Title

Parietal and Frontal Cortex Encode Stimulus-Specific Mnemonic Representations during Visual Working Memory

Journal

Neuron, 87(4)

ISSN

0896-6273

Authors

Ester, Edward F
Sprague, Thomas C
Serences, John T

Publication Date

2015-08-01

DOI

10.1016/j.neuron.2015.07.013

Peer reviewed

HHS Public Access

Author manuscript

Neuron. Author manuscript; available in PMC 2016 August 19.

Published in final edited form as:

Neuron. 2015 August 19; 87(4): 893–905. doi:10.1016/j.neuron.2015.07.013.

Parietal and Frontal Cortex Encode Stimulus-Specific Mnemonic Representations during Visual Working Memory

Edward F. Ester[1,*], **Thomas C. Sprague**[2], and **John T. Serences**[1,2,*]

[1]Department of Psychology, University of California, San Diego, La Jolla, CA, 92093, USA

[2]Neurosciences Graduate Program, University of California, San Diego, La Jolla, CA, 92093, USA

Summary

Working memory (WM) enables the storage and manipulation of information in an active state. WM storage has long been associated with sustained increases in activation across a network of frontal and parietal cortical regions. However, recent evidence suggests that these regions primarily encode information related to general task goals rather than feature-selective representations of specific memoranda. These goal-related representations are thought to provide top-down feedback that coordinates the representation of fine-grained details in early sensory areas. Here, we test this model using fMRI-based reconstructions of remembered visual details from region-level activation patterns. We could reconstruct high-fidelity representations of a remembered orientation based on activation patterns in occipital visual cortex and in several sub-regions of frontal and parietal cortex, independent of sustained increases in mean activation. These results challenge models of WM that postulate disjoint frontoparietal "top-down control" and posterior sensory "feature storage" networks.

Introduction

Visual working memory (WM) enables the representation and manipulation of information over short temporal intervals. This system is critical for bridging temporal gaps in visual processing that arise due to eye movements, occlusion, or the physical removal of stimuli from the visual field (Irwin, 1991; Hollingworth et al., 2008), and individual variability in WM ability is strongly correlated with general cognitive aptitudes such as IQ (Engle et al., 1999). Single-unit recordings in non-human primates suggest that WM is mediated by a broad network of frontal and parietal cortical regions. For example, many neurons in subregions of frontal and parietal cortex show elevated responses during tasks requiring the active storage of feature- or stimulus-specific visual information (e.g., Miller et al., 1996; Bisley and Pasternak, 2000; Mendoza-Halliday et al., 2014) or spatial information (Fuster and Alexander, 1971; Funahashi et al., 1989). Qualitatively similar results have been

*Correspondence: eester@ucsd.edu (E.F.E.), jserences@ucsd.edu (J.T.S.).

Supplemental Information: Supplemental Information includes Supplemental Experimental Procedures, five figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.neuron.2015.07.013.

Author Contributions: E.F.E., T.C.S., and J.T.S. conceived and designed the experiment. E.F.E. collected and analyzed the data. E.F.E., T.C.S., and J.T.S. wrote the manuscript.

obtained in humans using non-invasive neuroimaging techniques such as fMRI (e.g., Courtney et al., 1997; Pessoa et al., 2002; Srimal and Curtis, 2008). These sustained increases in activity (or activation) are regarded as a defining characteristic of cortical regions that support WM.

More recent human neuroimaging studies have used multivariate analyses to successfully decode simple visual features or spatial positions held in WM from delay-period multivoxel activation patterns in regions of posterior occipital and parietal cortex (see Serences et al., 2009; Harrison and Tong, 2009; Ester et al., 2009; Riggall and Postle, 2012; Emrich et al., 2013; Christophel et al., 2012, 2015; Jerde et al., 2012; Saber et al., 2015). Importantly, sustained activity changes can be dissociated from information storage during WM, as decoding is often successful even though the amplitude of the blood-oxygenation-level-dependent (BOLD) response typically returns to baseline levels during the memory delay period (Serences et al., 2009; Harrison and Tong, 2009; Riggall and Postle, 2012; Emrich et al., 2013). Additional studies have extended this work by using inverted encoding models (IEMs; Brouwer and Heeger 2009, 2011; see Sprague et al., 2015) to recover representations of remembered features based on delay period activation patterns within retinotopically organized occipital and posterior parietal cortex (Ester et al., 2013; Sprague et al., 2014).

Although there is a broad consensus that WM storage is mediated by a broad network of frontoparietal and sensory cortical areas, there is active debate about the general functional role(s) of these regions. According to one account, sustained increases in neural activity within frontoparietal cortical regions such as dorsolateral prefrontal cortex (dlPFC), superior precentral sulcus (sPCS; the putative human homolog of the macaque frontal eye fields), and portions of intra- and lateral parietal cortex encode representations of task-general information (e.g., which class of stimulus needs to be remembered, stimulus-response mappings, decision criteria, etc.). These task-general representations, in turn, are thought to coordinate highly detailed feature-specific representations in posterior sensory regions via top-down feedback (e.g., Sreenivasan et al., 2014a; D'Esposito and Postle, 2015). This model is supported by studies suggesting that single-unit and population-level responses in subregions of frontal and parietal cortex encode task-level variables such as rules (Warden and Miller, 2010; Riggall and Postle, 2012; Lee et al., 2013), category membership (Freedman et al., 2001), and stimulus-response mappings (Rowe et al., 2008). However, other studies have also demonstrated stimulus-specific responses in prefrontal and parietal regions during WM (e.g., Miller et al., 1996; Mendoza-Halliday et al., 2014), and recent evidence indicates that population-level responses in subregions of frontal and parietal cortex can encode both task-general and feature-specific representations in a high-dimensional and dynamic state space (Mante et al., 2013; Rigotti et al., 2013; Raposo et al., 2014; Stokes et al., 2013). Thus, sustained activity changes that are typically seen in frontoparietal cortex during WM might reflect the representation of task-related and feature-specific information.

Here, we tested this hypothesis by examining the information content of delay-period multivoxel fMRI activation patterns across all of human cortex. Participants were asked to remember the orientation of a peripheral grating across a 10 s delay period. We used an inverted encoding model to quantify representations of the remembered grating based on

delay period activation patterns in multiple retinotopically organized subregions of visual (V1-hV4v/V3a) and posterior parietal cortex (IPS0-3). Next, we used a traditional univariate analysis to identify subregions of frontoparietal cortex that showed a sustained increase in activation during the delay period, long considered a defining characteristic of regions that support WM. Inverted encoding models revealed robust representations of remembered orientation in a subset of these regions. Finally, we combined an IEM with a roving "searchlight" analysis (Kriegeskorte et al., 2006) to examine the information content of local activation patterns across the entire cortical sheet. This analysis revealed robust representations of the remembered orientation across a broad network of posterior visual and parietal regions, as well as portions of dlPFC, and ventral lateral prefrontal cortex (vlPFC). Collectively, these results show that representations of remembered visual features are encoded in both posterior and frontal cortex and challenge models of WM that postulate completely disjoint frontoparietal "top-down control" and posterior sensory "feature storage" networks.

Results

We collected BOLD fMRI data (see Supplemental Experimental Procedures) while six volunteers performed a delayed orientation recall task (Figure 1A and Experimental Procedures; each subject participated in 2–3 scanning sessions). Each trial began with the presentation of two "sample" gratings. Participants were subsequently cued to remember one of the two gratings over a 10 s blank delay (indicated by the green half of fixation diamond). Hereafter we will refer to the cued and non-cued gratings as the remembered and non-remembered orientations, respectively. Participants then adjusted the orientation of a probe grating to match the orientation of the remembered orientation as precisely as possible. The initial orientation of the probe grating was randomized with respect to the remembered orientation on every trial to ensure that participants could not anticipate the direction or the magnitude of the required rotation. Memory performance was quantified as the mean absolute value of the angular difference between the reported and actual stimulus orientations. Average recall error across participants (±1 SEM) was 14.63° (±2.86°). Each participant's recall error distribution was clustered around 0°, confirming that participants were storing accurate representations of the remembered grating (Figure 1B).

To assess feature-selective responses during WM, we used an inverted orientation-encoding model (IEM) to reconstruct representations of the remembered and non-remembered gratings based on activation patterns in several cortical regions of interest (ROIs; Experimental Procedures). For each ROI, we first divided the data into two sets—one used to train the model (the training set), and one used to compute a reconstruction of the remembered orientation (the test set). In the first phase of the analysis, delay period responses in each voxel measured during training blocks (Figure 2A) were modeled as a weighted sum of nine orientation-selective channels (Figure 2B), resulting in a matrix of weights that characterize the contribution of each orientation channel to the response of each voxel (Figure 2C). In the second phase of the analysis, we estimated the response of each orientation channel by combining these weights with the delay-period multivoxel activation patterns from each trial in the test set. This procedure yields a reconstructed representation of the remembered orientation on each trial (Figure 2D). We circularly shifted these trial-by-

trial reconstructions to a common orientation (0°) and averaged them to generate a single reconstructed representation. If delay period activation patterns within a ROI carry information about the remembered orientation, then the IEM should reveal a graded response function with a clear peak. Conversely, if these activation patterns do not represent the remembered orientation, then the IEM should reveal a flat response function. Note that this method converts BOLD activation patterns measured in voxel space back into stimulus space and can be conceptualized as a form of targeted dimensionality reduction that isolates orientation-specific representations from representations of other task-relevant and task-irrelevant factors.

