BINDING OF WHAT AND WHERE
DURING WORKING MEMORY MAINTENANCE

Joseph B. Sala[1,2] **and Susan M. Courtney**[3,4,5]

([1]Psychology Department, Stanford University, Stanford, CA, USA; [2]Department of Neurobiology, Stanford University, Stanford, CA, USA; [3]Department of Psychological and Brain Sciences, Johns Hopkins University, Baltimore, MD, USA; [4]Department of Neuroscience, Johns Hopkins University, Baltimore, MD, USA; [5]F.M. Kirby Research Center for Functional Brain Imaging, Kennedy Krieger Institute, Baltimore, MD, USA)

ABSTRACT

Prefrontal cortex (PFC) supports the maintenance of currently relevant information in working memory (WM). How the PFC is organized for the maintenance of disparate information, how this information is conjoined into a unified whole, and how the representation may change with task demands is still debated. The pattern of neural activity during maintenance of either abstract visual patterns, locations, or their "conjunction" was measured in two experiments using functional magnetic resonance imaging (fMRI). During delays, common regions in PFC were active, but a dorsal-ventral/spatial-nonspatial functional topography distinguished among the three delay types. During conjunction delays, no additional neural architecture was recruited. Instead, conjunction delays were characterized by a significant reduction compared to the response of that cortical region while maintaining its "preferred" information. A model is presented, extending the principles of "biased competition" to the PFC and the dynamic maintenance of information in WM, that accounts for current and seemingly contradictory previous results from both imaging and physiological studies. In this schema, the PFC is not only the source of biasing signals targeting earlier processing regions, but is also the target of these signals. This model stands as an alternative to traditional "domain specific" and "domain general" models of frontal organization of WM, and as an extension of earlier models of PFC mechanisms related to the cognitive control of goal directed behavior.

Key words: fMRI, spatial, nonspatial, visual, prefrontal

INTRODUCTION

Functional magnetic resonance imaging (fMRI) experiments in humans clearly demonstrate that both dorsal and ventral frontal cortex show activation during both object and spatial working memory (WM) tasks relative to non-mnemonic control tasks (Baker et al., 1996; Owen et al., 1996; D'Esposito et al., 1998; Nystrom et al., 2000; Postle et al., 2000a, 2000b; Stern et al., 2000; Postle and D'Esposito, 1999). In addition, single-cell recording experiments in monkeys demonstrate that cells in both dorsal and ventral frontal cortex can show selective responses to either objects, locations, or both depending on what information is currently task-relevant (Rao et al., 1997; Rainer et al., 1998). These findings are supported by evidence from other laboratories demonstrating the ability of single prefrontal cortex (PFC) neurons to integrate color and location information according to a pre-learned "rule" (White and Wise, 1999) or over the delays of delayed response tasks (Quintana et al., 1988; Yajeya et al., 1988). PFC neurons also appear to make cross-modal associations (Fuster et al., 2000). Furthermore, prefrontal cooling implicates the PFC specifically in the cross-temporal integration of spatial and nonspatial information (Quintana and Fuster, 1993).

The representation of information for maintenance within the PFC has been a matter of some controversy, however. Experimental results from both monkeys (for review see Levy and Goldman-Rakic, 2000) and humans (Courtney et al., 1996, 1998; Munk et al., 2002; Sala et al., 2003; Alain et al., 2001; Carlesimo et al., 2001; Bechara et al., 1998; Mottaghy et al., 2002; Rama et al., 2004) have indicated that posterior dorsal frontal cortex is preferentially involved during maintenance of spatial information and posterior ventral frontal cortex is preferentially involved during maintenance of nonspatial information. This has lead some (Goldman-Rakic, 1987, 1995a; Wilson et al., 1993) to propose a model of domain specificity, in which information is segregated in WM and the PFC according to the information content. This idea of domain specificity is similar to Baddeley's (Baddeley and Hitch, 1974) model in which separate neural systems are responsible for the phonological loop and the visuospatial sketchpad, although spatial and nonspatial WM tasks may share common executive resources.

Others have failed to observe any dorsal-ventral organization within the PFC related to maintenance of object *versus* spatial information (Petrides, 1995a, 1995b, 1996; D'Esposito et al., 1998; Nystrom et al., 2000; Stern et al., 2000; Postle and D'Esposito, 1999). Instead, it has been suggested that visual object and spatial information are maintained in a domain-general store (e.g., see Fuster, 2001 for review), Some have suggested that

instead there exists a dorsal-ventral segregation based on the WM process engaged (Petrides, 1995b; Owen et al., 1996; D'Esposito et al., 1999). However, segregation within the PFC based upon process is largely orthogonal to questions relating to whether the information is stored in a domain-specific or domain-general fashion as has been discussed by Johnson et al. (2003). Indeed some "process" models have acknowledged the possible existence of a spatially specialized area (Petrides, 2000).

Largely, the results on PFC response during the maintenance of object and spatial information exclusively support neither a strict domain-specific segregation nor an equipotent domain-general organization. Questions remain, then, as to how information is represented across dorsal and ventral lateral frontal cortex for WM maintenance, how this representation may change to reflect task demands, and what kind of cellular network could account for these seemingly contradictory results.

Previous studies that have examined cortical activity during maintenance of both an object's identity and its location have provided some insight into these questions. Rainer et al. (1998) showed that when both object and location are task relevant the majority of cells respond best when the cell's preferred object *and* preferred location match the sample, indicating an integrated representation. It has been suggested that such a representation could arise from either intrinsic connections within the frontal cortex or those that develop from learned associations (Rao et al., 1997; White and Wise, 1999). Such "adaptive coding" does not necessarily imply equipotentiality (Duncan, 2001), but dual spatial-object selectivity in single PFC cells has been frequently cited as supporting, and is consistent with, domain-general models of PFC organization (e.g., Nystrom et al., 2000; Stern et al., 2000; Postle and D'Esposito, 1999). An intrinsic integrated representation via omnipotent cell types equally distributed across dorsal and ventral frontal cortex, however, would not be in agreement with other data from lesions (Carlesimo et al., 2001; Bechara et al., 1998; Mottaghy et al., 2002) , single cell recording (Wilson et al., 1993; O'Scalaidhe et al., 1999), and fMRI (Courtney et al., 1996, 1998; Munk et al., 2002; Sala et al., 2003; Sakai and Passingham, 2003; Rama et al., 2004) showing a dorsal/ventral spatial/nonspatial functional topography. Two fMRI studies have suggested that activation of additional frontal cortical regions might mediate the interactions required to create such an integrated representation (Prabhakaran et al., 2000; Munk et al., 2002). However, the regions showing such additional activation were not the same across the two studies.

The current research sought to reconcile these disparate findings within a single model of the organization of PFC for maintenance in WM, by using fMRI to test the predictions of three types of possible organizational principles of PFC during WM for objects (the WHAT task), locations (the WHERE task), and objects-in-locations (the AND or BOTH tasks). We investigated whether the regions active during WM maintenance for any of these tasks showed varying degrees of engagement according to the stimulus representation necessary to fulfill the task demands, which would indicate a domain-dependent functional topography within this network. Of specific interest was the leverage that the *conjoined* representation of objects-in-locations could provide in discriminating between models of PFC organization. Comparison of the pattern of activity for these conjoined tasks *versus* that for the single information tasks could help determine whether the organization of the PFC during maintenance is stable and information-domain sensitive, stable and domain insensitive, or dynamically dependent on task demands.

These three "organizational principles" each suggest a different pattern across the frontal cortex for the three tasks and each garners support from various previous research efforts. A simple model positing stable and information-domain sensitive response across dorsal and ventral frontal cortex would predict that dorsal frontal cortex would respond preferentially during WHERE tasks and ventral frontal cortex would respond preferentially during WHAT tasks. If instead the organization of the frontal cortex is domain insensitive, then one would predict that dorsal and ventral frontal cortex would respond equally for WHAT and WHERE trials. Both the stable domain-sensitive and the stable domain-insensitive models would predict that during simultaneous maintenance of objects *and* their locations dorsal and ventral frontal cortices would respond either to the same or a greater degree as they did for each single information task. This latter prediction follows from the domain-sensitive model because in the conjunction task both cell populations (those representing objects and those representing locations) would need to be active. This same prediction follows from the stable domain-insensitive model because the same cell populations are presumed to represent either objects or locations or both. In a domain-sensitive model the relationship between an object and its location could be considered to be an additional type of information and thus one might predict an addition frontal region would come online during object-in-location WM to represent the *binding* of the two information types.

The alternative organizational principle for the neural systems responsible for WM is a dynamic system in which the nature of the representation of object and spatial information changes, being either segregated or integrated, according to task demands. Dynamic changes in information representation have already been observed in

studies of attentional selection of perceptual information. The highly influential model of "biased competition" (Desimone and Duncan, 1995) states that domain-dependent neural representations of perceptual information are mutually inhibitory and that biasing signals (presumably from prefrontal cortex) influence this competition in order to select those aspects of the sensory input that will be further processed. If such a mechanism were also at work to dynamically select the information representation to be maintained in WM, such that within-domain competition among objects or locations is biased by task-dependent influences from the other (spatial or nonspatial) information domain, the prediction for activation during the object-in-location WM task could be rather different. As explained in the discussion, such as system could create an integrated representation of objects-in-locations that would result in an overall lower level of activation in both dorsal and ventral frontal regions than would be observed during maintenance of only objects or only locations.

Behaviorally, conceptual models of WM have also debated the integration of visual information in WM. While efforts have been made to demonstrate the separable nature of visual spatial and nonspatial information (Logie, 1995; Reisberg and Logie, 1993), evidence suggests that when multiple features are task relevant, they are not stored separately, but rather are represented as integrated objects (Vogel et al., 2001; Jiang et al., 2000), bound by their common location. However, it appears as if this integration may be an additional and possibly an attentionally demanding step (Wheeler and Treisman, 2002). If objects and locations are stored separately or if their integration requires additional processing, one might predict that remembering objects-in-locations would be more demanding than remembering either the objects or their locations alone, leading to poorer performance on such tasks. A dynamic WM representation created through biased competition, however, suggests that both separate and bound representations would result in a similar WM load and that performance on the three tasks should be equivalent. Such a system would also agree with models, based mainly on behavioral data, positing a domain-general resource demanded by all three types of WM tasks that functions to select and create the most appropriate representation of task-relevant information to be maintained (Salway and Logie, 1995; Anderson et al., 2004).

