

ORIGINAL ARTICLE

Shared Representational Formats for Information Maintained in Working Memory and Information Retrieved from Long-Term Memory

Vy A. Vo^{1,2,3}, David W. Sutterer⁴, Joshua J. Foster^{5,6}, Thomas C. Sprague⁷, Edward Awh^{8,9} and John T. Serences^{2,3,10}

¹Brain-Inspired Computing, Emerging Technologies Research, Intel Labs, Hillsboro, OR 97124, USA,

²Neurosciences Graduate Program, University of California San Diego, La Jolla, CA 92093, USA, ³Department of Psychology, University of California San Diego, La Jolla, CA 92093, USA, ⁴Department of Psychological Sciences,

Vanderbilt University, Nashville, TN 37235, USA, ⁵Department of Psychological and Brain Sciences, Boston

University, Boston, MA 02215, USA, ⁶Center for Systems Neuroscience, Boston University, Boston, MA, USA,

⁷Department of Psychological and Brain Sciences, University of California, Santa Barbara, CA 93106, USA,

⁸Department of Psychology, The University of Chicago, Chicago, IL 60637, USA, ⁹Institute for Mind and Biology,

The University of Chicago, Chicago, IL 60637, USA and ¹⁰Kavli Foundation for the Brain and Mind, University of California San Diego, La Jolla, CA 92093, USA

Address correspondence to Dr Vy A. Vo. Email: vy.vo@intel.com

Abstract

Current theories propose that the short-term retention of information in working memory (WM) and the recall of information from long-term memory (LTM) are supported by overlapping neural mechanisms in occipital and parietal cortex. However, the extent of the shared representations between WM and LTM is unclear. We designed a spatial memory task that allowed us to directly compare the representations of remembered spatial information in WM and LTM with carefully matched behavioral response precision between tasks. Using multivariate pattern analyses on functional magnetic resonance imaging data, we show that visual memories were represented in a sensory-like code in both memory tasks across retinotopic regions in occipital and parietal cortex. Regions in lateral parietal cortex also encoded remembered locations in both tasks, but in a format that differed from sensory-evoked activity. These results suggest a striking correspondence in the format of representations maintained in WM and retrieved from LTM across occipital and parietal cortex. On the other hand, we also show that activity patterns in nearly all parietal regions, but not occipital regions, contained information that could discriminate between WM and LTM trials. Our data provide new evidence for theories of memory systems and the representation of mnemonic content.

Key words: fMRI, long-term memory, multivariate pattern analysis, working memory

Introduction

It is well established that performance on a working memory (WM) or a long-term memory (LTM) task relies on the partial reactivation, or reinstatement, of patterns of cortical activity that were present during the initial perception of the remembered item (Squire and Wixted 2011; D'Esposito and Postle 2015; Serences 2016; Xue 2018). For example, viewing a blue couch elicits feature-specific patterns of activity in visual regions selective for color and complex objects. Then, when trying to maintain the colored object in WM, or retrieve it from LTM, patterns of neural activity that resemble those at encoding are reinstated. That said, evidence for feature-specific cortical reinstatement has typically been assessed separately in studies of LTM (Ritchey et al. 2013; Bosch et al. 2014; Kuhl and Chun 2014; Xiao et al. 2017; Favila et al. 2018) and studies of WM (Harrison and Tong 2009; Serences et al. 2009; Riggall and Postle 2012; Emrich et al. 2013; Ester et al. 2013, 2016; Lee et al. 2013; Sprague et al. 2014, 2016). Several prominent theories suggest that both WM and LTM retrieval rely on shared neural substrates (Atkinson and Shiffrin 1968; Kosslyn 1980; Cowan 1995; Baddeley 2000), but direct tests of this hypothesis are rare. One study reported evidence that there is an overlap between representations of stimulus categories held in WM or activated from LTM (Lewis-Peacock and Postle 2008). Furthermore, some of these theories propose that retrieving information from LTM would simply place it in short term or WM, implying that the representations between WM and LTM would be equivalent and indistinguishable (Atkinson and Shiffrin 1968; Cowan 1995). We are unaware of any tests of equivalence between representations maintained in WM and retrieved from LTM, perhaps because this would require carefully matching the task performance.