Reconstructions of Orientation in Retinotopically Organized Visual and Posterior Parietal Cortex

Previous studies have successfully decoded and/or reconstructed representations of remembered features based on activation patterns in occipital cortex (V1-hV4v/V3a; Serences et al., 2009; Harrison and Tong, 2009; Riggall and Postle, 2012; Ester et al., 2013; Emrich et al., 2013) and posterior parietal cortex (Christophel et al., 2012, 2015). Consequently, we first attempted to reconstruct representations of the remembered and non-remembered orientations within these regions. Reconstructions were computed separately for each visual area (e.g., V1, V2, etc.) and posterior parietal subregion (e.g., IPS0, IPS1, etc.). We also accounted for the retinotopic location of each ROI with respect to the remembered orientation (i.e., contralateral versus ipsilateral) as prior work has revealed spatially global representations of simple features within subregions of visual cortex during WM (Ester et al., 2009; Pratte and Tong, 2014). To generate the plots shown in Figures S1 and S2, we averaged reconstructions within each subregion across all scan sessions (n = 2 or 3 for each participant) and then averaged across participants. Prospective differences between reconstructions of the remembered and non-remembered orientations were then evaluated using a bootstrapping procedure across participants and sessions (see Quantification and Comparison of Reconstructed Representations in Experimental Procedures for details).

Reconstructions of the remembered and non-remembered orientation from visual areas V1-hV4v/V3a are plotted as a function of retinotopic location (i.e., contralateral or ipsilateral relative to the location of the remembered or non-remembered orientation) in Figure S1. Analogous data from IPS subregions 0–3 are plotted in Figure S2. In contralateral V1, we observed robust representations of the remembered orientation (permutation test; $p < 0.001$) but not of the non-remembered orientation ($p = 0.663$). Moreover, reconstruction amplitudes were reliably higher for the remembered relative to the non-remembered orientation ($p = 0.002$). There was also a significant delay-period representation of the remembered item in ipsilateral V1 ($p < 0.01$), but it was not significantly different from the representation of the non-remembered item ($p = 0.191$).

We also evaluated reconstructions averaged across visual areas V1-hV4v/V3a (averaging was performed separately for each participant, scan session, and location, i.e., contralateral or ipsilateral). In line with the general pattern observed in V1, there was a robust representation in contralateral visual areas ($p = 0.029$) and a trend toward a robust delay-

period representation of the remembered orientation in ipsilateral visual areas ($p = 0.095$). We could not recover a significant representation of the non-remembered orientation in either contralateral or ipsilateral visual cortex (both p values > 0.22), and overall representations of the remembered orientation were more robust than representations of the non-remembered orientation in contralateral visual areas ($p = 0.037$), but not in ipsilateral areas ($p = 0.285$). Contralateral and ipsilateral representations of the remembered orientation were statistically indistinguishable ($p = 0.311$). Statistics for all the individual areas are reported in Table S1. Collectively, these findings replicate earlier work (Ester et al., 2009; Pratte and Tong, 2014) and suggest that voxel activation patterns in visual cortex encode a spatially global representation of the remembered orientation during WM.

Next, we examined representations of the remembered and non-remembered orientations averaged across IPS subregions 0–3 (see Figure S2). We observed a robust representation of the remembered orientation in ipsilateral IPS (permutation test; $p = 0.005$), but not contralateral IPS ($p = 0.187$). We were unable to recover representations of the non-remembered orientation in any IPS subregion (all p values > 0.10). In addition, representations of the non-remembered orientation were statistically indistinguishable from representations of the remembered orientation in contralateral IPS ($p = 0.601$), while representations of the remembered orientation were marginally stronger than representations of the non-remembered orientation in ipsilateral IPS ($p = 0.112$). Finally, there was a modest trend toward stronger representations of the remembered orientation in ipsilateral relative to contralateral IPS (permutation test; $p = 0.062$). Statistics for each IPS subregion can be found in Table S1. Collectively, these results are consistent with earlier studies documenting stimulus-specific representations in posterior IPS (e.g., Christophel et al., 2012, 2015).

Feature-Selective Activation Patterns in Regions with Elevated Delay Period Activation

Recent studies have documented an apparent dissociation between the univariate response amplitude and the feature-selective information content of BOLD activation patterns during WM storage. For example, several studies indicate that while specific features of a remembered stimulus (e.g., motion direction, orientation, color) can be successfully decoded using activation in posterior visual areas that do not show an elevated mean response during WM, feature information cannot not be decoded in subregions of frontoparietal cortex that do show an elevated mean response during WM (Riggall and Postle, 2012; Emrich et al., 2013; Lee et al., 2013; Sreenivasan et al., 2014b). To examine whether feature-selective representations might also be present in these frontoparietal areas, we next examined multivoxel activation patterns related to remembered and non-remembered orientations in cortical areas with elevated delay period activation. Following earlier work (e.g., Zarahn et al., 1997; Riggall and Postle, 2012), we identified regions with elevated delay period activation with a random-effects general linear model (GLM) that included separate regressors marking the sample, delay, and probe epochs (see Experimental Procedures). A statistical parametric map (SPM) showing cortical areas with elevated delay period activity is shown in Figure 3. From this analysis, we identified a set of 14 ROIs with elevated delay period activity, including bilateral portions of lateral and medial frontal cortex, superior parietal lobule, and lateral occipioparietal cortex (Figure 3; Table 1).

Figure 4A shows event-related averaged BOLD responses from a subset of four representative ROIs with elevated delay period activation (Figure S3A shows responses from the remaining ROIs). Next, we attempted to classify the remembered and non-remembered orientations by applying a support vector machine (SVM) to delay-period activation patterns measured in each ROI (see Supplemental Experimental Procedures). We made no assumptions regarding the retinotopic organization of voxels within each ROI; consequently data were sorted based only on the remembered or non-remembered orientations, irrespective of their locations (i.e., left or right visual field). As shown in Figures 4B and S3B, decoding accuracy rarely exceeded chance levels at either the group or single participant levels. These results are consistent with previously reported dissociations between the amplitude and feature-selective information content of the BOLD response in many frontoparietal ROIs (Riggall and Postle, 2012; Emrich et al., 2013; Lee et al., 2013; Sreenivasan et al., 2014b).

To provide a more direct test of the hypothesis that frontoparietal regions encode representations of the remembered orientation during WM, we applied an IEM to delay-period activation patterns in ROIs with elevated delay period responses. Unlike multivariate decoding analyses, this method maps BOLD activation patterns in voxel space into a pre-defined information space that specifies how one or more stimulus variables (in this case, orientation) might be encoded across a population of voxels. This step can be conceptualized as a form of targeted dimensionality reduction that may help to disentangle weak or sparsely distributed feature-specific representations from representations of other task-relevant factors.

Figures 4C and S3C plot reconstructions of the remembered and non-remembered orientations in each ROI that showed elevated BOLD activation during the WM delay. Robust representations of the remembered orientation were observed in a subset of these ROIs, including left superior precentral sulcus (sPCS), bilateral superior parietal lobule (SPL), and right parieto-occipital cortex (Figures 4 and S3; Table 1). However, other regions did not contain robust representations of the remembered or non-remembered orientations. We also computed reconstructed representations of the remembered and non-remembered orientations for each fMRI image obtained during the WM delay (Figures 4D and S3D). In regions containing a robust representation of the remembered orientation (e.g., left sPCS and bilateral SPL), representations of the remembered orientation appeared shortly after the offset of the sample display and persisted until the presentation of the probe display. Conversely, representations of the non-remembered orientation were observed early during the delay period (e.g., samples acquired 2 or 4 s after the start of the trial), but were absent at later samples. This result suggests that participants completed the behavioral task by initially encoding representations of both the remembered- and non-remembered orientations, then purging the representation of the non-remembered orientation following the onset of the postcue.

Collectively, the results shown in Figures 4C and 4D suggest that at least some frontoparietal cortical regions with elevated delay period activity represent elementary feature properties during WM. Next, we asked whether representations encoded by these regions are categorical or continuous in nature. Although the graded shape of the

reconstructions shown in Figures 4C and S3C are nominally consistent with a continuous representation, they were generated using a basis set of nine overlapping sinusoids (Figure 2B). This overlap ensures the responses of neighboring points along each curve will be correlated and will confer smoothness to the reconstructions even if the underlying feature representation is categorical. We therefore recomputed reconstructions of the remembered orientation using a basis set containing nine orthogonal Kronecker delta functions, where each function was centered on one of the nine possible remembered orientation values (see Saproo and Serences, 2014). If representations of the remembered orientation are categorical, then we should recover a representation with a sharp peak at the remembered orientation and a uniformly small response to all other orientations. Conversely, if the representations are continuous, then we should observe a graded response function similar to those shown in Figures 4C and S3C.

Figures 4E and S3E plot reconstructed representations of the remembered orientation obtained using this method in delay period ROIs. In regions containing a robust representation of the remembered orientation (e.g., bilateral SPL and left sPCS), reconstructions peaked at the remembered orientation and gradually decreased with the angular distance from this orientation. This result suggests that representations of the remembered orientation are continuous rather than categorical.