Here, we report results from two fMRI experiments that examined the activation patterns during maintenance of an abstract image, a location, or the combination of the two. Experiment 1 consisted of three delayed-recognition tasks presented in a pseudo-random order (Figure 1A). During the WHAT task, subjects were instructed to remember the identity of two sample images,
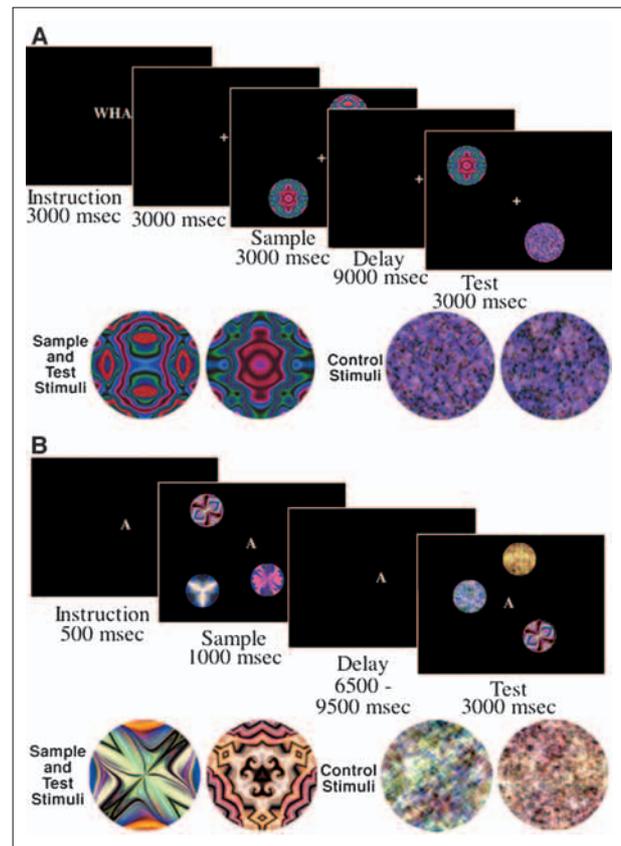

Fig. 1 – *Example stimuli and task schematic for A) Experiment 1 and B) Experiment 2.*

regardless of their position on the screen, and at test indicated whether a test image matched either of the samples. During the WHERE task, subjects needed to remember the locations of two sample images regardless of image identity. For the BOTH task, subjects were again presented with two sample images, but were now instructed that they needed to remember each of the patterns in the location in which it originally appeared. At test, during a BOTH trial, a matching test image consisted of one of the sample images in the location in which it originally appeared. A nonmatching test image for the BOTH task consisted of one of the sample images in a new location, a new image in one of the sample locations, or one of the sample images in the other sample image's location. For all tasks, it was explained to subjects that at test, a "scrambled" control image would be presented along with the "test image" and that this image was in no way related to the task.

Experiment 2 (Figure 1B), while similar to Experiment 1, contained several minor and one significant change. The major difference between the two tasks consisted of a change in the task requiring maintenance of a conjoined representation. In Experiment 2 the BOTH task was replaced with the AND task. The key distinction between these tasks was in the possible test images that could appear. For the AND task in Experiment 2, the test image was always either one

of the sample images in its original location (match) or one of the sample images in a location belonging to a different sample image. In this sense, the AND task was a specific test of the "bind" between an image and location. Minor differences included using three sample memorandum rather than two, a variable delay rather than a fixed delay length, scanning at 3T rather than 1.5T, inclusion of a sensorimotor control task, an instruction cue present throughout the trial, and blocking the three task conditions.

The results of these two experiments support the idea of a dynamic representation of information within WM. A model is presented in which competitive interactions within domain-sensitive dorsal and ventral regions of PFC are biased by task-dependent interactions between these regions. This Biased-Competition in Working Memory (BCWM) model provides a possible neural mechanism that would explain both the current experimental results and those of previous studies.

MATERIALS AND METHODS

Participants

Subjects (Experiment 1: n = 8, ages 18-24, 4 female; Experiment 2: n = 10, ages 18-26, 6 female) were non-smokers in good health with no reported history of head injury, neurological or mental disorders, or substance abuse, not currently using medications that affect the central nervous system or cardiovascular function. The experimental protocol was approved by the institutional review boards of both the Johns Hopkins University and the Johns Hopkins Medical Institutions. Participants were compensated with US$ 50. All the subjects gave written informed consent and were provided with written and oral debriefing after completion of the study.

Design

Stimuli

To minimize the verbalizability of sample images, colorful abstract patterns were created as stimuli using commercial software, Artmatic 2.1 (U&I Software). Control stimuli were these same images phase scrambled, maintaining the same overall wavelength, luminance, and frequency information as the original (Figure 1). In Experiment 1, a fixed set of 12 images was used throughout the experiment for all subjects. For Experiment 2 a random sample of 12 images were drawn from a set of 100 images for each subject. All images measured approximately 3.5° of visual angle and were initially unfamiliar to subjects and appeared equally often in all experimental conditions.

Experiment 1 Tasks

There were three delayed-recognition tasks: a nonspatial WM task (WHAT), a spatial WM task (WHERE), and a WM task for the conjunction of nonspatial and spatial information (BOTH). Before the experiment, subjects were informed that the instruction WHAT meant they needed to remember what the images looked like regardless of where on the screen they appeared, WHERE meant they needed to remember where the images were presented regardless of what images were presented, and BOTH meant they needed to remember the visual image in its exact location on the screen (the conjunction of WHAT and WHERE information). Each trial consisted of 1) the presentation of a visual instruction (WHAT, WHERE, or BOTH) for 3 sec, 2) a 3 sec period of fixation, 3) the simultaneous presentation of 2 abstract images as samples for 3 sec, 4) a 9 sec delay period, and 5) the simultaneous presentation of 1 abstract test image and 1 scrambled control image for 3 sec (Figure 1). During the presentation of the test image the subject would indicate with a button press whether the abstract image presented matched either of the sample images according to the cued information (left button = match, right button = mismatch). Subjects performed six trials of each task type in a randomized order within each of 6 imaging runs. Between trials a random number of "null" fixation events were inserted, resulting in varying intertrial interval (ITI) lengths of 3-21 sec. Before participation in fMRI, each subject received two practice runs using the same set of stimuli.

For all tasks, sample images were presented approximately 6° of visual angle from fixation and at least 60 radial degrees apart from one another so that they would neither overlap in presentation, nor share a spatial non-match distractor location. During presentation of all test images, a scrambled control stimulus was simultaneously presented in a location that was neither in nor within 20 radial degrees of a sample location, and was at least 60 degrees from the test image. This control stimulus presented at test allowed the sample and test presentation periods to contain similar amounts of visual stimulation. At test, the WHERE non-match distractor was a non-sample image presented 20 radial degrees from one of the sample locations. The WHERE match was a non-sample image presented in one of the sample locations. The WHAT non-match distractor was a non-sample image pattern, presented in a location that was neither in, nor within 20 radial degrees of, one of the sample locations. A WHAT match was a sample image pattern, presented in a location that was neither in, nor within 20 radial degrees of, one of the sample locations. The BOTH task non-match distractor was either a sample image pattern presented 20 degrees from the location in which it appeared at sample, a non-sample image presented in one of the sample locations, or a sample image

pattern presented in the location of the other sample image pattern. A BOTH match was a sample image in the sample location in which it was originally presented.

Experiment 2 Tasks

There were three delayed-recognition tasks: a nonspatial WM task (the WHAT task), a spatial WM task (the WHERE task), and a WM task for the conjunction of nonspatial and spatial information (the AND task). Before the experiment subjects were informed that the instruction "I" meant that they needed to remember what the images looked like regardless of where on the screen they appeared, "L" meant that they needed to remember the locations where the images were presented regardless of what images were presented, and "A" meant that they needed to remember the visual image in its exact location on the screen (the conjunction of the WHAT and WHERE information). Sensorimotor control (NOTHING) trials were also included. Subjects were instructed that the instruction "N" meant that they simply needed to watch the presentation of 3 scrambled images, wait over a delay, and press both buttons when another set of 3 scrambled images was presented. For all tasks, stimuli were presented on one of two imaginary circles centered about a fixation cross that remained visible throughout the trial. Stimuli appeared in 18 locations, six on an inner and 12 on an outer circle centered approximately 3.5° and 6° of visual angle from fixation respectively.

Each trial consisted of: 1) the presentation of a visual instruction ('I', 'L', 'A' or 'N') for 500 msec which then remained on the screen throughout the trial to serve as fixation, 2) the simultaneous presentation of three abstract images as samples for 1 sec, 3) a delay period of 6.5, 8, or 9.5 sec, and 4) the simultaneous presentation of one abstract test image and two scrambled control images for 3 sec (Figure 1). During the presentation of this test image the subject would indicate with a button press whether the abstract image presented matched either of the sample image presentations according to the cued information. Unlike in Experiment 1, trials were not randomized, but rather grouped into distinct mini blocks to discourage subjects from adopting one strategy to perform on all trials regardless of task demands. Subjects performed two mini blocks of each of the experimental tasks within each run, with the order of the tasks counterbalanced within the run and across subjects. Each mini-block contained three trials, for a total of six trials of each type in each run. The first and last trial of each run was a NOTHING trial, and these control trials separated each of the mini-blocks as well, for a total of seven control trials per run. Between each trial a random number of "null" fixation events were inserted, resulting in varying ITI lengths of 4, 5.5, or 7 sec. Each run began and ended with 12 sec of fixation. A run was 7 min and 54 sec long and subjects underwent six imaging runs. Before participation in fMRI, each subject received two practice runs using the same set of stimuli.

For all tasks, sample images consisted of three of the twelve abstract patterns presented in three of the eighteen locations. During presentation of all test images, two scrambled control stimuli were simultaneously presented in two of the remaining 15 locations that did not match any of the locations used at sample. A non-match for the "WHERE" task was the presentation of a non-sample image in a location not used, but immediately adjacent to a location used during sample presentation. The "WHERE" match was a non-sample image presented in one of the sample locations. The "WHAT" non-match distractor was a non-sample image pattern, presented in a location that was neither in, nor immediately adjacent to one of the sample locations. A "WHAT" match was a sample image pattern, presented in a location that was neither in, nor immediately adjacent to one of the sample locations. The "AND" task non-match distractor was one of the sample image patterns presented in one of the sample locations, but not the location in which that image appeared at sample. An "AND" match was a sample image in the sample location in which it was originally presented.