Here, we strived for a direct comparison of cortical reinstatement of stimulus features between a WM task and a LTM task. Participants either maintained a spatial position in WM, or retrieved it from LTM and maintained it for the same duration (Fig. 1), while we measured responses with functional magnetic resonance imaging (fMRI). Importantly, we carefully matched behavioral performance across memory tasks and minimized any differences in visual stimulation between the tasks. Therefore, the remembered information was the same across conditions—it simply arose from a different source (perception for WM vs. an internal source for LTM). We used multivariate pattern analysis to examine the remembered representations in early occipital visual areas and across several regions in retinotopic and lateral parietal cortex (LPC). In this paper, we use the terms “representation” or “mnemonic representation” to specifically refer to information about the memory stimulus present in a given region during the maintenance interval of either task. The eight retinotopic regions that we defined (V1 through dorsal parietal regions in the intraparietal sulcus, IPS0–2) are known to have sensory-like WM representations (Riggall and Postle 2012; Emrich et al. 2013; Sprague et al. 2014, 2016; Ester et al. 2015) that are finely resolved enough to decode stimulus features such as orientation or motion direction. That is, a model trained on data from a sensory task can decode feature information during a WM task. Recent evidence suggests that in intraparietal sulcus (IPS), feature-specific WM representations can be encoded in a nonsensory format (Christophel et al. 2018; Rademaker et al. 2019; Iamshchinina et al. 2021). That is, only a model trained on data from a WM task can decode information during a held-out WM run. As for LTM, the data on reinstated stimulus features in retinotopic regions are limited to V1–V3 (Bosch et al. 2014) or rely on a single large occipitotemporal

region of interest (ROI) (Favila et al. 2018). However, work has shown that feature-specific LTM representations are encoded in LPC (Kuhl and Chun 2014; Lee et al. 2017), likely in a nonsensory format (Xiao et al. 2017; Favila et al. 2018). Therefore, we also examined a set of lateral parietal regions (dorsal and ventral lateral intraparietal sulcus, together lateral IPS; angular gyrus [AnG]) reported in those studies. The role of LPC in WM is not well understood (Xiao et al. 2017; Favila et al. 2018, 2020). Accordingly, we took advantage of our experimental design to test whether the reinstated information in these regions is similar between WM and LTM or if the feature-specific information is maintained via different representational formats.

First, we used an inverted encoding model (IEM) to quantify the fidelity of cortical reinstatement in both tasks, which was based on the amount of spatial information in fMRI activation patterns. Put another way, the IEM indexes the quality and strength of the mnemonic representation. To test whether the remembered spatial position was represented in a sensory-like format, we trained the IEM on an independent perceptual task and tested it on either the WM maintenance task or LTM retrieval task. To test if a mnemonic representation was in a nonsensory-like format, we trained the IEM on one of the memory tasks instead. First, we found that memories for spatial position maintained in WM or retrieved from LTM were represented in a sensory-like format across all retinotopic ROIs. Furthermore, the fidelity or quality of these representations was statistically indistinguishable between WM and LTM. We found that lateral IPS and AnG both tracked some spatial information held in memory. However, this information was represented in a nonsensory format.

Next, we used a two-way linear classifier to decode the memory task that subjects were performing on each trial. We found that the memory task could be decoded from activation patterns in parietal regions (both retinotopic and lateral) but not in early visual cortex. Taken together, our results provide direct evidence that information in maintained WMs and retrieved LTMs rely on similar neural representations in occipital and parietal cortex. However, parietal cortex, unlike early visual cortex, also encodes information about the source of the remembered information in addition to coding the remembered location.