Whole-Brain Identification of Feature-Selective WM Representations

Finally, we combined the IEM approach used in previous sections with a roving searchlight analysis (e.g., Kriegeskorte et al., 2006) to identify cortical regions representing the remembered orientation irrespective of changes in delay period activation. We first defined a spherical neighborhood with an 8 mm radius around each voxel in the cortical sheet. Voxels within each neighborhood were used to compute a reconstruction of the remembered orientation, and the amplitude of each reconstructed representation was estimated by fitting the reconstructed channel response function with an exponentiated cosine function (Equation 4). We then generated a separate SPM for each subject that marked clusters representing the remembered orientation by submitting neighborhood-level reconstruction amplitude estimates from the remaining five participants to a one-tailed t test against zero (see Figure 5A). Although this “hold-one-participant-out” approach yields a unique set of clusters for each participant, it avoids circularity by ensuring that reconstructions of the remembered and non-remembered orientations remain statistically independent from the criteria used to define these clusters. Finally, we projected each participant's SPM onto a computationally inflated representation of his or her gray-white matter boundary (Figure 5B). Clusters containing a robust representation of the remembered orientation ($t_{critical}$ = 2.778, with 4 degrees of freedom and $p = 0.05$, one-tailed and uncorrected for multiple comparisons in order to maximize sensitivity) were retained for subsequent analyses.

We observed robust representations of the remembered orientation across a broad network of cortical areas, including subregions of retinotopically organized visual and posterior parietal cortex, lateral occipital cortex (LOC) and bilateral dorsolateral prefrontal cortex (dlPFC; Figure 5B). Next, we identified clusters supporting a robust reconstruction of the remembered orientations located near three broad PFC cortical areas: left dlPFC, right

dlPFC, and left vlPFC. As shown in Figure 5, SPMs for each of our participants had at least one significant cluster in left dlPFC, while SPMs for four and five of our six participants had at least one cluster located near left vlPFC and right dlPFC, respectively. We also observed significant clusters located near anterior portions of medial prefrontal cortex, but these were only present in the SPMs for two or three participants (e.g., participants AB and AC; Figure 5B). Next, we combined clusters within the same general anatomical location (e.g., left dlPFC) to generate a set of three ROIs located in left dlPFC, right dlPFC, and left vlPFC. Although the precise location(s) of these ROIs varied across participants, they were generally located less than a few millimeters apart (see Table 2). We then extracted multivoxel activation patterns from each cluster located near left and right dlPFC and left vlPFC (separately for each participant).

Figure 6A plots event-related average BOLD responses time locked to the onset of the sample display in each ROI. In left dlPFC there was an initial transient response to the sample display, followed by a sustained lower-amplitude response that persisted until the onset of the probe display. This pattern was reminiscent of many regions demonstrating elevated delay period activation (Figure 4A). However, we observed no changes in the amplitude of either right dlPFC or left vlPFC. Regardless of overall changes in the BOLD response, we were unable to classify the identities of either the remembered or non-remembered orientations from delay period activation patterns in any ROI, replicating the general pattern seen in ROIs with elevated delay period activation (Figure 4B). However, an IEM revealed a robust representation of the remembered orientation in right dlPFC ($p = 0.013$) and left vlPFC ($p = 0.001$), but not left dlPFC ($p = 0.231$). Conversely, we could not recover a representation of the non-remembered orientation using activation patterns from any of these areas (all p values > 0.62). Representations of the remembered orientation in left vlPFC and right dlPFC were also reliably stronger than representations of the non-remembered orientation ($p = 0.018$ and 0.008, respectively). Figure 6D plots time-resolved reconstructions of the remembered and non-remembered orientations. With the exception of right vlPFC, representations of the remembered orientation emerged shortly after the onset of the sample display and persisted for the majority of the trial.

Finally, we reconstructed representations of the remembered and non-remembered orientations in each ROI using a basis set of delta functions. As shown in Figure 6E, reconstructions of the remembered orientation peaked at the remembered orientation and decreased gradually with angular distance from this value in in both right dlPFC and left vlPFC (a similar trend was also observed in left dlPFC, but this region did not contain a robust representation of the remembered orientation), consistent with a continuous rather than categorical representation.

Discussion

Recent models of WM postulate that storage is mediated by the coordination of neural activity in largely separable frontoparietal and posterior sensory cortical networks. According to one influential model (e.g., Sreenivasan et al., 2014a; D'Esposito and Postle, 2015), subregions of frontoparietal cortex encode representations of task-relevant factors (e.g., task sets and stimulus-response mappings) rather than feature-selective information.

These representations, in turn, serve to coordinate the creation and maintenance of stimulus- or feature-specific representations in posterior sensory areas. Evidence supporting this model comes primarily from studies suggesting that it is possible to decode the attributes of a remembered stimulus in posterior sensory cortex during WM, but not in frontoparietal regions that show elevated delay-period activation (e.g., Riggall and Postle, 2012; Emrich et al., 2013; Lee et al., 2013; Sreenivasan et al., 2014b).

The present data challenge this framework by demonstrating that representations of a remembered feature are distributed throughout the cortical hierarchy, including many retinotopically organized regions of visual and posterior parietal cortex (Figures S1 and S2), and subregions of frontoparietal cortex defined by elevated delay-period activation (Figures 3, 4, and S3) or local information content as indexed by a multivariate searchlight analysis (Figures 5 and 6). Collectively, these results suggest that frontoparietal cortical areas contribute to WM storage by both actively representing task-relevant information (e.g., Figures 3, 4, 5, and 6) and coordinating the representation of this information and/or modulating responses to incoming sensory signals in posterior sensory cortex via top-down feedback signals. Our data also reveal a very nuanced pattern of function across sub-regions of frontoparietal cortex: some regions show only sustained delay period activation (e.g., left middle frontal gyrus; Figure 3 and Table 1), some show only feature-selective response patterns (e.g., left ventrolateral and right dorsolateral prefrontal cortex; Figure 6), and other regions show both sustained activation and feature selectivity (e.g., right superior parietal lobule; Figure 4). The co-existence of these three response patterns suggests the possibility of distinct functional networks that operate to jointly mediate both top-down cognitive control, broadly construed, as well as the maintenance of feature-specific information about currently relevant stimuli.

Although we report robust representations of a remembered feature in several subregions of PFC (see Figures 4 and 6), there is ample evidence suggesting that many of these subregions also modulate perceptual and WM representations in posterior sensory cortical areas. For example, a recent transcranial magnetic stimulation (TMS) study in humans demonstrated that stimulating the lateral prefrontal cortex during the encoding period of a WM task modulates the selectivity of responses in visual cortex during a subsequent memory delay (Lee and D'Esposito, 2012). In related work, Ekstrom et al. (2008) demonstrated that stimulating the frontal eye fields during a challenging perceptual task has a systematic effect on exogenously driven responses in early visual cortex of macaque monkeys (see also Ruff et al., 2006; Moore and Armstrong, 2003). Finally, patient studies suggest that recognition and recall performance on tasks requiring participants to manipulate information held in WM (e.g., recalling a list of words or digits in reverse order; D'Esposito and Postle, 1999) or suppress distracting information during storage (Chao and Knight, 1998) are impaired following lesions to PFC. These results imply that PFC plays an integral role in controlling access to WM (see also Miller et al., 1996; McNab and Klingberg, 2008). Our results also suggest that some feature-specific information about WM representations is also encoded in many posterior cortical regions, including subregions of retinotopically organized visual cortex and posterior parietal cortex (see Figures S1 and S2 and Figure 5). This distributed code may at least partially explain why WM storage is largely unaffected following lesions to PFC.

How Are WM Representations Encoded?

In a recent study, Mendoza-Halliday et al. (2014) recorded from three interconnected cortical regions implicated in motion processing in macaque monkeys—the middle temporal area (MT), the medial superior temporal area (MST), and the lateral prefrontal cortex (lPFC)—while monkeys remembered the direction of a moving dot stimulus over a short delay. Large increases in spiking activity over the delay interval that encoded the remembered motion direction were observed in areas MST and lPFC, but not area MT. Conversely, these authors observed sustained direction-selective changes in MT local field potential power (LFP; particularly for low frequencies in the theta, alpha, and beta bands) as well as robust spike-field coherence between spikes recorded from lPFC and MT LFP power in the β band. Given these results, Mendoza-Halliday et al. proposed that feedback signals generated in MST or lPFC modulate subthreshold synaptic activity in MT, thereby biasing responses to subsequent sensory inputs. This proposal is broadly consistent with a model of WM in which memoranda are initially encoded by transient spiking activity in posterior sensory cortex and stored by sustained spiking activity in anterior association regions, including lPFC. This sustained activity also acts as a "top-down" feedback mechanism that induces subthreshold changes in the activity of visual cortical neurons in a manner that biases responses to further sensory input.

Why, then, have human neuroimaging studies consistently failed to find stimulus- or feature-specific activation patterns in frontoparietal cortex during WM? Critically, the model proposed by Mendoza-Halliday et al. (2014) hinges on the assumption that WM representations are encoded primarily by sustained changes in patterns of spiking activity. However, WM representations might also be encoded by "subthreshold" changes in neural membrane potentials below the spiking threshold or other neural properties that are not reflected in action potentials (Stokes, 2015). For example, a recent theoretical paper suggests that WM representations could be sustained by changes in synaptic weights within a recurrent neural network that could be read out by a sweep of spiking activity (e.g., Mongillo et al., 2008). A similar principle might hold for changes in sub-threshold membrane potentials, which, once elevated, reduce the input required to produce spikes. Assuming that these non-spiking response properties can be detected in large-scale neural activity measures such as the LFP, and given known links between the LFP and the BOLD signal (Logothetis et al., 2001; Magri et al., 2012), it is plausible that neuroimaging methods may be particularly useful in exploring networks that support WM via these mechanisms (see, for example, Boynton, 2011).