Imaging Protocol

All scans were performed at the F.M. Kirby Research Center for Functional Brain Imaging. Experiment 1 utilized a 1.5 Tesla Philips Gyroscan ACS NT Powertrak 6000, while Experiment 2 was performed on a 3 Tesla Philips Intera Quasar. A T1-weighted high-resolution anatomical image was obtained halfway through the acquisition of functional data to allow for anatomical localization (Experiment 1: MP-RAGE, 70 coronal slices, $2.5 \times .9 \times .9$ mm, TR = 20 msec, TE = 4.60 msec, FA = 30°; Experiment 2: 200 coronal slices, $1.0 \times 1.0 \times 1.0$ mm, TR = 8.13 msec, TE = 3.7 msec, FA = 8°). During the performance of the cognitive tasks, a T2* weighted, interleaved gradient echo, echo planar imaging sequence [Experiment 1: 35 coronal slices, $5 \times 3.59 \times 3.59$ mm (FOV = 240, 64×64 acquisition and reconstruction matrix), TR = 3000 msec, TE = 40 msec, FA = 90°; Experiment 2: 27 axial slices, $3.0 \times 1.875 \times 1.875$ mm (FOV = 240, 80×80 acquisition matrix, 128×128 reconstruction matrix), 1 mm gap, TR = 1500 msec, TE = 30 msec, FA = 65°] was used to obtain functional images.

fMRI Analysis: Experiment 1

General

Individual subject data was slice-time corrected, motion corrected, and analyzed via multiple

regression using analysis of functional neuroimaging (AFNI) software (Cox, 1996). The model for multiple regression consisted of twelve experimental regressors and six regressors derived from motion correction parameters to help account for residual variance due to uncorrected motion. A time series representing the occurrence of each of the predefined cognitive events (instruction, sample presentation, delay, and test presentation) for each of the tasks was separately convolved with a model hemodynamic response function to create the experimental regressors. This resulted in a matrix with each column indicating the experimental condition (i.e., 'WHAT – instruction', 'WHAT – sample', 'WHAT – delay', 'WHAT – test', 'WHERE – instruction', 'WHERE – sample', 'WHERE – delay', 'WHERE – test', 'BOTH – instruction', 'BOTH – sample', 'BOTH – delay', 'BOTH – test'), and each row indicating an individual scan volume. The value at each cell indicated the predicted response at a given scan due to the occurrence of an experimental condition eliciting a model hemodynamic response. These regressors combined correct and incorrect trials due to the limited number of trials of an individual task type that we were able to scan. Scalar beta coefficients for each of these regressors were obtained via multiple regression and were converted into percent signal change from the average baseline coefficient (comprised of unmodeled timepoints, i.e. ITI) for each of the runs.

Following model estimation, beta coefficients were converted into percent signal change and transformed into Talairach space in AFNI for group averaging. The percent signal change within each voxel for each of the delay types was entered into an analysis of variance (ANOVA) with subjects as a random factor. Those areas showing significant WHAT, WHERE, or BOTH delay activity relative to baseline (ITI) across subjects, were used to create a mask for between-task comparisons of differences. Percent signal change indicating significant increases in blood oxygen level-dependent (BOLD) signal were analyzed separately from those representing significant decreases in BOLD signal. For all comparisons, individual voxel significance was held to $p < .01$ and corrected for multiple comparisons via spatial extent of activation. This spatial extent correction for multiple comparisons was based on Monte Carlo simulations estimating the size of a "cluster" of active voxels necessary to satisfy an experiment wise $p < .05$ (Ward, 2000) (inclusion in masks of significant activity for delays compared to baseline required 11 suprathreshold contiguous voxels, 704 mm^3; comparisons within these masks required 4 voxels, 256 mm^3).

Functional topography maps were created by independently assigning the R – G – B channels of a voxel's color according to the relative strength of response during each of the WHAT – WHERE – BOTH delay periods respectively. The relative strength of response for each of the delay periods was calculated by dividing its beta weight by the sum of the three delay beta weights. This fraction was then converted to a scale from 0 to 255, and this number was set as the value of the color channel. If a voxel was equally responsive to all memory delays, then the relative strength of response for the WHAT, WHERE, and BOTH delays would each be .33, translating into a value of 84 for the Red, Green, and Blue color channels (resulting in a gray color). However, should the response from within a voxel be greater for one delay over the others, the resulting color blend will be closer to that of the dominant channel. For example, in the extreme case that a voxel is only responsive during WHAT delays, the relative strengths for each delay would be: WHAT = 1.0 (255 on red channel), WHERE = 0 (0 on green channel), and BOTH = 0 (0 on blue channel). The resulting color would therefore be pure red.

The estimated "time-course" of activity for each trial type was produced by performing deconvolution analysis on the individual subjects' data (Ward, 2002). This analysis produces an estimate of the magnitude of the response function for each of eleven time-points (33 sec) following the beginning of the presentation of the instruction across trials, for each of the three tasks. This estimate of response at each time point was then converted into percent signal change from the average baseline estimate of all the runs. These percent signal change values were then averaged across all voxels within a region for each subject, and then averaged across subjects to create the trial-averaged time-courses. The error bars reflect the standard error of the interaction between subject and delay type at each time point. These "timecourses" are presented to demonstrate the signal across each of the tasks. None of the statistics cited in the results or discussion were based on the estimates from this deconvolution analysis and this analysis was not used to separately model delay period activity.

ROI Analysis

The regions selected for region of interest (ROI) analyses were defined either functionally, based upon areas showing a significant main effect of experimental delay type (WHAT, WHERE, BOTH/AND), or anatomically based upon functional areas reported in previously published material (Prabhakaran et al., 2000; Munk et al., 2002). The percent signal change during the delay was averaged across all the voxels within a region for each subject. These values were then entered into random effects tests for significance (either paired t-tests, or repeated measures ANOVA).

Laterality

To test for presence of within-task laterality of response in anterior and posterior cortex, all voxels showing significantly greater response during delay than baseline were grouped into four regions based upon position left or right of midline, and anterior or posterior of post-central gyrus. The percent signal change for these regions was obtained as in the other ROI analyses, and the values for left and right hemispheres were entered into paired t-tests for each of these regions.

fMRI Analysis: Experiment 2

General

Individual subject data was analyzed similarly to Experiment 1. Experimental regressors were created for each of the predefined cognitive events (instruction and sample presentation, delay, and test presentation) for each of the tasks, including the sensorimotor control, separately.

The percent signal change for each of the delays was entered into an ANOVA with subjects as a random factor. Those areas showing significant "WHAT", "WHERE", or "AND" delay activity relative to "NOTHING" delay activity across subjects, were used to create a mask for between task comparisons. For all comparisons, individual voxel significance was held to p < .01 and corrected for multiple comparisons via spatial extent of activation, holding each cluster of voxels to an experiment wise p < .05 (inclusion in masks of significant activity for WM delays compared to control delays required 92 suprathreshold contiguous voxels, 1294 mm³; comparisons within these masks required 51 voxels, 717 mm³). Functional topography maps were created using the method presented for Experiment 1.

Hemodynamic response functions time-locked to the beginning of the instruction and sample presentation, were estimated for each of the tasks separately by performing deconvolution analysis on the individual subjects' data (Ward, 2002). An estimate of the response for nine time-points (13.5 sec) following the presentation of the instruction in each trial was obtained for each of the four tasks. These estimated hemodynamic response functions were averaged similarly as in Experiment 1. However, the estimated response during each corresponding time point of the control task was subtracted from that of each of the WM tasks at each time point. These resulting time-courses represent the *difference* in response between each of the tasks and sensorimotor control at each time point, starting at the beginning of the instruction presentation. As for Experiment 1, these "timecourses" are presented to demonstrate the activation across the trial during each of the tasks. Again none of the statistics cited in the results or

discussion were based on the estimates from this deconvolution analysis and this analysis was not used to separately model delay period activity.

ROI Analysis

Regions of interest were defined similarly to those for Experiment 1, both functionally and anatomically.

RESULTS

Behavioral

Prior to the collection of fMRI data, the WM tasks were behaviorally tested on separate groups of subjects (Experiment1: n = 30, Experiment 2: n = 15) to ensure that, within experiments, the tasks were equated for overall difficulty. A one-way ANOVA on accuracy measures (Experiment 1: WHAT = 79%, WHERE = 75%, BOTH = 75%; Experiment 2: WHAT = 89%, WHERE = 91%, AND = 86%) revealed no differences between the three tasks [Experiment 1: $F_{(2, 58)} = 1.95$, p >.1; Experiment 2: $F_{(2, 28)} = 1.95$, p >.1]. Similarly, the same test on accuracy scores from fMRI subjects (Experiment 1: n = 8, WHAT = 81%, WHERE = 75%, BOTH = 75%; Experiment 2: n = 10, WHAT = 83%, WHERE = 82%, AND = 80%), again revealed no significant differences [Experiment 1: $F_{(2, 14)} = .87$, p > .1; Experiment 2: $F_{(2, 18)} = 1.08$, p > .1].

Voxelwise fMRI Analysis

Voxelwise comparisons (corrected for multiple comparisons) of each delay type *versus* fixation baseline (ITIs) during Experiment 1, or the control delay periods during the sensorimotor control task of Experiment 2, revealed activity in a large expanse of cortex spanning temporal, parietal and frontal areas. Areas showing significantly greater activity during the delay of each task separately *versus* baseline are depicted in two slices through the brain (Figure 2). Also depicted in Figure 2 is the functional topography of these areas, represented by the relative contribution of each task's delay activity towards the total signal observed over all delays. Despite the large degree of overlap observable in these maps, the areas implicated do not respond in an equipotential fashion to the three delay types. Further voxelwise direct comparisons between delay activity levels for the three tasks were performed within the areas that demonstrated greater activity during any delay than during baseline.

Pair-wise comparisons within areas showing greater activation for delays than baseline in Experiment 1 revealed differences between WHAT and WHERE and BOTH delays (Table I and Figure
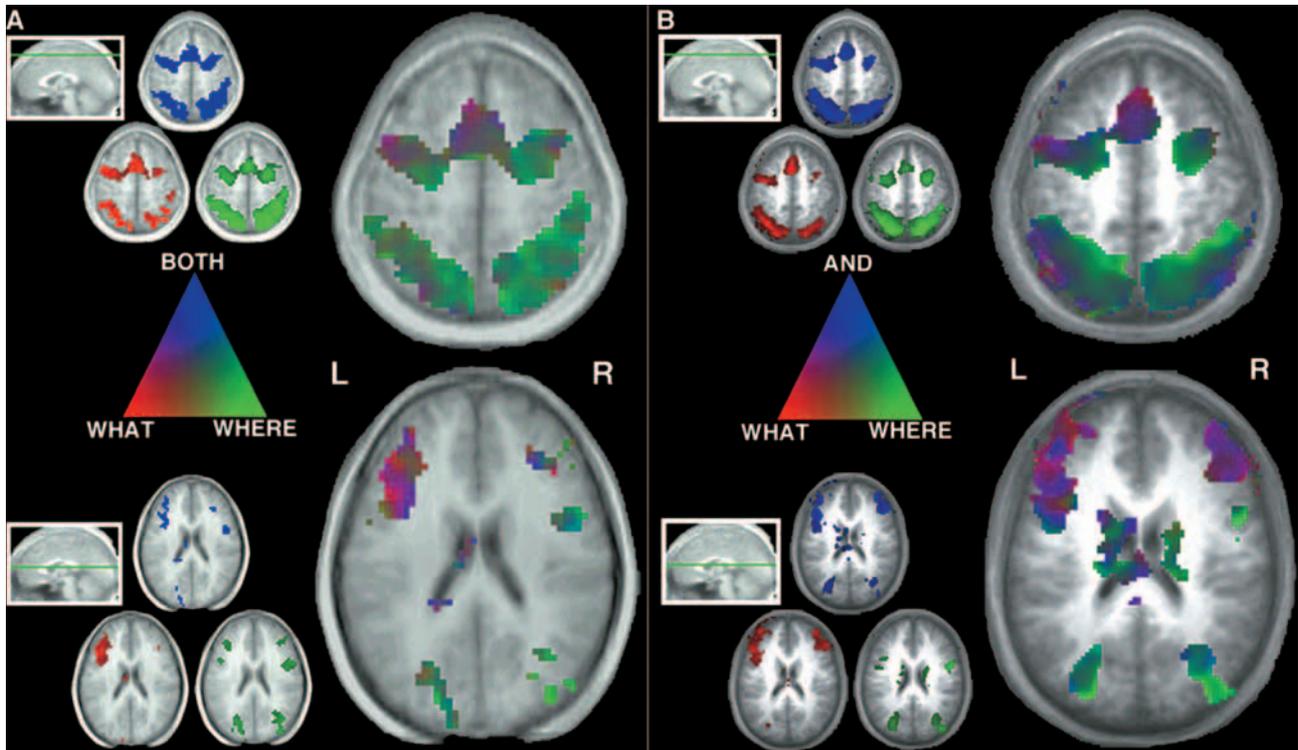
Fig. 2 – *Areas of significant response during each of the delay periods versus A) fixation (Experiment 1) or B) control delay periods (Experiment 2), and the functional topography of these regions. Small maps arranged in a triangle display (clockwise from top) are the areas showing significantly greater response during BOTH/AND (blue), WHERE (green) and WHAT (red) delays compared to baseline. Larger image depicts the functional topography across voxels active during any of the delays. The color indicates the relative strength of response during each of the delays. A voxel heavily dominated by signal during a WHAT delay appears red, WHERE delay green, and BOTH delay blue. As a voxel's response profile becomes more equipotential for the three delays, the color fades to a neutral gray. The green line on the sagittal images appearing next to the maps indicates the approximate slice positioning.*

