–840.
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202.
- Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Holt, Rinehart and Winston.
- Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, *7*, 134–140.
- Moore, C. I., Stern, C. E., Corkin, S., Fischl, B., Gray, A. C., Rosen, B. R., et al. (2000). Segregation of somatosensory

- activation in the human rolandic cortex using fMRI. *Journal of Neurophysiology*, *84*, 558–569.
- Newell, A. (1990). *Unified theories of cognition*. Cambridge: Harvard University Press.
- O'Reilly, R. C., & Frank, M. J. (2006). Making working memory work: A computational model of learning in the prefrontal cortex and basal ganglia. *Neural Computation*, *18*, 283–328.
- O'Reilly, R. C., Noelle, D. C., Braver, T. S., & Cohen, J. D. (2002). Prefrontal cortex and dynamic categorization tasks: Representational organization and neuromodulatory control. *Cerebral Cortex*, *12*, 246–257.
- Passingham, R. E., & Rowe, J. B. (2002). Dorsal prefrontal cortex: Maintenance in memory or attentional selection? In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 221–232). Oxford: Oxford University Press.
- Petrides, M. (2005). Lateral prefrontal cortex: Architectonic and functional organization. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, *360*, 781–795.
- Petrides, M. (2006). The rostro-caudal axis of cognitive control processing within lateral frontal cortex. In S. Dehaene, J.-R. Duhamel, M. D. Hauser, & G. Rizzolatti (Eds.), *From monkey brain to human brain: A Fyssen Foundation Symposium* (pp. 293–314). Cambridge: MIT Press.
- Picard, N., & Strick, P. L. (2001). Imaging the premotor areas. *Current Opinion in Neurobiology*, *11*, 663–672.
- Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage*, *10*, 15–35.
- Sakai, K., & Passingham, R. E. (2003). Prefrontal interactions reflect future task operations. *Nature Neuroscience*, *6*, 75–81.
- Saltzman, E. L. (1979). Levels of sensorimotor representation. *Journal of Mathematical Psychology*, *20*, 91–163.
- Shimamura, A. P. (1995). Memory and frontal lobe function. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 803–813). Cambridge: MIT Press.
- Sohn, M. H., Goode, A., Stenger, V. A., Carter, C. S., & Anderson, J. R. (2003). Competition and representation during memory retrieval: Roles of the prefrontal cortex and the posterior parietal cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, *100*, 7412–7417.
- Stuss, D. T., & Benson, D. F. (1987). The frontal lobes and control of cognition and memory. In E. Perecman (Ed.), *The frontal lobes revisited* (pp. 141–158). New York: The IRBN Press.
- Tomaiuolo, F., MacDonald, J. D., Caramanos, Z., Posner, G., Chiavaras, M., Evans, A. C., et al. (1999). Morphology, morphometry and probability mapping of the pars opercularis of the inferior frontal gyrus: An in vivo MRI analysis. *European Journal of Neuroscience*, *11*, 3033–3046.