Conclusions

Multiple neuroimaging studies have identified feature- and stimulus-specific WM representations in visual and posterior parietal cortex (Serences et al., 2009; Harrison and Tong, 2009; Ester et al., 2013; Christophel et al., 2012; 2015), but not frontal and anterior parietal cortical areas (e.g., Riggall and Postle, 2012; Emrich et al., 2013; Lee et al., 2013; Sreenivasan et al., 2014b). In light of these findings, current models (e.g., Sreenivasan et al., 2014a; D'Esposito and Postle, 2015) propose that WM storage is mediated by the coordinated activity of two largely disjoint networks: a frontoparietal network that encodes task goals and abstract representations of memoranda, and a posterior sensory "feature

storage" network that enables the storage of detailed visual representations. Our findings instead suggest that feature-specific WM representations are encoded by a broadly distributed network of sensory and frontoparietal cortical areas. Representations of memoranda in frontoparietal cortical regions may be multiplexed with representations of other task-relevant information such as motor programs, stimulus-response mappings, and decision criteria (e.g., Mante et al., 2013; Rigotti et al., 2013; Stokes et al., 2013; Raposo et al., 2014), thereby enabling the flexible control of behavior in response to changing task demands.

Experimental Procedures

Participants

Six neurologically intact volunteers (3 females, mean age 26.83 years, all right handed) from the University of California San Diego community participated in two (n = 2) or three (n = 4) 2-hr experimental scanning sessions. One participant was author T.C.S. Each participant also completed a single 2-hr retinotopic mapping scan session; data from this session were used to define retinotopically organized regions of visual and posterior parietal cortex (see Retinotopic Mapping and ROI definition, Supplemental Experimental Procedures). Participants also completed a short (∼30 min) behavioral training session prior to being scanned in order to familiarize them with the WM task. All participants reported normal or corrected-to-normal and gave both written and oral informed consent as required by the local Institutional Review Board. Participants were compensated at a rate of $10/hr for behavioral training and $20/hr for scanning.

Orientation WM Task

Stimuli were generated in MATLAB using the Psychophysical Toolbox software package (Brainard, 1997; Pelli, 1997) and projected onto a 110 cm (width) display located at the base of the magnet bore. Participants viewed the display from a distance of approximately 370 cm via a mirror attached to the scanner's head coil. A representative trial of the task is depicted in Figure 1. Participants saw two "sample" gratings (radius 1.88°, 3 cycles/degree) to the left and right of a fixation diamond (width 0.6°) along the horizontal meridian (6.15° eccentricity). Each grating flickered at 3 Hz (i.e., 167 ms on, 167 ms off) for a total of 1,000 ms. Each sample grating was assigned an orientation drawn from a uniform distribution over 0°–160° in 20° increments, plus a small angular jitter (±1°–5°; randomly chosen on each trial). Immediately after offset, one half of the fixation diamond changed colors from black to green; this change served as a "post-cue" and indicated which of the two gratings was to be remembered over a subsequent 10 s blank delay. We refer to the postcued grating as the "remembered" grating and the non-postcued grating as the "non-remembered" grating. The delay period was followed by the presentation "probe" grating. The initial orientation of the probe grating was randomized with respect to the remembered orientation on each trial to prohibit anticipatory motor responses. Participants were instructed to adjust the orientation of the probe (using an MR-compatible button box) to match that of the remembered sample. Participants were given 3 s to adjust the probe, and the probe's orientation at the end of this interval was taken as the participant's final response. Trials were separated by a 4 or 6 s inter-trial interval (pseudorandomly chosen after each trial). Each experimental block

contained 18 trials and lasted 378 s. Stimulus orientations and locations (i.e., left or right visual field) were fully crossed within a single block of trials. Each participant completed nine blocks per scanning session, and the orientations of the remembered and non-remembered gratings were fully crossed across these nine blocks.

Orientation Encoding Model

A linear encoding model was used to characterize orientation-selective responses in each functionally defined ROI. This model rests on the assumptions that the measured response in a given voxel is an approximately linear sum of underlying neural activity, and that at least some of the voxels within a given ROI exhibit a non-uniform response profile across orientations (e.g., Brouwer and Heeger, 2009, 2011).

We began by modeling the response of each voxel within a given ROI as a linear sum of 9 information channels. Following Brouwer and Heeger (2009, 2011), we let B_1 (m voxels × n trials) be the observed signal in each voxel in each trial, C_1 (k channels × n trials) be a matrix of predicted responses for each information channel on each trial, and W (m voxels × k channels) be a weight matrix that characterizes the mapping from "channel space" to "voxel space." The relationship between B_1, C_1, and W can be described by a general linear model of the form:

$$B_1 = WC_1. \quad \text{(Equation 1)}$$

C_1 reflects the predicted response in each modeled information channel on each trial. For most analyses (Figures S1 and 2; panels C and D in Figures 4, 6, and S3), we generated a basis set containing nine half-wave rectified sinusoids centered at different orientations (0°, 20°, 40°, etc) and raised to the 8th power. These functions were chosen because they approximate the shape of single-unit tuning functions in V1, where the half-bandwidth of orientation-selective cells has been estimated to be approximately 20° (though there is substantial variability in bandwidth; see Ringach et al., 2002; Gur et al., 2005). We used these functions and the remembered orientation estimate the responses of each channel during WM. In other analyses (panel E of Figures 4, 6, and S4), we modeled the response of each information channel using a delta function centered at one of the 9 orientations used in the task, so that each column of C_1 was 1 at the relevant orientation for that trial and 0 elsewhere.

Given B_1 and C_1, we estimated the weight matrix $\hat{W}$ (m voxels × k channels) using ordinary least-squares regression:

$$\hat{\mathrm{W}} = B_1 C_1^T \left(C_1 C_1^T \right)^{-1}. \quad \text{(Equation 2)}$$

Given these weights and voxel responses observed in an independent "test" dataset, we invert the model to transform the observed test data B_2 (m voxels × n trials) into a set of estimated channel responses, C_2 (k channels × n trials):

$$C_2 = \left(\hat{W}^T \hat{W}\right)^{-1} \hat{W}^T B_2. \quad \text{(Equation 3)}$$

The estimated channel responses were circularly shifted to a common center (0°) and averaged across trials. To generate the smooth, 180-point functions shown in Figures 3 and 4, we repeated the encoding model analysis a total of 19 times and shifted the centers of the orientation channels by 1° on each iteration.

We implemented a "leave-one-out" cross-validation routine such that data from all but one experimental block acted as B_1 and were used to estimate $\hat{W}$, while data from the remaining scan acted as B_2 and were used to estimate C_2. This approach ensures that the data used to estimate the weight matrix $\hat{W}$ (B_1) and channel responses (B_2) were statistically independent. The entire analysis was repeated until all blocks within a given scanning session were held out as a test set, and the resulting channel responses were concatenated across trials. Channel response estimation was performed separately for each 2-hr experimental session, and the results were averaged across sessions.

With the exception of the sample-by-sample analyses shown in Figures 4, 6, S1, S2, and S3, all multivariate analyses were based on data averaged across three TRs beginning 8, 10, and 12 s after the start of each trial to account for hemodynamic lag. Similar findings were obtained when we used data from TRs beginning 6, 8, and 10 s after the start of each trial. Time-resolved reconstructions (Figures 4D, 6D, S1, S2, and S3D) were computed by applying an IEM to activation patterns measured at samples obtained 2–12 s after the onset of the sample display. Note that the average of sample-by-sample reconstructions obtained at 8, 10, and 12 s following the onset of the probe display need not match the reconstructions obtained by applying an IEM to activation patterns averaged across these samples (Figures 4C, 6C, S1, S2, and S3D).

Quantification and Comparison of Reconstructed Representations

Reconstructed representations of the remembered and non-remembered orientations were quantified using bootstrapping. For a given ROI, we began by computing representations of the remembered and non-remembered orientations separately for each participant (n = 6) and experimental session (n = 2 or 3). These reconstructions were stacked, yielding a set of two 16 (number of sessions across all participants) by 180 (smoothed orientation channels) data matrices. In retinotopically organized visual and posterior parietal areas, we created separate matrices for contralateral and ipsilateral representations. Next, we randomly sampled (with replacement) and averaged 16 rows from each matrix, yielding one representation of the remembered orientation and one representation of the non-remembered orientation. Each representation was fit with an exponentiated cosine function of the form:

$$f(x) = \alpha(e^{k(\cos(\mu - x) - 1)}) + \beta \quad \text{(Equation 4)}$$

where x is a vector of channel responses. μ, k, and β control the center (i.e., mean), concentration (i.e., inverse of width) and baseline (i.e., vertical offset) of the function, while α corresponds to the amplitude of the function (i.e., vertical stretching/scaling; signal above

a noisy baseline). We used the latter to define a measure of the robustness of the reconstructed representation. Fitting was performed by combining a general linear model with a grid search procedure (see Supplemental Experimental Procedures).