3). For the comparison of WHAT and WHERE delays, greater WHAT activity was found in a region centered in the left middle frontal gyrus (MFG), extending into the inferior frontal gyrus (IFG) and anterior insula. Greater WHERE activity clusters were found bilaterally at the junction of the superior frontal sulcus (SFS) and precentral sulcus, bilateral superior parietal lobule, left

TABLE I

Activation clusters and center of mass coordinates in Talairach space for the direct comparison of WHAT and WHERE delays and a main effect of delay type for Experiments 1 and 2

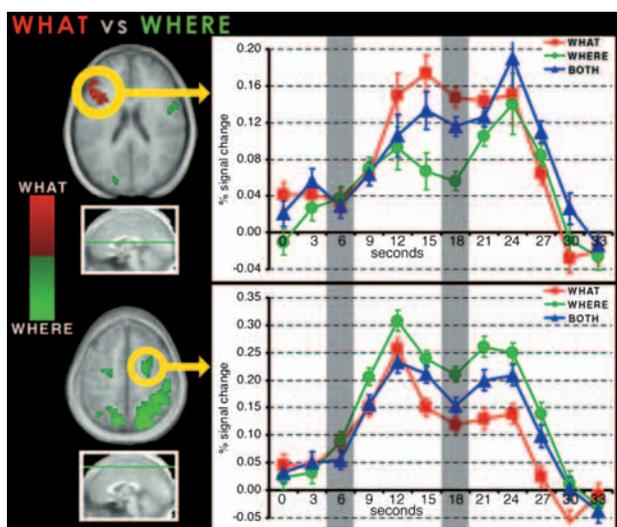|	Experiment 1			Experiment 2		
	x	y	z	x	y	z
What delay > Where delay						
Mid-frontal gyrus (inferior frontal gyrus and anterior insula)	− 40	21	22	− 41	23	26
Cingulate				0	23	41
Where delay > What delay						
Precentral sulcus	51	5	24	51	2	25
Superior frontal sulcus	24	− 6	53	27	− 5	53
Superior frontal sulcus	− 23	− 12	53	− 22	− 8	50
Superior parietal (precuneus, inferior parietal, post central gyrus and cuneus)	29	− 58	48	26	− 63	45
Superior parietal (precuneus, inferior parietal, post central gyrus and cuneus)	− 23	− 60	55	− 20	− 59	51
Middle occipital gyrus				− 26	− 79	24
Middle occipital gyrus	35	− 86	13			
Main effect of delay type						
Mid-frontal gyrus (inferior frontal gyrus and anterior insula)	− 39	19	22	− 42	18	29
Superior frontal sulcus	24	− 6	54	27	− 6	53
Superior frontal sulcus				− 22	− 8	51
Superior parietal (precuneus, inferior parietal, post central gyrus and cuneus)	30	− 60	44	23	− 61	51
Superior parietal (precuneus, inferior parietal, post central gyrus and cuneus)	− 15	− 66	52	− 20	− 61	52
Middle occipital gyrus				35	− 80	22
Middle occipital gyrus				− 25	− 81	25
Cingulate				− 1	21	43

Fig. 3 – *Direct comparison of delay activity in Experiment 1. Areas depicting greater WHAT than WHERE activity are depicted in red while areas showing greater WHERE than WHAT delay activity are depicted in green (all areas p < .05, corrected for multiple comparisons). Graphed to the right of all images is the trial-averaged timecourse of the estimated response across all significant voxels within the region circled in yellow from the presentation of the instruction. Gray bars indicate the presentation of sample and test images respectively.*

cuneus, right precentral sulcus, and right middle occipital gyrus. Depicted in Figure 3 are the trial averaged timecourses for the three tasks averaged across all voxels within left IFG and MFG and right SFS clusters from Experiment 1. These timecourses suggest that the activity within the regions shows a graded response profile across the three delay types, with the activity during the "conjunction" delays less than that of the preferred but more than the nonpreferred information. Similar tests were run on the areas showing greater activity during baseline than during task delay periods, which included parts of cingulate, medial frontal, and auditory and somatosensory cortices. However no differences between WM delay types were found in these areas. Direct pair wise comparisons within regions demonstrating greater activity during any delay than during control delay for Experiment 2 revealed a similar pattern of results (Table I).

ROI Analyses

As the above regions were selected based upon specific pair-wise difference criteria, further consideration of pair-wise differences between the three tasks within these areas are biased. To create regions unbiased for a particular pair-wise difference, a statistical map of voxels displaying a main effect of delay type between the experimental conditions was computed, and corrected for multiple comparisons. In Experiment 1, this analysis revealed activation foci in the left IFG/MFG, right SFS, and bilateral superior parietal cortex. These two frontal regions were found to

encompass both the respective IFG/MFG and SFS regions resulting from the previous comparisons, and were defined as ROIs. Pair wise differences between the three tasks within these ROIs were tested. The IFG/MFG region displayed significant differences for all comparisons with WHAT greater than WHERE, BOTH greater than WHERE, and WHAT greater than BOTH [$t(7)$ = 8.1, 2.6, 3.9; $p < .01$, .05, .01 respectively]. The SFS region displayed significant differences for all comparisons as well, with WHERE greater than WHAT, BOTH greater than WHAT, and WHERE greater than BOTH [$t(7)$ = 4.0, 3.1, 3.6; $p < .01$, .05, .01 respectively]. Similar paired t-tests were run on the left and right parietal regions identified by this analysis. The left parietal region displayed significantly greater activity for the WHERE delays than WHAT delays, and WHERE delays than BOTH delays [$t(7)$ = 4.5, 4.8 respectively; $p < .01$], but no difference between activity levels for WHAT and BOTH delays [$t(7)$ = − 1.41, $p > .1$]. The right parietal cortex, however, displayed significant differences for all comparisons with WHERE greater than WHAT, BOTH greater than WHAT, and WHERE greater than BOTH [$t(7)$ = 5.6, 3.9, 3.9 respectively; $p < .01$].

In Experiment 2, significant foci for this main effect of delay type were identified in the left IFG/MFG, bilateral SFS, bilateral superior parietal lobule, bilateral middle occipital gyrus and the supplementary motor area (SMA) extending into the cingulate. Within these regions, pair wise differences between our tasks were again tested and the timecourse of response time-locked to the instruction and sample relative to that of the control was extracted (Figure 4). Within the frontal cortex, the left IFG/MFG and bilateral SFS showed the graded response for all three tasks as found in Experiment 1. The IFG/MFG region displayed WHAT greater than WHERE, AND greater than WHERE, and WHAT greater than AND delays [$t(9)$ = 5.1, 4.7, 2.6; $p < .01$, .01, .05 respectively]. Both left and right SFS regions displayed WHERE greater than WHAT, AND greater than WHAT, and WHERE greater than AND [$t(9)$ = 4.9, 5.5, 2.5, $p < .01$, .01, .05 respectively for the left and $t(9)$ = 9.4, 4.0, 2.6, $p < .01$, .01, .05 respectively for the right]. The SMA/cingulate however did not show this pattern. While showing greater activity for WHAT than WHERE and AND than WHERE [$t(9)$ = 4.4, 5.1 respectively, $p < .01$] no such difference was found between WHAT and AND [$t(9)$ = 1, $p > .1$].