Methods

Participants

We trained 12 human participants (8 females) on both the LTM and the WM tasks. Two participants dropped out of the study during behavioral training. This left 10 participants (6 females) from whom we collected both behavioral and fMRI data, 2 of whom were the coauthors of this paper. Six of these participants (participant IDs BI, BJ, BO, and BC) had previously completed a set of retinotopic mapping scans in the lab for other studies (Henderson and Serences 2019; Rademaker et al. 2019). For the remaining four participants, we collected new retinotopy data to define each ROI (see ROI Definition). All participants provided written informed consent. Participants were compensated for their time (\$20/h for fMRI, \$10/h for behavior) unless they were this paper's coauthors. These procedures were approved by the local UC San Diego Institutional Review Board.

Prescan Behavioral Training

We trained participants to form LTM pairings between 24 unique clip art items and 24 unique spatial positions (Fig. 1). After 1 day

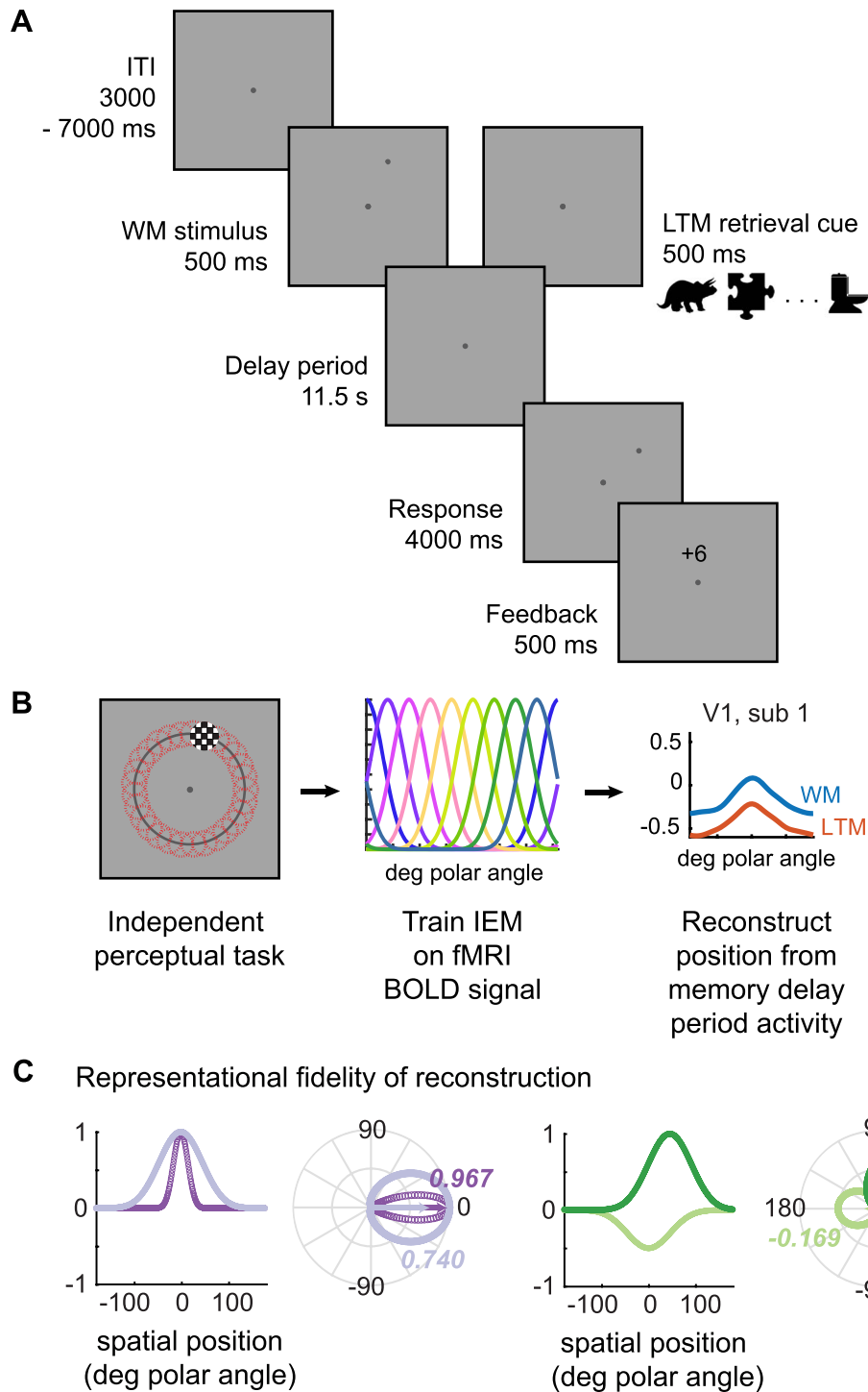


Figure 1. Memory tasks and perceptual task presented to participants in the scanner. (A) The two memory tasks were identical except for the memory cue—in the WM task, participants saw a dot indicating the position along an annulus, along with an irrelevant clip art item; in the LTM task, participants instead saw a previously studied clip art item that they had learned to associate with a spatial position during behavioral training. After an 11.5-s delay period, participants reported the remembered spatial position by rotating a randomly placed dot to the correct position. (B) An independent task was used to map the position selectivity of voxels in retinotopically defined regions of visual and parietal cortex. Participants detected an occasional dimming of a checkerboard stimulus that appeared at randomly ordered locations along a ring. This independent task was used to train the IEM, which was then tested on the memory tasks. (C) Example spatial representations and their corresponding fidelity values, a single number which characterizes the quality of the representation of the remembered position (always set to 0° here). Each spatial representation was plotted in polar space. The fidelity metric is equivalent to the length of the bold horizontal vectors. Dark purple: The best model-based spatial representation is narrow and centered exactly at 0°. Light purple: A broader spatial representation has a shorter mean vector, capturing the fact that less “energy” is at 0°. Dark green: A spatial representation that is slightly offset from zero has a short x-component and therefore a lower fidelity value, as compared to a representation centered at zero. Light green: An inverted spatial representation has a mean vector that points in the opposite direction, resulting in a negative fidelity value.