This entire procedure was repeated 2,500 times, yielding 2,500 amplitude estimates for representations of the remembered and non-remembered orientations. We determined whether a given ROI contained a significant representation of the remembered or non-remembered orientation by computing the proportion of resampled amplitude estimates for each stimulus that exceeded 0 ($p < 0.05$; one-tailed). We also examined whether amplitude estimates were higher for the remembered relative to the non-remembered orientations by computing the proportion of times that the difference between resampled amplitude estimates for the remembered and non-remembered orientations were less than 0.

Finally, because each scan was treated independently, participants who completed three scan sessions ($n = 4$) will exert a greater impact on the outcome of this analysis than those who completed two scan sessions ($n = 2$). We therefore confirmed that the effects reported here generalized when we excluded data from the final scan for each of the four participants who completed three sessions (see Table S2).

Searchlight Analysis

An IEM was combined with a roving “searchlight” procedure to identify regions representing the remembered orientation across the entire brain. For each participant, we first generated a cortical mask marking only gray matter voxels. We then defined a spherical “neighborhood” (radius 8.0 mm) centered on voxel in the mask. Neighborhoods containing fewer than 100 voxels (e.g., due to cortical folding patterns) were discarded, resulting in an average cluster size of 198 voxels (with a maximum size of 257 voxels). Within each of these neighborhoods, we used an IEM to estimate the responses of nine hypothetical orientation channels corresponding to the possible orientations of the remembered stimulus (see Figure 2) and fit the resulting reconstructions with the function described in Equation 4. Estimates of α obtained from fitting were then used to define a set of candidate ROIs for each participant via a “hold-one-participant-out” cross-validation routine. For each participant, submitted neighborhood-level reconstruction amplitude estimates from the remaining five participants to a one-tailed t test against a distribution with a mean of zero. Thus, for participant AA, we retained data from participants AB, AC, AI, AL, and AP, while for participant AC, we retained data from participant AA, AB, AI, AL, and AP (and so forth). We then generated a statistical parametric map (SPM) marking voxels whose amplitude estimates were reliably greater than 0 [$t(4) = 2.78$, $p < 0.05$, one-tailed]. We then projected each participant's SPM onto a computationally inflated representation of his or her gray-white matter boundary, and used BrainVoyager's “Create POIs from Map Clusters” function with an area threshold of 20 mm^2 to identify ROIs containing a robust representation of the remembered stimulus.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Supported by NIH R01 MH092345 and James S. McDonnell Foundation Scholar Award to J.T.S.

References

Bisley JW, Pasternak T. The multiple roles of visual cortical areas MT/MST in remembering the direction of visual motion. Cereb Cortex. 2000; 10:1053–1065. [PubMed: 11053227]

Boynton GM. Spikes, BOLD, attention, and awareness: a comparison of electrophysiological and fMRI signals in V1. J Vis. 2011; 11:12. http://dx.doi.org/10.1167/11.5.12.

Brainard DH. The psychophysics toolbox. Spat Vis. 1997; 10:433–436. [PubMed: 9176952]

Brouwer GJ, Heeger DJ. Decoding and reconstructing color from responses in human visual cortex. J Neurosci. 2009; 29:13992–14003. [PubMed: 19890009]

Brouwer GJ, Heeger DJ. Cross-orientation suppression in human visual cortex. J Neurophysiol. 2011; 106:2108–2119. [PubMed: 21775720]

Chao LL, Knight RT. Contribution of human prefrontal cortex to delay performance. J Cogn Neurosci. 1998; 10:167–177. [PubMed: 9555105]

Christophel TB, Hebart MN, Haynes JD. Decoding the contents of visual short-term memory from human visual and parietal cortex. J Neurosci. 2012; 32:12983–12989. [PubMed: 22993415]

Christophel TB, Cichy RM, Hebart MN, Haynes JD. Parietal and early visual cortices encode working memory content across mental transformations. Neuroimage. 2015; 106:198–206. [PubMed: 25463456]

Courtney SM, Ungerleider LG, Keil K, Haxby JV. Transient and sustained activity in a distributed neural system for human working memory. Nature. 1997; 386:608–611. [PubMed: 9121584]

D'Esposito M, Postle BR. The dependence of span and delayed-response performance on prefrontal cortex. Neuropsychologia. 1999; 37:1303–1315. [PubMed: 10530730]

D'Esposito M, Postle BR. The cognitive neuroscience of working memory. Annu Rev Psychol. 2015; 66:115–142. [PubMed: 25251486]

Ekstrom LB, Roelfsema PR, Arsenault JT, Bonmassar G, Vanduffel W. Bottom-up dependent gating of frontal signals in early visual cortex. Science. 2008; 321:414–417. [PubMed: 18635806]

Emrich SM, Riggall AC, Larocque JJ, Postle BR. Distributed patterns of activity in sensory cortex reflect the precision of multiple items maintained in visual short-term memory. J Neurosci. 2013; 33:6516–6523. [PubMed: 23575849]

Engle RW, Tuholski SW, Laughlin JE, Conway ARA. Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. J Exp Psychol Gen. 1999; 128:309–331. [PubMed: 10513398]

Ester EF, Serences JT, Awh E. Spatially global representations in human primary visual cortex during working memory maintenance. J Neurosci. 2009; 29:15258–15265. [PubMed: 19955378]

Ester EF, Anderson DE, Serences JT, Awh E. A neural measure of precision in visual working memory. J Cogn Neurosci. 2013; 25:754–761. [PubMed: 23469889]

Freedman DJ, Riesenhuber M, Poggio T, Miller EK. Categorical representation of visual stimuli in the primate prefrontal cortex. Science. 2001; 291:312–316. [PubMed: 11209083]

Funahashi S, Bruce CJ, Goldman-Rakic PS. Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. J Neurophysiol. 1989; 61:331–349. [PubMed: 2918358]

Fuster JM, Alexander GE. Neuron activity related to short-term memory. Science. 1971; 173:652–654. [PubMed: 4998337]

Gur M, Kagan I, Snodderly DM. Orientation and direction selectivity of neurons in V1 of alert monkeys: functional relationships and laminar distributions. Cereb Cortex. 2005; 15:1207–1221. [PubMed: 15616136]

Harrison SA, Tong F. Decoding reveals the contents of visual working memory in early visual areas. Nature. 2009; 458:632–635. [PubMed: 19225460]

Hollingworth A, Richard AM, Luck SJ. Understanding the function of visual short-term memory: transsaccadic memory, object correspondence, and gaze correction. J Exp Psychol Gen. 2008; 137:163–181. [PubMed: 18248135]

Irwin DE. Information integration across saccadic eye movements. Cognit Psychol. 1991; 23:420–456. [PubMed: 1884598]

Jerde TA, Merriam EP, Riggall AC, Hedges JH, Curtis CE. Prioritized maps of space in human frontoparietal cortex. J Neurosci. 2012; 32:17382–17390. [PubMed: 23197729]

Kriegeskorte N, Goebel R, Bandettini P. Information-based functional brain mapping. Proc Natl Acad Sci USA. 2006; 103:3863–3868. [PubMed: 16537458]

Lee TG, D'Esposito M. The dynamic nature of top-down signals originating from prefrontal cortex: a combined fMRI-TMS study. J Neurosci. 2012; 32:15458–15466. [PubMed: 23115183]

Lee SH, Kravitz DJ, Baker CI. Goal-dependent dissociation of visual and prefrontal cortices during working memory. Nat Neurosci. 2013; 16:997–999. [PubMed: 23817547]

Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A. Neurophysiological investigation of the basis of the fMRI signal. Nature. 2001; 412:150–157. [PubMed: 11449264]

Magri C, Schridde U, Murayama Y, Panzeri S, Logothetis NK. The amplitude and timing of the BOLD signal reflects the relationship between local field potential power at different frequencies. J Neurosci. 2012; 32:1395–1407. [PubMed: 22279224]

Mante V, Sussillo D, Shenoy KV, Newsome WT. Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature. 2013; 503:78–84. [PubMed: 24201281]

McNab F, Klingberg T. Prefrontal cortex and basal ganglia control access to working memory. Nat Neurosci. 2008; 11:103–107. [PubMed: 18066057]

Mendoza-Halliday D, Torres S, Martinez-Trujillo JC. Sharp emergence of feature-selective sustained activity along the dorsal visual pathway. Nat Neurosci. 2014; 17:1255–1262. [PubMed: 25108910]

Miller EK, Erickson CA, Desimone R. Neural mechanisms of visual working memory in prefrontal cortex of the macaque. J Neurosci. 1996; 16:5154–5167. [PubMed: 8756444]

Mongillo G, Barak O, Tsodyks M. Synaptic theory of working memory. Science. 2008; 319:1543–1546. [PubMed: 18339943]

Moore T, Armstrong KM. Selective gating of visual signals by microstimulation of frontal cortex. Nature. 2003; 421:370–373. [PubMed: 12540901]

Pelli DG. The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat Vis. 1997; 10:437–442. [PubMed: 9176953]

Pessoa L, Gutierrez E, Bandettini P, Ungerleider L. Neural correlates of visual working memory: fMRI amplitude predicts task performance. Neuron. 2002; 35:975–987. [PubMed: 12372290]

Pratte MS, Tong F. Spatial specificity of working memory representations in the early visual cortex. J Vis. 2014; 14:22. http://dx.doi.org/10.1167/14.3.22. [PubMed: 24648192]

Raposo D, Kaufman MT, Churchland AK. A category-free neural population supports evolving demands during decision-making. Nat Neurosci. 2014; 17:1784–1792. [PubMed: 25383902]