The results of Prabhakaran et al. (2000) suggested that an area of activation within the right middle and superior frontal gyri might represent a WM buffer for the bound representation of letters and spatial locations. An anatomical ROI was drawn about these two gyri. The percent signal change during each of the delays was averaged over all voxels within this ROI for each subject
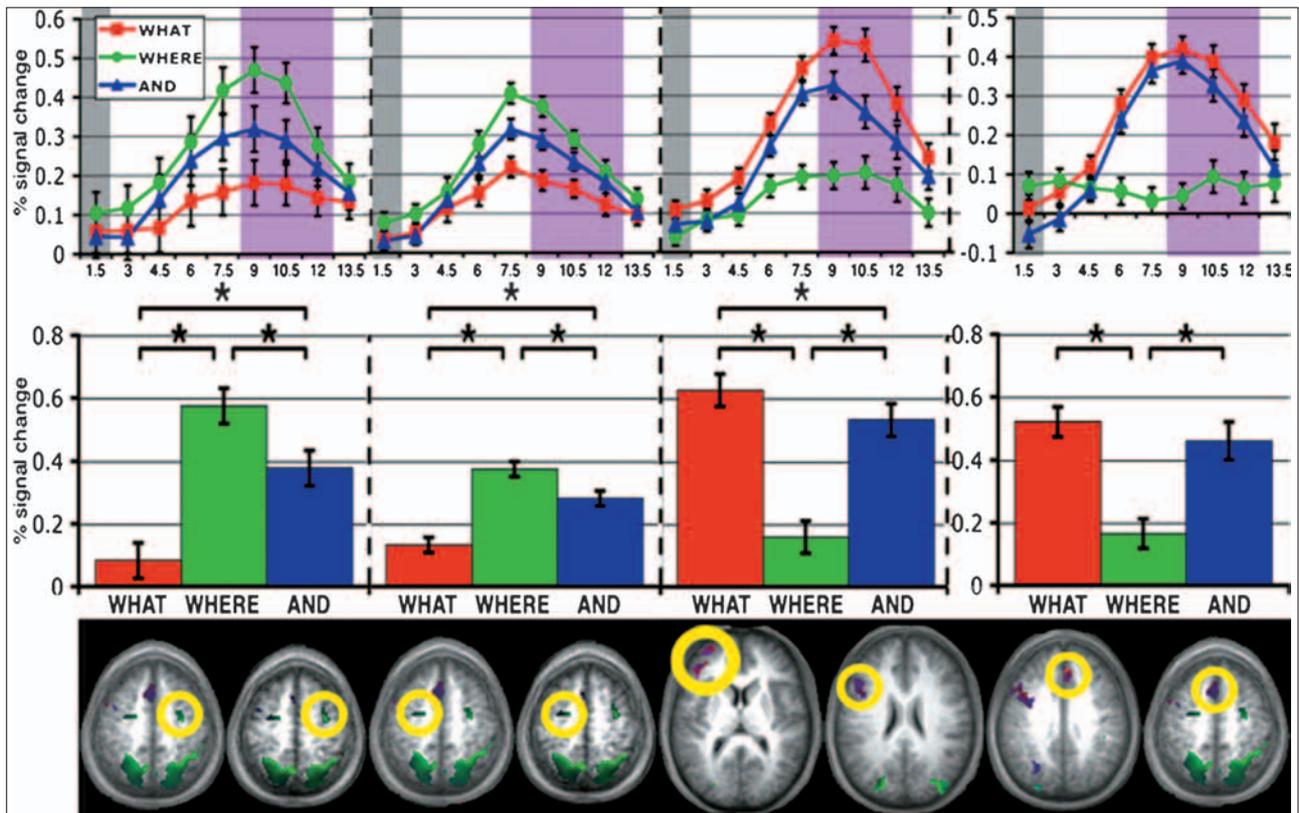
Fig. 4 – *Percent signal change from control delay for each of the three delay types for regions displaying a main effect of delay in Experiment 2. Line graphs plot the difference in response to each working memory (WM) task instruction and sample and that of the control instruction and sample averaged across each of the regions. Gray bars indicate presentation of instruction/sample, and pink bar indicates window during which the test image could appear. Functional topography maps underneath each bar graph illustrate voxels included in the functionally defined region of interest (ROI). All colored voxels within the yellow circle were part of the individual ROI.*

from Experiment 1 and these values were entered into a repeated measures ANOVA. In an attempt to ensure that differences within this ROI were attributed to a positive maintenance signal relative to the nonmnemonic baseline as is normally associated with delay period activity, the ROI was further restricted to include only voxels that showed a significant increase from baseline in any of the tasks. No main effect of delay type (WHAT, WHERE, BOTH/AND) was found [F (2, 14) = .087, $p > .1$], indicating that there were no differences in delay period activity among the three delay types within this ROI (the same test performed on the anatomical region without the "positive activation" restriction lead to equivalent results). This procedure was repeated for the corresponding area from Experiment 2, and again no differences were observed [F (2, 28) = .73, $p > .1$]. A similar analysis was performed on the voxels within an anatomically defined ROI based upon the Munk et al. (2002) report of a medial frontal region, extending from the SMA to the anterior cingulate that showed greater activity for the maintenance of the combination of the identity of vegetables and their locations than for either information type alone. Again, however, no main effect of delay type was observed [F (2, 14) = .626, $p > .1$]. In Experiment 2 the functionally defined medial frontal (SMA/cingulate) cluster of

activation was largely overlapping with this anatomical ROI. As described above, this functionally defined region did show differences between WHAT and WHERE, and AND and WHERE delays, but not between WHAT and AND (see Table I and Figure 4).

Laterality

Results of the direct comparisons between WHAT and WHERE suggested that there might be a laterality difference between the tasks in addition to the dorsal/ventral difference. To test whether the response elicited during a given delay type was lateralized to the left or right hemisphere in anterior and posterior cortex, activity for each delay during Experiment 1 was extracted from all clusters of voxels showing significant delay response for any of the tasks. For each delay type a comparison of percent signal change within the left *versus* right anterior cortex, and left *versus* right posterior cortex was performed. Within anterior activations, percent signal change was greater in the left hemisphere than right during WHAT delays, greater in the right hemisphere than left during WHERE delays, and not significantly different in right and left hemispheres during BOTH delays [$t(7) = 5.4, -2.4, 1.0; p < .01, p < .05, p > .1$ respectively]. Posterior activations did not show significant differences in

laterality of response within WHAT, WHERE, or BOTH delays [t(7) = 1.0, -1.9, .7; p > .1, .05, .1 respectively] though there was a trend for percent signal change to be greater in right posterior areas than left posterior areas for the WHERE task. Similar analyses were performed on the data from Experiment 2. Within anterior activations, percent signal change was greater in the left hemisphere than right during WHAT delays, but no differences were observed between left and right hemispheres for the WHERE or AND delays [t(9) = 2.60, -1.06, 1.63; p < .05, p > .1, .1 respectively]. No significant laterality of response for WHAT, WHERE, or AND was observed in posterior regions [t(9) = 1.51, -1.07, 1.86 respectively; all p > .1].

<center>DISCUSSION</center>

<center>*Summary*</center>

In the current study, as frequently reported in other studies, activation across both dorsal and ventral frontal cortex is significantly greater than a fixation baseline or a sensorimotor control delay during all WM delay types. However, this does not lead to an equipotential response across the frontal cortex (see Figure 2). Rather, differences in the type of information that must be maintained in each of the tasks reveal a relative dorsal-ventral/spatial-nonspatial functional topography across the frontal cortex, as has also been previously reported (Courtney et al., 1996, 1998; Munk et al., 2002; Sala et al., 2003; Sakai and Passingham, 2003). A hemispheric laterality difference was also observed between the tasks as has been previously noted (D'Esposito et al., 1998; Smith and Jonides, 1997) with WHERE right lateralized and WHAT left lateralized, but this was in addition to, rather than instead of, a dorsal/ventral topography. These comparisons demonstrate that even though both dorsal and ventral areas are responding to both WHAT and WHERE delay types, there exists a "preference" for a particular type of information within each of these regions. An analysis concern for studies of WM that interrogate "delay" period activity in delayed recognition tasks is that this signal can be contaminated by activity at sample. However, it is unlikely that the entire effect that we are observing is due to contamination from the sample presentation. Previously, others (Pessoa et al., 2002) and ourselves (Sala et al., 2003) have taken steps to provide a greater degree of separation for the regressors corresponding to sample and delay. This additional analysis however returned qualitatively equal results. Furthermore, although no statistical inferences are drawn from the timecourses presented, they are included to visualize the activity elicited over time, and demonstrate that the differences between the

conditions exist late in the delay as well as early.

The magnitude of the response within regions of the PFC can distinguish among all of the three tasks' delay periods. During BOTH/AND delays, signal is significantly greater than that of the "non-preferred" single information, and significantly less than that of the "preferred" single information within a cortical area. This change in signal, in contrast to a simple addition of the spatial and nonspatial responses within an area or a response equal to that of the preferred information, suggests that the same regions and cell populations may integrate spatial and nonspatial information when this integration is task relevant as explained below.

Though the BOTH task encourages the formation of a conjoined representation of spatial and nonspatial information, subjects could have formed separate representations of these information types and rapidly switched between them, dividing their attention between the two information domains. Divided attention between spatial and nonspatial information could result in the reduced activation observed in the BOTH task relative to the focused attention on either spatial or nonspatial information in the WHERE and WHAT tasks, respectively. However, had subjects been spending half of the time rehearsing spatial and half the time rehearsing nonspatial information, one might expect a decrement in behavioral performance on the BOTH task. This was not the case in the accuracy measures for the three tasks. Furthermore, neurophysiological evidence indicates that the information carried by single cells in PFC during BOTH delays can be selective for the conjunction of spatial and nonspatial information continuously throughout the delay. Those results also argue against the idea that an object and its location are maintained simultaneously via divided attention.

Behavioral evidence suggests that the relationship between an object and its location may be held in WM, or the two pieces of information may be maintained separately, depending on task demands (Sala, 2003). No area demonstrated greater activity for the BOTH delays as compared to WHAT or WHERE, despite the apparent additional information that needed to be maintained. The conjunction task of Experiment 2 was altered to ensure that subjects were maintaining a representation of the "bind" between spatial and non-spatial information. During the AND task, knowledge of spatial and nonspatial attributes of the samples alone were not sufficient to perform correctly, as distractor test images matched both these individual pieces of information. There did not appear to be any qualitative difference in the patterns of activation observed for the BOTH task as compared to the AND task, although a direct comparison across experiments was not possible. In both experiments there was no evidence for BOTH/AND delay

activity being an additive measure of the responses during the WHAT and WHERE tasks (see also Munk et al., 2002). Rather, within areas demonstrating differences among the three delays, the activity during BOTH/AND delays was less than that of the preferred and greater than that of the non-preferred information (Figure 4).

Implications for Understanding the Functional Topography of Frontal Cortex

The results of the current and previously reported experiments in both humans and monkeys indicate that the PFC conforms to neither a stable domain-sensitive nor a stable domain-insensitive organizational principle. As in previous studies, significant activity relative to baseline or the sensorimotor control task was found for each PFC region's nonpreferred information, a result sometimes thought to be at odds with a static and domain-sensitive functional organization. This need not be the case, however, if this activity represents some incidental encoding of the non-relevant information. Alternatively, this activation above baseline for the nonpreferred task could be interpreted as indicating the existence of intermixed spatially and nonspatially-selective cell populations with a greater percentage of primarily spatially-tuned neurons dorsally and a greater percentage of primarily object-tuned neurons ventrally. This view has been previously proposed both by one of the authors of the current study (Haxby et al., 2000) and by others (e.g., see Duncan, 2001). The relative distributions account, however, is at odds with the results of the current experiments. If dorsal and ventral PFC areas contained a mixture of "what cells" and "where cells", then activity during BOTH/AND delays would be greater than the activity during maintenance of either information alone, because both cell populations would now be active instead of only one or the other. The result that BOTH/AND activity was significantly less than the preferred information type suggests a more selective, integrated representation of the information during this task, perhaps with a smaller number of cells active than during the region's preferred task, but with those cells more active than during the region's nonpreferred task. Therefore, while the results of the current and previous studies showing regional preferences for spatial *versus* nonspatial information are at odds with a stable domain-general organization, a stable domain-sensitive, or even relative domain-sensitive organization, has no way to account for a decrease during "bound" maintenance.