of LTM study and retrieval, we continued LTM training while also measuring their performance on an analogous WM task. In both tasks, participants were asked to report the remembered spatial positions as precisely as possible. To minimize differences in effort and difficulty of recall between the two memory tasks at the time of scanning, participants were trained until their error on the LTM task was stable.

For each participant, the 24 spatial positions were selected randomly from 24 bins evenly spaced along an isoecentric ring at 3.9° visual angle from a central fixation point. As a result, the set of locations was roughly uniformly distributed around a circle for each participant, and each participant had a slightly different set of positions. Each spatial position was rendered as a dot subtending 0.2° visual angle (in polar co-ordinates of the 360° circle, this subtended 1.4° polar angle). The clip art items (diameter: 1.5° visual angle) were filled outlines of real-world objects taken from a bank of 48 images that all subtended the same area of retinal space (Fig. 1). These stimuli have been used in previous LTM studies and are all royalty-free images (Sutterer and Awh 2015; Sutterer et al. 2019).

The LTM study-retrieval task on the first day of training consisted of six alternating runs of a study phase and a retrieval phase. During the study phase, the participant was shown each clip art item at the center of the screen and its associated spatial position. The trial order was randomized. This study phase was self-paced—while each pairing was always presented for 500 ms followed by at least a 1000-ms blank period, participants chose when to advance to the next item by hitting the space bar. During the retrieval phase of the task, each clip art item was presented at the center in a randomized order, and participants were asked to report the associated spatial position as precisely as possible by clicking on a part of the visible isoecentric ring. After each retrieval trial, participants received two forms of feedback about their performance: The true position was shown in dark gray on the ring, and the error was shown as a signed integer at the center. This number indicated how far their report deviated from the true position (+, if report was too far clockwise; −, for counterclockwise).