Rowe J, Hughes L, Eckstein D, Owen AM. Rule-selection and action-selection have a shared neuroanatomical basis in the human prefrontal and parietal cortex. Cereb Cortex. 2008; 18:2275–2285. [PubMed: 18234684]

Riggall AC, Postle BR. The relationship between working memory storage and elevated activity as measured with functional magnetic resonance imaging. J Neurosci. 2012; 32:12990–12998. [PubMed: 22993416]

Rigotti M, Barak O, Warden MR, Wang XJ, Daw ND, Miller EK, Fusi S. The importance of mixed selectivity in complex cognitive tasks. Nature. 2013; 497:585–590. [PubMed: 23685452]

Ringach DL, Shapley RM, Hawken MJ. Orientation selectivity in macaque V1: diversity and laminar dependence. J Neurosci. 2002; 22:5639–5651. [PubMed: 12097515]

Ruff CC, Blankenburg F, Bjoertomt O, Bestmann S, Freeman E, Haynes JD, Rees G, Josephs O, Deichmann R, Driver J. Concurrent TMS-fMRI and psychophysics reveal frontal influences on human retinotopic visual cortex. Curr Biol. 2006; 16:1479–1488. [PubMed: 16890523]

Saber GT, Pestilli F, Curtis CE. Saccade planning evokes topographically specific activity in the dorsal and ventral streams. J Neurosci. 2015; 35:245–252. [PubMed: 25568118]

Saproo S, Serences JT. Attention improves transfer of motion information between V1 and MT. J Neurosci. 2014; 34:3586–3596. [PubMed: 24599458]

Serences JT, Ester EF, Vogel EK, Awh E. Stimulus-specific delay activity in human primary visual cortex. Psychol Sci. 2009; 20:207–214. [PubMed: 19170936]

Sprague TC, Ester EF, Serences JT. Reconstructions of information in visual spatial working memory degrade with memory load. Curr Biol. 2014; 24:2174–2180. [PubMed: 25201683]

Sprague TC, Saproo S, Serences JT. Visual attention mitigates information loss in small- and large-scale neural codes. Trends Cogn Sci. 2015; 19:215–226. [PubMed: 25769502]

Sreenivasan KK, Curtis CE, D'Esposito M. Revisiting the role of persistent neural activity during working memory. Trends Cogn Sci. 2014a; 18:82–89. [PubMed: 24439529]

Sreenivasan KK, Vytlacil J, D'Esposito M. Distributed and dynamic storage of working memory stimulus information in extrastriate cortex. J Cogn Neurosci. 2014b; 26:1141–1153. [PubMed: 24392897]

Srimal R, Curtis CE. Persistent neural activity during the maintenance of spatial position in working memory. Neuroimage. 2008; 39:455–468. [PubMed: 17920934]

Stokes MG. 'Activity-silent' working memory in prefrontal cortex: a dynamic coding framework. Trends Cogn Sci. 2015; 19:394–405. http://dx.doi.org/10.1016/j.tics.2015.05.004. [PubMed: 26051384]

Stokes MG, Kusunoki M, Sigala N, Nili H, Gaffan D, Duncan J. Dynamic coding for cognitive control in prefrontal cortex. Neuron. 2013; 78:364–375. [PubMed: 23562541]

Warden MR, Miller EK. Task-dependent changes in short-term memory in the prefrontal cortex. J Neurosci. 2010; 30:15801–15810. [PubMed: 21106819]

Zarahn E, Aguirre GK, D'Esposito M. Empirical analyses of BOLD fMRI statistics. I Spatially unsmoothed data collected under null-hypothesis conditions. Neuroimage. 1997; 5:179–197. [PubMed: 9345548]

Highlights

- Activity in frontoparietal (FP) cortex is elevated during active memory storage
- FP activity is thought to reflect top-down biasing signals rather than storage
- We show that sub-regions of FP cortex encode feature-specific information during WM
- These representations are dissociable from overall changes in mean activation

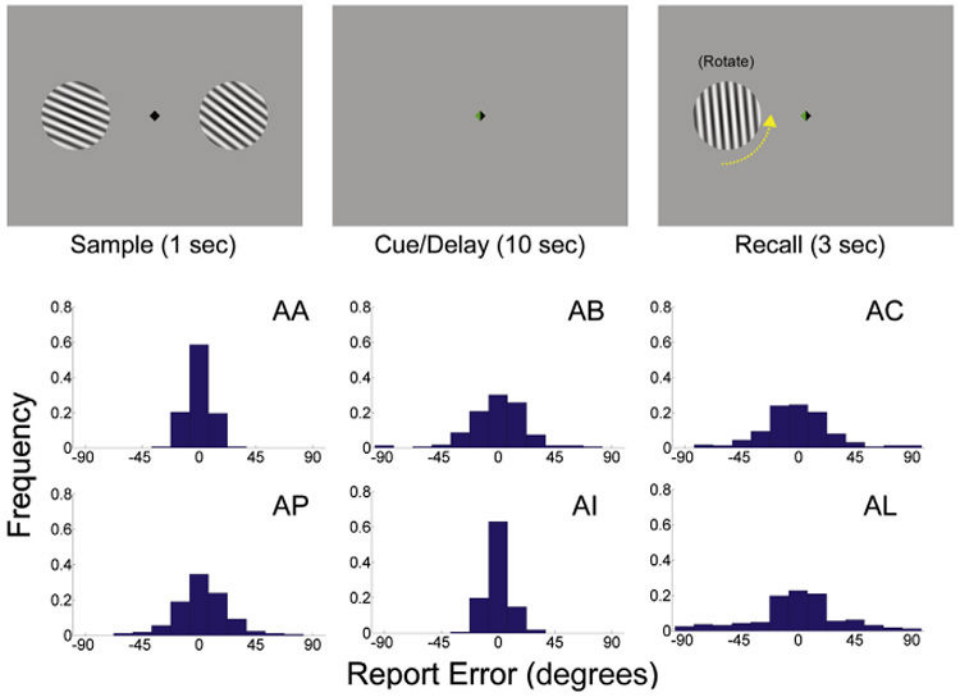

Figure 1. Behavioral Task and Behavioral Performance

(A) Participants viewed displays containing two lateralized gratings for 1,000 ms and were immediately post-cued to remember the orientation of the grating on the left and right side of fixation (indicated by the green half of fixation diamond). Following a 10 s delay, a randomly oriented probe grating appeared at the location of the remembered grating and participants were given 3 s to adjust its orientation to match that of the remembered grating using a button box (one button rotated the grating clockwise, the other rotated the grating counterclockwise, as illustrated by the dashed yellow arrow, not present on the visual display). The initial orientation of the probe was randomized with respect to the remembered orientation on each trial.

(B) Histograms of absolute recall error (i.e., reported minus actual orientation) for each of the six participants.

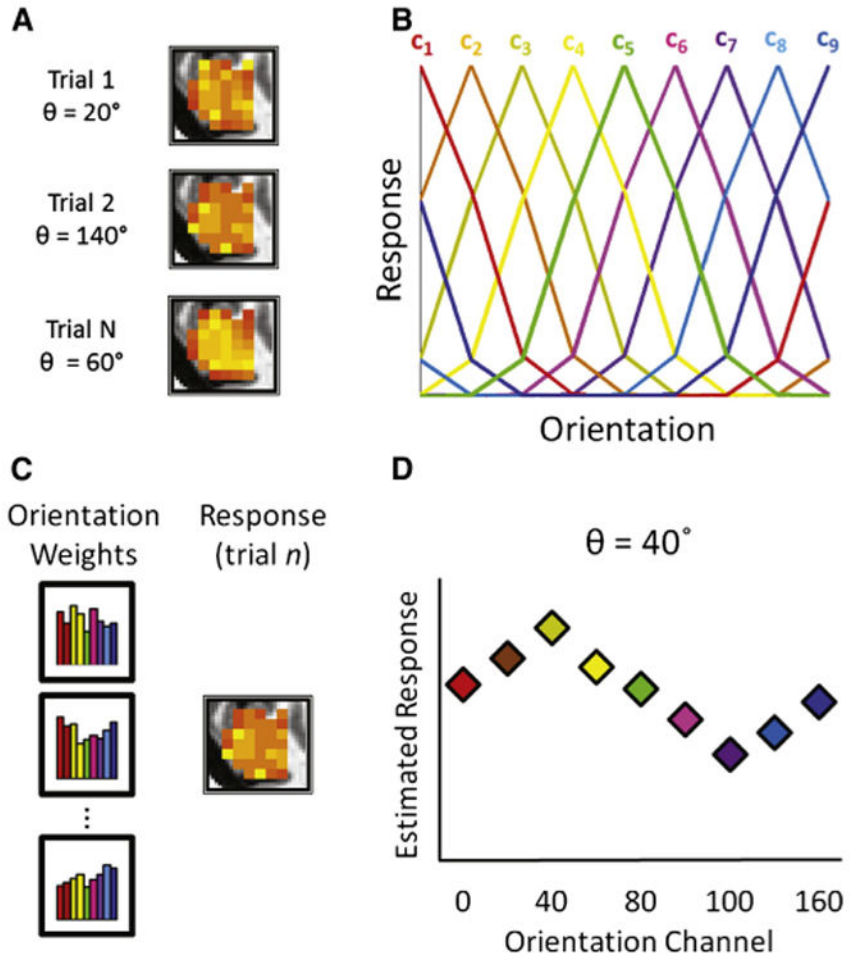

Figure 2. Inverted Encoding Model for Orientation Enables Reconstruction of Working Memory Representations from fMRI Activation Patterns

(A) On each trial, we measured delay period activation levels from a population of voxels within a given cortical area.