Biased Competition Model of WM

The results of both the current and previous research suggest that frontal cortex can maintain an integrated representation of spatial and nonspatial information when this information is task relevant, but when an integrated representation is not task relevant, dorsal and ventral frontal regions distinguish between spatial and nonspatial information. Furthermore, the bound representation of these information types is produced through an overall reduction in response, distributed over the same cortex responsible for representing the individual information types. Such an integrated representation could be created from intrinsic and experience-dependent connections between regions of frontal cortex that enable learning of associations between stimuli and task relevant behavior (Asaad et al., 2000; Miller, 2000a; White and Wise, 1999). The pattern of results from both the current and previous studies can be realized by extending the concepts of biased competition (Desimone, 1996; Desimone and Duncan, 1995). While it is often proposed that the PFC is the source of "biasing signals" that target earlier processing areas and representations of information (Miller and Cohen, 2001; Deco and Rolls, 2003; Chelazzi et al., 1998, 2001; Desimone, 1998; Reynolds and Desimone, 1999; White and Wise, 1999), the PFC has not traditionally been considered as a target of "biasing signals" (though see Almeida et al., 2004). This extension of biased competition within the frontal cortex (depicted in Figure 5), however, allows for flexible, dynamic frontal representations to emerge dependent upon the task requirements. Domain specific regions of the visual system initially innervate dorsal and ventral regions of frontal cortex. In the proposed model, this information is then distributed across the frontal cortex through the combination of local (within a PFC region) competitive interactions and inter-regional (across regions of PFC) excitatory signals.

During the presentation of an item for maintenance, the dorsal frontal cortex receives the majority of spatially selective inputs while ventral frontal cortex receives the majority of nonspatially selective inputs due to connections with domain specific visual processing areas. Through connections between these dorsal and ventral frontal regions, both intrinsic and learned through experience, spatial inputs are passed to ventral frontal regions and nonspatial inputs are passed to dorsal frontal regions, distributing the "to be remembered" information across the frontal cortex and enabling single cells to have dual object-in-location selectivity. It is worth noting that, within this framework, the information carried across the frontal cortex is the task relevant information, regardless of where the response is measured. Therefore, during a WHAT task, activity in both dorsal and ventral PFC represents primarily nonspatial information as this is currently task-relevant, just as during the BOTH/AND task the activity would reflect the integrated representation
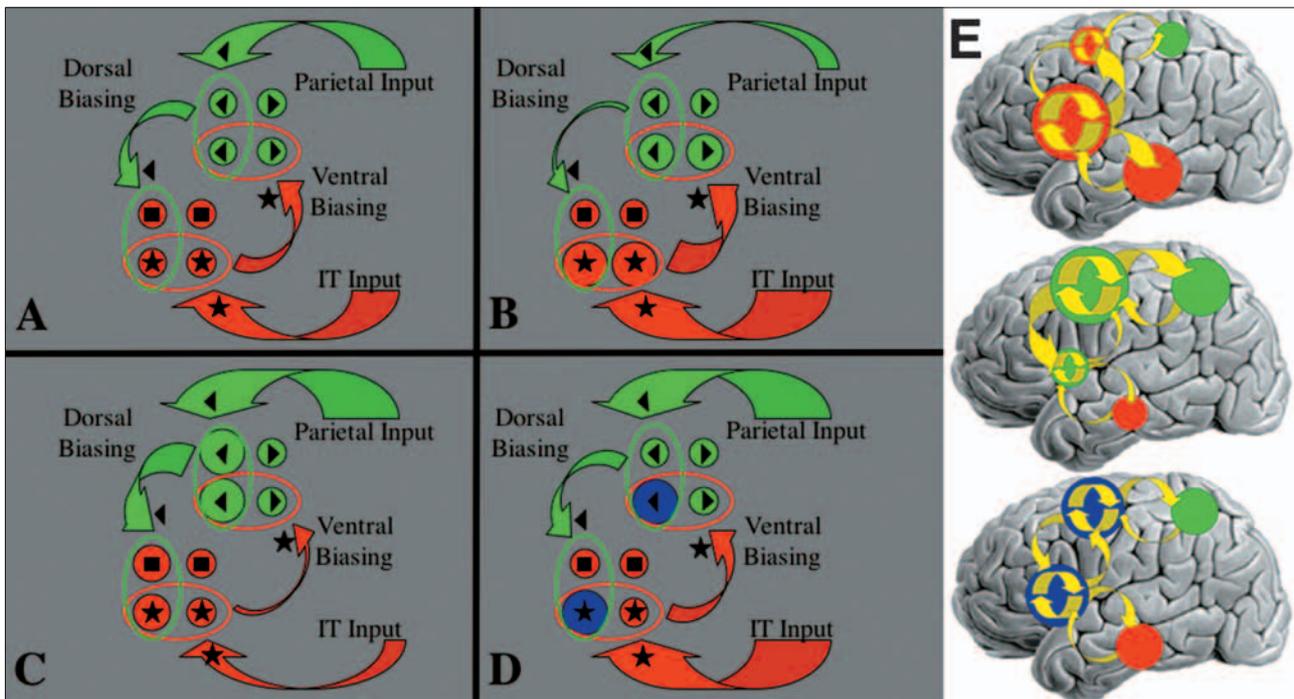
Fig. 5 – A) A model extending the principles of biased competition to the frontal cortex for the maintenance of information in working memory (WM). Cells within the dorsal cortex are depicted in green, cells within the ventral cortex are depicted in red. Within a region, cells are connected in a competitive fashion. Biasing signals flow between dorsal and ventral frontal regions, allowing cells processing "left" (outlined by green oval) in dorsal cortex to influence processing of ventral cells processing both stars and squares (also outlined by green oval). Similarly biasing signals originating from cells processing "star" (outlined by red oval) in ventral cortex can influence "left" and "right" cells (outlined again in red). B) State of dynamic system when nonspatial information is to be maintained. C) State of system when spatial information is to be maintained. D) State of system when the conjunction of nonspatial and spatial information is to be maintained. E) Illustration of the synaptic and cellular activity and the type of selectivity predicted by the model during the WHAT, WHERE, and BOTH/AND tasks. Each circle represents the elicited fMRI activity from a region during a specific task. The color of the circle identifies the type of information represented (red = nonspatial, green = spatial, and blue = integrated) while the size of the circle indicates the relative amount of activity measured. Yellow arrows indicate the synaptic connections contributing measurable input to a region, again with size indicating the relative amount of this input.

of spatial and nonspatial information across dorsal and ventral PFC. Incidental encoding and maintenance of non-relevant information could also be observed, but this would have a less robust representation (i.e., more broadly tuned selectivity; Duncan, 2001) than the relevant information. The strength and direction of these inter-regional excitatory connections would presumably be governed by the representation of task instructions, goals, and context represented hierarchically in more anterior portions of the PFC (Koechlin et al., 2003; Braver and Bongiolatti, 2002).

For example, consider Figure 5 under the case when a star is presented on the left during a WM task. When only nonspatial information is task relevant (5B), the ventral frontal cortex receives the task relevant inputs via direct innervation from ventral extrastriate cortex and imparts this information to the dorsal frontal cortex. Spatial inputs to dorsal frontal cortex via direct innervation from parietal cortex have little effect on the activity in frontal cortex as maintenance of this information is attenuated, perhaps through reduced functional connectivity (Koechlin et al., 2003; Sakai and Passingham, 2003). All cells that are selective for the sample stimulus shape are active in both dorsal and ventral regions, independent of

the location of the stimulus. However, when only spatial information is task relevant (5C), the dorsal frontal cortex becomes more activated by these direct inputs and imparts this information to the ventral frontal cortex. All cells that are selective for the location of the sample stimulus are active in both dorsal and ventral regions, independent of the object used to mark that location. In both these cases, the task relevant information is distributed across the PFC. When the conjunction of spatial and nonspatial information is task relevant (5D), these excitatory inputs flow both directions between dorsal and ventral frontal cortex, biasing competition within each region and resulting in a smaller, more selective, subset of cells (i.e., only those that prefer both the object and it's location) being maximally active. The resulting pattern of fMRI response is an overall decrease of activity summed across all cells within each of the regions. This activity now integrates the two information types, and demonstrates both spatial and nonspatial selectivity.

Current research suggests that the fMRI signal reflects primarily synaptic activity rather than spiking activity (Logothetis et al., 2001), though spiking activity is not to be ruled out as a component of the fMRI signal (for discussion see

Bandettini and Ungerleider, 2001). If the system producing sustained activity enabling visual WM is viewed as a distributed system with reentrant processing, then the interactions between the frontal regions, local processing within a region, and each region's interaction with domain specific processing streams will affect the synaptic activity observed during fMRI experiments of WM (Haxby et al., 2000; Fuster, 2000; Zipser et al., 1993; Goldman-Rakic, 1995b; Tagamets and Horwitz, 2000). A pictorial representation of the fMRI signal from areas involved in such a network is depicted in Figure 5E. Each circle represents the elicited fMRI activity from a region during a specific task. The color of the circle identifies the type of information represented (red = nonspatial, green = spatial, and blue = integrated) while the size of the circle indicates the relative amount of activity measured. Yellow arrows indicate the synaptic connections contributing measurable input to a region, again with size indicating the relative amount of this input.

The model stresses the direct input from the domain specific processing regions, as well as the strength of response within a region due to that input. It is likely that these components contribute more to the fMRI signal than do the inter-regional biasing signals. Results from lesion studies (Carlesimo et al., 2001; Mottaghy et al., 2002; Bechara et al., 1998) are in accord with this model, as disruption of dorsal or ventral regions of frontal cortex would lead to a difficulty in communicating the domain specific information across the frontal cortex to create this distributed, integrated representation. Hence, a dorsal frontal lesion would limit the amount of spatial information that can gain access to frontal resources for maintenance, while the majority of nonspatial inputs would still be unaffected.

These results and the interpretation presented here also coincide with findings from single unit recordings and flesh out a mechanism of prefrontal interactions suggested by these studies. Rao et al. (1997) trained monkeys to first remember an object, then to find that object in an array and remember the location in which it was presented. They found that some of the cells appear to "switch modes". After conveying object selective activity during the WHAT delay, they then began to code location selective information over WHERE delays, mirroring the demands of the task. The proposed model is in agreement with these results because the task relevant information, spatial or nonspatial, is distributed across the frontal cortex via the biasing signals, which would change according to task demands. On one trial a cell may show activity for the memory of "star" during an initial WHAT delay, yet during a WHERE delay it may respond for memory of "left" regardless of what item was cueing that memory. In addition, the model predicts differential

time courses for the development of object and spatially selective responses in ventral and dorsal frontal cortices. In dorsal frontal cortex, spatial selectivity in the response would appear earlier than selectivity for an object. The opposite pattern should be seen in ventral frontal cortex.