On all remaining days of prescan training, participants practiced both the LTM retrieval task and a similar WM maintenance task (similar to Fig. 1A, with different timings). The order of the tasks was counterbalanced across participants. In the LTM retrieval task, the participant was shown one of the 24 clip art items at the center (500 ms). After a delay (1000 ms), they were presented with the response probe: A dot appeared at a random position along the (now invisible) isoecentric ring, and participants used four keyboard keys to rotate this dot along the ring until it matched the item's associated position as precisely as possible. The two outer buttons moved the dot along the ring at a fast rate, whereas the two inner buttons moved it along a slower rate. This allowed participants to make fast but precise responses. After the 4000-ms response period, participants received numerical feedback on their response error (Fig. 1). We chose to only present numerical feedback to minimize any interference between the retrieved memory representation and feedback given at the end of the trial. The intertrial interval (ITI) varied between 200 and 500 ms. Participants repeated this set of 24 trials per run, for a total of six runs in the LTM task block.

The WM maintenance task was nearly identical to the LTM task in timing and presentation. The one exception was that the initial memory array displayed a dot at a random spatial position on the ring in addition to a clip art item presented at the center, which was randomly selected from a separate set of 24 unique

clip art items. Participants were asked to remember the position of the dot during the 1-s delay and to use the keyboard keys to rotate the response dot to the position of the memorandum (Fig. 1). The clip art item was a task-irrelevant sensory control, and there was no predictive relationship between WM position and clip art item. As with the LTM task, a set of 24 trials was repeated for six times.

All participants performed six runs of each task on each training day. The first two participants (BG, BH) completed more training runs as we piloted the best way to collect response data (mouse or keyboard). Their behavioral performance in the scanner did not appear significantly different from the other subjects. All participants were trained on both the LTM and WM tasks for at least 5 days. At the end of each training day, the experimenter plotted the mean recall error on both tasks to assess whether the mean recall error on the LTM retrieval task was stable across days. Two different experimenters were involved in training, and the criterion was subjective for each experimenter. Importantly, however, the criterion does not enforce similar recall error between the WM and LTM tasks. Instead, it was simply a check for stable performance in the LTM task. One participant (BV) requested to repeat the initial LTM study training on day 4 of training.

Behavioral Tasks Performed in fMRI Scanner

Participants performed both LTM and WM tasks in the scanner (Fig. 1A). While these were very similar to the behavioral training task, they differed in several key respects: 1) instead of a 1000-ms delay between the cue and the response, we inserted a 11500-ms delay to accommodate the slow hemodynamic response; 2) LTM and WM runs were alternated instead of blocked in the order preserved from the training; 3) participants viewed the experiment through a head-coil-mounted mirror pointing at a gray rectangular screen (120 × 90 cm) at the foot of the scanner bore (~3.85 m viewing distance); 4) responses were made via a four-key button box; and 5) the ITIs were longer (uniformly distributed in 200-ms steps between 3000 and 7000 ms) to aid in estimating the stimulus-evoked blood oxygen level-dependent (BOLD) response. Participants continued to receive feedback on every trial that indicated the signed response error.