(B) We modeled the response of each voxel to different orientations over trials as a weighted sum of nine hypothetical orientation channels, each with an idealized response function.

(C) The result of this operation is a set of channel weights that characterize the orientation selectivity of each voxel.

(D) We then use the pattern of channel weights across all voxels within an ROI and a novel activation pattern from those voxels from a single trial to estimate the response of each orientation channel on that trial. See text and Experimental Procedures for further information. Trial-by-trial reconstructions were coregistered to a common orientation (0°) and averaged. Data in (A)–(D) are synthetic and for illustrative purposes only.

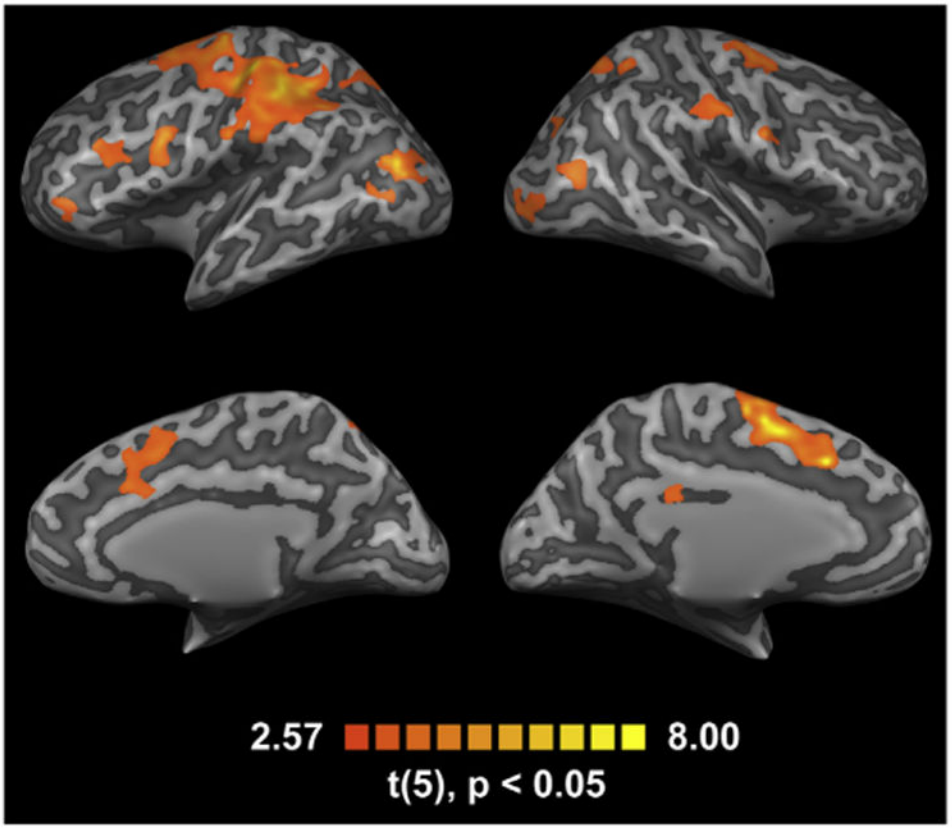

Figure 3. ROIs with Elevated Delay Period Activation

We used a random-effects general linear model to evaluate changes in the BOLD signal during the sample, delay, and probe epochs (see Experimental Procedures). We projected regions with elevated delay period activation onto a computationally inflated visualization of the cortical surface for a representative participant. This visualization was used to define the ROIs described in Figure 4, Figure S3, and Table 1 (see text and Experimental Procedures for details).

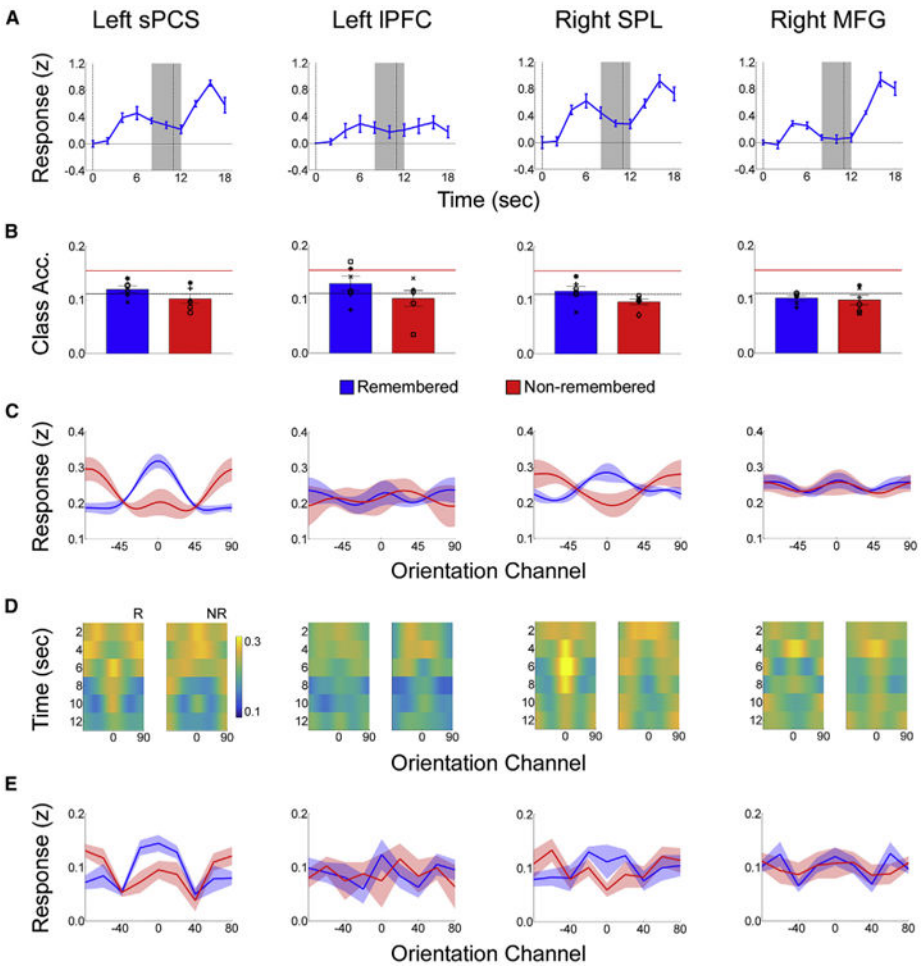

Figure 4. Univariate and Multivariate Analyses of BOLD Activation Patterns in a Subset of Regions with Elevated Delay Period Activity

(A) Estimated BOLD responses time locked to the onset of the sample array in each ROI. The vertical dashed lines at 0 and 11 s mark the onset of the sample and probe displays, respectively, and shaded regions mark the temporal epoch used in delay-period multivariate analyses (SVM classification and IEM reconstruction; B, C, and E).

(B) Multivariate classification accuracy for the remembered (blue) and non-remembered (red) orientations in each ROI. Horizontal dashed line at 0.1111 denotes theoretical chance classification accuracy assuming an infinite number of trials, and the solid red line at approximately 0.15 depicts empirically estimated chance decoding accuracy given the number of observations in each testing session (see Supplemental Experimental Procedures). Symbols correspond to individual participants.

(C) Reconstructions of the remembered and non-remembered orientations in each ROI. Data were averaged across samples obtained 8, 10, and 12 s following the onset of the sample display before modeling began.

(D) Time-resolved reconstructions of the remembered (“R”) and non-remembered (“NR”) orientations obtained by applying an IEM independently to data from each sample across an interval spanning 2 to 12 s after the onset of the sample display. All panels have the same color scale (see color bar).

(E) Reconstructions of the remembered and non-remembered orientations obtained using a basis set of nine Kroeneker delta functions (rather than the smooth sinusoids shown in Figure 2). Smooth reconstructions that peak at the remembered orientation (0°) are consistent with a continuous (rather than categorical or discrete) representation. For data analyzed from all delay-period ROIs, see Figure S3. All error bars and shaded regions are ±1 within-participant SEM.