Another set of cells' activity was also described by Rao et al. (1997) as "predominately location tuned" but modulated by the object that cued the location. Rainer et al. (1998), using a task similar to our BOTH task, specifically studied these "conjunction" cells. They found that when monkeys were trying to remember objects in locations, individual neurons coded the conjunction of object and spatial information. A neuron that demonstrated activity preferential for a certain object was differentially active through a delay period over a specific and restricted region in space. Similarly, in the proposed model, when both spatial and nonspatial information are task relevant, the information from one input stream of processing will have a modulatory effect on the response to the other stream's input. The proposed model offers an explanation as to how the formation of these "conjunction" cells may arise through inter-regional biasing signals, local competitive interactions, and the task relevancy of information.

The dynamic nature of the representation of information for maintenance has behavioral implications as well. Despite what may appear to be an additional "load" for the BOTH/AND trials as compared to the single information trials of WHAT and WHERE, the biased competition model states that in each condition the load in each condition is three items. Accordingly, the behavioral performance is stable across the three tasks. The distinction between these three trial types is at the representational level. Rather than the BOTH/AND representation being 3 locations plus 3 objects, the representation is instead 3 integrated wholes (as in Vogel et al., 2001). However, in accordance with subsequent research by Wheeler and Triesman (2002), the creation of this integrated object and location is task dependent and not the default state of the system. Further behavioral studies with the WHAT, WHERE, and AND tasks offer confirmatory evidence that the representation formed during the delay periods differ according to the task demands as suggested by the current model (Sala, 2003).

Relevance to Other Models of PFC Function

Two influential models of PFC function have recently been proposed, the integrative model by Miller and Cohen (2001) and the Adaptive Coding model by Duncan (2001). These theories are part of a larger body of work (e.g., Fuster, 1985; Yajeya et al., 1988; Cohen and Servan-Schreiber, 1992; White and Wise, 1999) supporting the idea that the

PFC represents abstract representations that serve to bias activity in lower regions according to the precise goals and task contingencies. Under this schema the sustained activity within the PFC during our delayed recognition tasks represent the goals of the task, namely the production of a certain response if a stimulus matching along the relevant dimension appears at test. To achieve this, both these earlier models and the proposed Biased Competition Model of WM state that the PFC engages in cognitive control, issuing biasing signals favoring the information to be maintained. These theories invoke the PFC as an "activity-based" mechanism of control for attentional selection and agree that information that enters distinct portions of the PFC from separate processing streams is eventually integrated across the frontal cortex when such an integrated representation is task relevant. However, while the integrative model of PFC and the Adaptive Coding model remain relatively agnostic as to how the PFC itself is organized, the model proposed here directly addresses this issue, and thus serves to extend these models. The proposed model suggests that the task relevant information is represented across the frontal cortex, created through excitatory interactions between, and competitive interactions within, functionally specialized PFC regions as outlined above.

The current model considers the PFC as hierarchically organized, as has been frequently suggested (Koechlin et al., 2003; Fuster, 2001; Sakai and Passingham, 2003; Courtney, 2004) with a vital role in maintaining multiple types of information necessary for appropriate behavior. In this view, the PFC plays a role in maintaining both relevant sensory information and the abstract rules of the task. Sustained activity representing abstract task rules and contextual information can provide biasing signals directly to sensory and motor brain systems (as is the emphasis of the Miller-Cohen and Duncan models), or these abstract representations can provide biasing signals that influence the type of representation of the sensory information that will be most useful to be able to complete the task at some future time (after the delay period). It is this ability and the corresponding implications for the organization of the PFC that is emphasized in the current model.

Another model of selective representation of task-relevant information in WM based on the principles of biased competition was recently proposed by Almeida et al. (2004). Their neurodynamic computational model demonstrated the feasibility of such a mechanism for the selective representation and distribution of task relevant information across domain dependent modules. In their model, separate populations of "integrated", "object", and "spatial" neurons are assumed to exist. Through competitive interactions within these modules and cooperative interactions

between them, the response from one module (e.g., 'object') can influence the selective response within another (e.g., 'spatial'). Our biased competition model assumes a similar mechanism of competitive and cooperative (bias) signals. However, the current model does not propose distinct modules for "spatial", "object", or "integrated" representation. Rather, the task relevant information is distributed and integrated over the entire PFC. Indeed, the same cells that appear to be sensitive for one type of information (e.g., spatial), may "switch" their response profiles as the demands of the task change as has been observed previously (Rao et al., 1997). This model also accounts for the large degree of overlap between areas responsive during "object" and "spatial" WM delays observed in neuroimaging (Baker et al., 1996; Owen et al., 1996; D'Esposito et al., 1998; Postle and D'Esposito, 1999; Nystrom et al., 2000; Postle et al., 2000a, 2000b; Stern et al., 2000), as well as the regional dissociations observed within the PFC during these delay types (Courtney et al., 1996, 1998; Munk et al., 2002; Sala et al., 2003; Sakai and Passingham, 2003; Rama et al., 2004). Furthermore, In the proposed model, it is the communication between regions of PFC receiving the majority of inputs from the visual processing streams, and the competitive interactions within these regions PFC that create a reentrant circuit to most accurately maintain only the task relevant information, and enable integration of multiple types of information at the individual neuron level (Rainer et al., 1998).

The suggestion of a functional topography based upon stimulus properties in no way precludes other organizational schemas from being represented in the frontal cortex (e.g., Owen et al., 1998; Smith and Jonides, 1999; White and Wise, 1999; Asaad et al., 2000; D'Esposito et al., 2000; Petrides, 2000; Johnson et al., 2003; Koechlin et al., 2003). The tasks used in the present experiment and in others having found this topography (Courtney et al., 1996, 1998; Munk et al., 2002; Sala et al., 2003), have held other cognitive demands constant across tasks and therefore cannot address the existence of other functional topographies. The Biased Competition Model of WM proposed here would certainly allow for, *and indeed predicts*, the existence of additional topographies. In the current experiments, however, all three WM tasks required maintenance of task instructions, and therefore no differences in the amount of activation would be expected in PFC regions representing this non-domain-specific information. Other organizational schemas can utilize the same principles of competition and biasing signals to convey information for maintenance, processing, and rule guided behavior. Such a model would suggest that as regions of the frontal cortex interact, more selective and complex response patterns can be constructed throughout the

frontal cortex in order to achieve goal-oriented actions (Miller, 2000b; Miller and Cohen, 2001; Koechlin et al., 2003). Indeed the proposed model places great demand on the task goals to shape the way in which information will be integrated within the PFC. A recent study by Koechlin et al. (2003) suggesting a hierarchical model of PFC functioning, demonstrated that anterior PFC can effect the processing of lateral PFC for specific tasks in accordance with manipulations of high level instructional sets. A similar mechanism could be at work in our tasks, on each trial, allowing for the instructional cue to effect or adjust the processing and interactions within the PFC in order to represent the task demands. The proposed model of frontal organization extends previous theories and research by suggesting a novel way in which regions of the PFC may interact with one another. These hierarchically organized interactions within the PFC result in flexible, dynamic, and sustainable representations of information necessary for the completion of goal directed behavior.

REFERENCES

ALAIN C, ARNOTT SR, HEVENOR S, GRAHAM S and GRADY CL. "What" and "where" in the human auditory system. *Proceedings of the National Academy of Sciences of the USA, 98:* 12301-12306, 2001.
ALMEIDA R, DECO G and STETTER M. Modular biased-competition and cooperation: A candidate mechanism for selective working memory. *European Journal of Neuroscience, 20:* 2789-2803, 2004.
ANDERSON JR, BOTHELL D, BYRNE MD, DOUGLASS S, LEBIERE C and QIN Y. An integrated theory of the mind. *Psychological Review, 111:* 1036-1060, 2004.
ASAAD WF, RAINER G and MILLER EK. Task-specific neural activity in the primate prefrontal cortex. *Journal of Neurophysiology, 84:* 451-459, 2000.
BADDELEY A and HITCH GJ. Working memory. In Bower GA (Ed), *Recent Advances in Learning and Memory.* New York: Academic Press, 1974.
BAKER SC, FRITH CD, FRACKOWIAK RS and DOLAN RJ. Active representation of shape and spatial location in man. *Cerebral Cortex, 6:* 612-619, 1996.
BANDETTINI PA and UNGERLEIDER LG. From neuron to bold: New connections. *Nature Neuroscience, 4:* 864-866, 2001.
BECHARA A, DAMASIO H, TRANEL D and ANDERSON SW. Dissociation of working memory from decision making within the human prefrontal cortex. *Journal of Neuroscience, 18:* 428-437, 1998.
BRAVER TS and BONGIOLATTI SR. The role of frontopolar cortex in subgoal processing during working memory. *NeuroImage, 15:* 523-536, 2002.
CARLESIMO GA, PERRI R, TURRIZIANI P, TOMAIUOLO F and CALTAGIRONE C. Remembering what but not where: Independence of spatial and visual working memory in the human brain. *Cortex, 37:* 519-534, 2001.
CHELAZZI L, DUNCAN J, MILLER EK and DESIMONE R. Responses of neurons in inferior temporal cortex during memory-guided visual search. *Journal of Neurophysiology, 80:* 2918-2940, 1998.
CHELAZZI L, MILLER EK, DUNCAN J and DESIMONE R. Responses of neurons in macaque area v4 during memory-guided visual search. *Cerebral Cortex, 11:* 761-772, 2001.
COHEN JD and SERVAN-SCHREIBER D. Context, cortex, and dopamine: A connectionist approach to behavior and biology in schizophrenia. *Psychological Review, 99:* 45-77, 1992.
COURTNEY SM. Attention and cognitive control as emergent properties of information representation in working memory. *Cognitive, Affective, and Behavioral Neuroscience, 4:* 501-516, 2004.
COURTNEY SM, PETIT L, MAISOG JM, UNGERLEIDER LG and HAXBY JV. An area specialized for spatial working memory in human frontal cortex. *Science, 279:* 1347-1351, 1998.
COURTNEY SM, UNGERLEIDER LG, KEIL K and HAXBY JV. Object and spatial visual working memory activate separate neural systems in human cortex. *Cerebral Cortex, 6:* 39-49, 1996.
COX RW. Afni: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research, 29:* 162-173, 1996.
DECO G and ROLLS ET. Attention and working memory: A dynamical model of neuronal activity in the prefrontal cortex. *European Journal of Neuroscience, 18:* 2374-2390, 2003.
DESIMONE R. Neural mechanisms for visual memory and their role in attention. *Proceedings of the National Academy of Sciences of the USA, 93:* 13494-13499, 1996.
DESIMONE R. Visual attention mediated by biased competition in extrastriate visual cortex. *Philosophical Transactions of the Royal Society of London B, 353:* 1245-1255, 1998.
DESIMONE R and DUNCAN J. Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18:* 193-222, 1995.
D'ESPOSITO M, AGUIRRE GK, ZARAHN E, BALLARD D, SHIN RK and LEASE J. Functional MRI studies of spatial and nonspatial working memory. *Cognitive Brain Research, 7:* 1-13, 1998.
D'ESPOSITO M, POSTLE BR, BALLARD D and LEASE J. Maintenance versus manipulation of information held in working memory: An event-related fMRI study. *Brain and Cognition, 41:* 66-86, 1999.
D'ESPOSITO M, POSTLE BR and RYPMA B. Prefrontal cortical contributions to working memory: Evidence from event-related fMRI studies. *Experimental Brain Research, 133:* 3-11, 2000.
DUNCAN J. An adaptive coding model of neural function in prefrontal cortex. *Nature Reviews Neuroscience, 2:* 820-829, 2001.
FUSTER JM. The prefrontal cortex, mediator of cross-temporal contingencies. *Human Neurobiology, 4:* 169-179, 1985.
FUSTER JM. Executive frontal functions. *Experimental Brain Research, 133:* 66-70, 2000.
FUSTER JM. The prefrontal cortex – An update: Time is of the essence. *Neuron, 30:* 319-333, 2001.
FUSTER JM, BODNER M and KROGER JK. Cross-modal and cross-temporal association in neurons of frontal cortex. *Nature, 405:* 347-351, 2000.
GOLDMAN-RAKIC PS. Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In Plum F and Mountcastle F (Eds), *Handbook of Physiology.* Washington D.C.: The American Phsyiological Society, 1987.
GOLDMAN-RAKIC PS. Architecture of the prefrontal cortex and the central executive. *Annals of the New York Academy of Sciences, 769:* 71-83, 1995a.
GOLDMAN-RAKIC PS. Cellular basis of working memory. *Neuron, 14:* 477-485, 1995b.
HAXBY JV, PETIT L, UNGERLEIDER LG and COURTNEY SM. Distinguishing the functional roles of multiple regions in distributed neural systems for visual working memory. *NeuroImage, 11:* 145-156, 2000.
JIANG Y, OLSON IR and CHUN MM. Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26:* 683-702, 2000.
JOHNSON MK, RAYE CL, MITCHELL KJ, GREENE EJ and ANDERSON AW. fMRI evidence for an organization of prefrontal cortex by both type of process and type of information. *Cerebral Cortex, 13:* 265-273, 2003.
KOECHLIN E, ODY C and KOUNEIHER F. The architecture of cognitive control in the human prefrontal cortex. *Science, 302:* 1181-1185, 2003.
LEVY R and GOLDMAN-RAKIC PS. Segregation of working memory functions within the dorsolateral prefrontal cortex. *Experimental Brain Research, 133:* 23-32, 2000.
LOGIE RH. *Visuo-Spatial Working Memory.* Hove: Lawrence Erlbaum Associates Inc., 1995.
LOGOTHETIS NK, PAULS J, AUGATH M, TRINATH T and OELTERMANN A. Neurophysiological investigation of the basis of the fMRI signal. *Nature, 412:* 150-157, 2001.

MILLER EK. Organization through experience. *Nature Neuroscience, 3:* 1066-1068, 2000a.

MILLER EK. The prefrontal cortex and cognitive control. *Nature Reviews Neuroscience, 1:* 59-65, 2000b.

MILLER EK and COHEN JD. An integrative theory of prefrontal cortex function. *Annual Review Neuroscience, 24:* 167-202, 2001.

MOTTAGHY FM, GANGITANO M, SPARING R, KRAUSE BJ and PASCUAL-LEONE A. Segregation of areas related to visual working memory in the prefrontal cortex revealed by rtms. *Cerebral Cortex, 12:* 369-375, 2002.

MUNK M, LINDEN D, MUCKLI L, LANFERMANN H, ZANELLA F, SINGER W and GOEBEL R. Distributed cortical systems in visual short-term memory revealed by event-related functional magnetic resonance imaging. *Cerebral Cortex, 12:* 866-876, 2002.

NYSTROM LE, BRAVER TS, SABB FW, DELGADO MR, NOLL DC and COHEN JD. Working memory for letters, shapes, and locations: fMRI evidence against stimulus-based regional organization in human prefrontal cortex. *NeuroImage, 11:* 424-446, 2000.

O'SCALAIDHE SP, WILSON FA and GOLDMAN-RAKIC PS. Face-selective neurons during passive viewing and working memory performance of rhesus monkeys: Evidence for intrinsic specialization of neuronal coding. *Cerebral Cortex, 9:* 459-475, 1999.

OWEN AM, EVANS AC and PETRIDES M. Evidence for a two-stage model of spatial working memory processing within the lateral frontal cortex: A positron emission tomography study. *Cerebral Cortex, 6:* 31-38, 1996.

OWEN AM, STERN CE, LOOK RB, TRACEY I, ROSEN BR and PETRIDES M. Functional organization of spatial and nonspatial working memory processing within the human lateral frontal cortex. *Proceedings of the National Academy of Sciences of the USA, 95:* 7721-7726, 1998.

PESSOA L, GUTIERREZ E, BANDETTINI PA and UNGERLEIDER LG. Neural correlates of visual working memory: fMRI amplitude predicts task performance. *Neuron, 35:* 975-987, 2002.

PETRIDES M. Functional organization of the human frontal cortex for mnemonic processing. Evidence from neuroimaging studies. *Annals of the New York Academy of Sciences, 769:* 85-96, 1995a.

PETRIDES M. Impairments on nonspatial self-ordered and externally ordered working memory tasks after lesions of the mid-dorsal part of the lateral frontal cortex in the monkey. *Journal of Neuroscience, 15:* 359-375, 1995b.

PETRIDES M. Specialized systems for the processing of mnemonic information within the primate frontal cortex. *Philosophical Transactions of the Royal Society of London B, 351:* 1455-1461 (discussion 1461-1462), 1996.

PETRIDES M. The role of the mid-dorsolateral prefrontal cortex in working memory. *Experimental Brain Research, 133:* 44-54, 2000.

POSTLE BR and D'ESPOSITO M. "What" – then – "where" in visual working memory: An event-related fMRI study. *Journal of Cognitive Neuroscience, 11:* 585-597, 1999.

POSTLE BR, STERN CE, ROSEN BR and CORKIN S. An fMRI investigation of cortical contributions to spatial and nonspatial visual working memory. *NeuroImage, 11:* 409-423, 2000a.

POSTLE BR, ZARAHN E and D'ESPOSITO M. Using event-related fMRI to assess delay-period activity during performance of spatial and nonspatial working memory tasks. *Brain Research Protocols, 5:* 57-66, 2000b.

PRABHAKARAN V, NARAYANAN K, ZHAO Z and GABRIELI JD. Integration of diverse information in working memory within the frontal lobe. *Nature Neuroscience, 3:* 85-90, 2000.

QUINTANA J and FUSTER JM. Spatial and temporal factors in the role of prefrontal and parietal cortex in visuomotor integration. *Cerebral Cortex, 3:* 122-132, 1993.

QUINTANA J, YAJEYA J and FUSTER JM. Prefrontal representation of stimulus attributes during delay tasks. I. Unit activity in cross-temporal integration of sensory and sensory-motor information. *Brain Research, 474:* 211-221, 1988.

RAINER G, ASAAD WF and MILLER EK. Memory fields of neurons in the primate prefrontal cortex. *Proceedings of the National Academy of Sciences of the USA, 95:* 15008-15013, 1998.

RAMA P, POREMBA A, SALA JB, YEE L, MALLOY M, MISHKIN M and COURTNEY SM. Dissociable functional cortical topographies for working memory maintenance of voice identity and location. *Cerebral Cortex, 14:* 768-780, 2004.

RAO SC, RAINER G and MILLER EK. Integration of what and where in the primate prefrontal cortex. *Science, 276:* 821-824, 1997.

REISBERG D and LOGIE RH. The ins and outs of working memory: Overcoming the limits on learning from imagery. In Roskos-Ewoldsen B and Intons-Peterson MJ (Eds), *Imagery, Creativity, and Discovery: A Cognitive Perspective.* Amsterdam: North-Holland/Elsevier Science Publishers, 1993.

REYNOLDS JH and DESIMONE R. The role of neural mechanisms of attention in solving the binding problem. *Neuron, 24:* 19-29 (111-125), 1999.

SAKAI K and PASSINGHAM RE. Prefrontal interactions reflect future task operations. *Nature Neuroscience, 6:* 75-81, 2003.

SALA JB. *Active Maintenance and the Binding of Information during Working Memory.* A dissertation published in conformity with the requirements for the degree of Doctor of Philosophy, Johns Hopkins University, Baltimore, MD, 2003.

SALA JB and COURTNEY SM. Working memory maintenance reflects task goals and guides selective attention. *Visual Cognition* (submitted).

SALA JB, RAMA P and COURTNEY SM. Functional topography of a distributed neural system for spatial and nonspatial information maintenance in working memory. *Neuropsychologia, 41:* 341-356, 2003.

SALWAY AF and LOGIE RH. Visuospatial working memory, movement control and executive demands. *British Journal of Psychology, 86:* 253-269, 1995.

SMITH EE and JONIDES J. Working memory: A view from neuroimaging. *Cognitive Psychology, 33:* 5-42, 1997.

SMITH EE and JONIDES J. Storage and executive processes in the frontal lobes. *Science, 283:* 1657-1661, 1999.

STERN CE, OWEN AM, TRACEY I, LOOK RB, ROSEN BR and PETRIDES M. Activity in ventrolateral and mid-dorsolateral prefrontal cortex during nonspatial visual working memory processing: Evidence from functional magnetic resonance imaging. *NeuroImage, 11:* 392-399, 2000.

TAGAMETS MA and HORWITZ B. A model of working memory: Bridging the gap between electrophysiology and human brain imaging. *Neural Networks, 13:* 941-952, 2000.

VOGEL EK, WOODMAN GF and LUCK SJ. Storage of features, conjunctions and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 27:* 92-114, 2001.

WARD BD. Simultaneous inference for fMRI data. *http://afni.nimh.nih.gov/afni/docpdf/AlphaSim.pdf*, 2000.

WARD BD. Deconvolution analysis of fMRI time series data. *AFNI 3dDeconvolve Documentation, Medical College of Wisconsin*, 2002.

WHEELER ME and TREISMAN AM. Binding in short-term visual memory. *Journal of Experimental Psychology: General, 131:* 48-64, 2002.

WHITE IM and WISE SP. Rule-dependent neuronal activity in the prefrontal cortex. *Experimental Brain Research, 126:* 315-335, 1999.

WILSON FA, SCALAIDHE SP and GOLDMAN-RAKIC PS. Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science, 260:* 1955-1958, 1993.

YAJEYA J, QUINTANA J and FUSTER JM. Prefrontal representation of stimulus attributes during delay tasks. II. The role of behavioral significance. *Brain Research, 474:* 222-230, 1988.

ZIPSER D, KEHOE B, LITTLEWORT G and FUSTER JM. A spiking network model of short-term active memory. *Journal of Neuroscience, 13:* 3406-3420, 1993.

Susan M. Courtney, Department of Psychological and Brain Sciences, Ames Hall, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218-2686, USA. e-mail: Courtney@jhu.edu