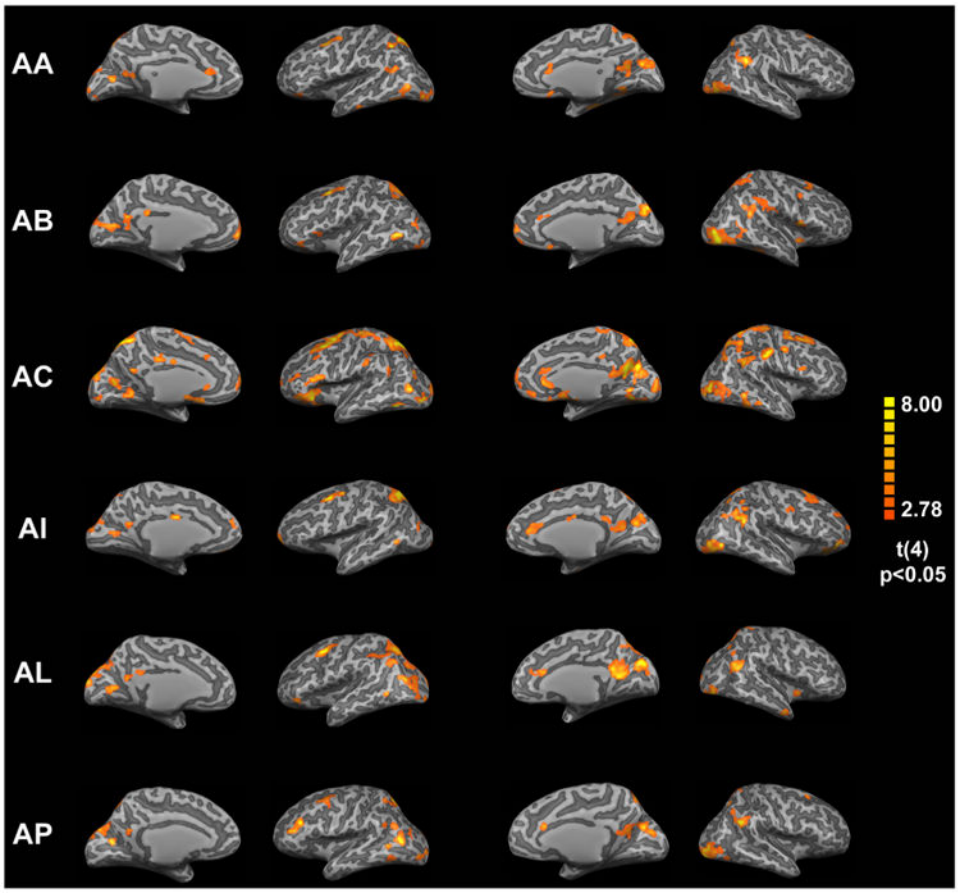

Figure 5. Searchlight-Defined ROIs Representing the Remembered Orientation

(A) Schematic of "leave-one-participant-out" cross-validation procedure. We generated an SPM marking neighborhoods containing a robust representation of the remembered orientation for each participant (e.g., AA) by submitting neighborhood-level amplitude estimates from the remaining five participants (e.g., AB-AP) to a t test against a distribution with a mean of 0. Thus, each participant's map was generated using data from the five other participants, but not his or her own data.

(B) Clusters containing a robust representation of the remembered orientation generated using the leave-one-participant-out approach described in (A). Each row corresponds to a different participant. These visualizations were used to define ROIs in bilateral dlPFC and left vlPFC (see text for details).

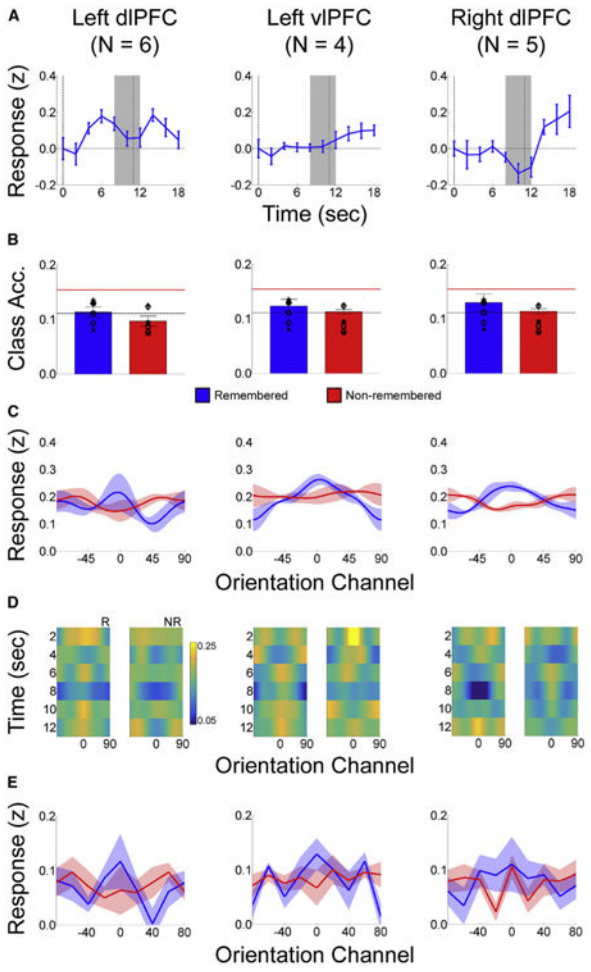

Figure 6. Reconstructions of the Remembered and Non-remembered Orientations in Searchlight-Defined Subregions of Pre-frontal Cortex

All conventions are identical to those shown in Figure 4.

(A) Estimated BOLD responses time locked to the onset of the sample display.

(B) Multivariate classification accuracy.

(C) Reconstructions of the remembered and non-remembered orientations. Note that these ROIs were defined using a leave-one-participant-out cross-validation approach (Figure 5). This ensures that the reconstructions shown here are statistically independent of the criteria used to define each participant's ROIs.

(D) Representations of the remembered (“R”) and non-remembered (“NR”) orientations computed on a sample-by-sample basis. All panels are on the same color axis (see color bar).

(E) Reconstructions of the remembered and non-remembered orientations obtained using a basis set of Kroeneker delta functions.

Table 1

Summary of ROIs with Elevated Delay Period Activation

	X	Y	Z	Hemi	Name	Size (# Voxel)	R	NR	R < NR
1	−4.31 (±0.21)	−31.33 (±0.07)	24.92 (±0.04)	Left	Ventromedial Cingulate	44 (±1)	0.720	0.049	0.950
2	−19.44 (±0.56)	−58.82 (±0.29)	49.80 (±1.09)	Left	Superior Parietal Lobule	212 (±19)	0.117	0.810	0.060
3	−43.28 (±1.07)	14.12 (±0.87)	25.64 (±0.32)	Left	Middle Frontal Gyrus	256 (±24)	0.643	0.514	0.602
4	−40.41 (±0.28)	−33.49 (±0.99)	43.71 (±0.47)	Left	Postcentral Sulcus	918 (±56)	0.060	0.243	0.235
5	−38.23 (±1.80)	−17.60 (±1.14)	47.37 (±1.13)	Left	Central Sulcus	452 (±47)	0.295	0.259	0.516
6	−24.27 (±1.11)	−6.39 (±0.72)	55.41 (±0.53)	Left	Superior Precentral Sulcus	416 (±30)	<1e–05	0.916	2e–04
7	−5.76 (±0.27)	5.61 (±0.42)	44.88 (±0.44)	Left	Medial Superior FG	382 (±10)	0.463	0.794	0.215
8	−36.24 (±0.65)	−73.54 (±1.38)	18.64 (±0.75)	Left	Occipitoparietal Cortex	222 (±27)	0.052	0.217	0.179
9	49.62 (±0.27)	−19.15 (±0.38)	37.22 (±0.34)	Right	Postcentral Sulcus	146 (±11)	0.102	0.994	0.006
10	42.85 (±0.79)	−65.33 (±0.10)	13.61 (±0.44)	Right	Occipitoparietal Cortex	97 (±7)	0.039	0.865	0.019
11	26.48 (±0.17)	−4.58 (±0.26)	54.20 (±0.49)	Right	Superior Precentral Sulcus	172 (±10)	0.049	0.763	0.082
12	25.51 (±0.21)	−72.83 (±0.42)	32.37 (±0.76)	Right	Intraparietal Sulcus	82 (±3)	0.051	0.746	0.064
13	7.89 (±0.21)	10.46 (±0.50)	41.42 (±0.88)	Right	Medial Superior FG	300 (±14)	0.364	0.363	0.506
14	−20.95 (±0.46)	−54.17 (±0.36)	51.42 (±0.62)	Right	Superior Parietal Lobule	259 (±26)	0.062	0.894	0.029

X, Y, and Z are mean (±1 SEM) Talariach coordinates for each ROI. R and NR show p values depicting the robustness of the remembered and non-remembered orientations, respectively (see Quantification and Comparison of Reconstructed Representations in Experimental Procedures). For R, a p value < 0.05 corresponds to a robust (greater than zero) representation. For NR, a p value < 0.05 indicates that the amplitude of the non-remembered orientation reconstruction was reliably less than 0. R > NR shows p values comparing the strengths of remembered and non-remembered reconstructions. A p value < 0.05 means that reconstructions of the remembered orientation were significantly stronger than reconstructions of the non-remembered orientation. PFC, prefrontal cortex; FG, frontal gyrus.

Table 2

Searchlight ROI Coordinates and Sizes for Each Participant

	Left dlPFC				Right dlPFC				Left vlPFC			
	X	**Y**	**Z**	**Size**	**X**	**Y**	**Z**	**Size**	**X**	**Y**	**Z**	**Size**
AA	−30	5	49	198	22	14	51	55	−37	34	−3	100
AB	−29	3	47	210	27	10	45	176	−40	27	1	153
AC	−35	7	51	886	30	6	50	529	−38	27	1	961
AI	−29	4	48	209	22	14	49	273	–	–	–	N/A
AL	−30	7	47	228	–	–	–	N/A	−31	31	−6	134
AP	−33	6	50	226	22	14	52	49	–	–	–	N/A

Size refers to the number of 2 × 2 × 2 mm voxels within each ROI. X, Y, and Z are Talariach coordinates that correspond to the centroid of the ROI. We could not identify any clusters representing the remembered orientation located near dlPFC in one participant (AL). Similarly, we could not identify left any clusters representing the remembered orientation near vlPFC in two participants (AI and AP). Consequently, the ROIs described above were generated using clusters from the remaining five and four participants, respectively. See Figure 5 and text for more information.