In addition to the memory tasks, participants also performed an independent mapping task, which we used to train the spatial encoding model (see below) and to localize voxels responsive to the isoecentric ring (as before, 3.9° from central fixation point). In this task, participants fixated at a central point as a flickering stimulus (6 Hz) appeared at some location on the screen for 3000 ms. Participants monitored the flickering stimulus for an occasional contrast change (8 of 54 total trials per run) and reported whether the contrast increased or decreased on those trials (pedestal at 70% Michelson contrast). On most trials, the stimulus was a circular checkerboard (0.9° radius, 1.36 cycles/deg) centered at a point on the isoecentric ring. Each location was pseudorandomly drawn from one of 30 evenly spaced polar bins. On five trials, participants instead saw a ring aperture (1° wide) with a similar flickering checkerboard pattern (3.9° from center). On the target trials, either the entire circular checkerboard changed contrast, or a 0.681° (36° polar angle) wide area of the ring changed contrast for 500 ms. We also included six null trials to aid in the estimation of the stimulus-evoked BOLD response. Trial order was randomized, with ITIs varying between 2000 and 5000 ms (mean: 3500 ms). The magnitude of the contrast change was manually adjusted by the experimenter

on each run to keep task performance for each participant near 75% (participant averaged mean and bootstrapped 95% confidence interval [CI]: 71.7% accuracy [60.11%, 81.62%]; Michelson contrast change 0.50 [0.42, 0.58]).

Statistical Analysis

The majority of the statistical analyses reported in the paper and all CIs were generated by bootstrapping. We opted for a nonparametric approach because it carries fewer assumptions about the distribution of the underlying data.

Unless otherwise noted, the bootstrapping was done by resampling with replacement across the $N=10$ participants for 10 000 iterations. To correct the CIs for a bias toward narrowness in small samples, we drew $N-1$ samples on each iteration (Hesterberg 2011). Then, the mean of the value of interest (e.g., response error) was computed on each resampling iteration to generate a distribution of means for the sampled population. Generally, we calculated 95% CIs using the percentile method. When computing the fidelity metric, however, we also corrected for skewness in the CIs across participants by applying the bias correction and acceleration (BCa) adjustment (Davison and Hinkley 1997; Hesterberg 2014) after estimating bias and acceleration by jackknife. BCa-adjusted values also produced an accompanying P value.

To evaluate the evidence for the null hypothesis that there was no difference between the WM and LTM tasks, we computed a Bayes factor for the comparison between conditions. We used the “BayesFactor” package in R 3.5.3 to convert a paired-sample t -value into a Bayes factor, given a standard Cauchy prior on the effect size and Jeffreys prior on variance (Rouder et al. 2009). For a stable estimate of the t -value in our small sample, we computed a t -value on every iteration of the across-subject bootstrap described above. This generated a distribution of t -values. We took the mean t -value of the bootstrapped distribution to calculate the Bayes factor in favor of the null hypothesis of no difference between WM and LTM (BF_{01}).

To assess effects across ROIs, we ran repeated-measures ANOVAs on either the representational fidelity metric or the decoding accuracy. While it is challenging to control for differences in ROI size and for correlations between ROIs, we adopted a nonparametric randomization approach that should minimize any spurious effects from these factors. We compared the F -score from a typical ANOVA to a null distribution of F -scores that were generated by permuting the data across all factors (e.g., ROI, temporal epoch, and memory task), separately within each participant, for 10 000 iterations (Manly 2007).

When multiple statistical comparisons are made (e.g., repeated for each ROI), we corrected across all test P values using the false discovery rate procedure ($q=0.05$ unless otherwise noted; Benjamini and Yekutieli 2001).

Behavioral Data Analysis

To describe participant performance on each memory task, we examined each participant's trial-by-trial response error. This is the signed difference between the true location of the memorandum and the participant's response (between 0 and $\pm 180^\circ$). Previous findings have shown that a histogram of response errors can be described as a mixture of a uniform distribution, representing guesses, and a circular Gaussian distribution centered near the correct response (Zhang and Luck 2008). The two parameters of the circular Gaussian describe the mean response

(μ), which describes any systematic bias, and the standard deviation (SD) of the responses, which describes the variability of the memory reports. These parameters were fit using publicly available code (Schneegans and Bays 2016).

We also fit the data with a simpler single-parameter model. We set the mean of the circular Gaussian to be the sample mean and used a modified version of the same code to fit the SD. We then compared the single-parameter and two-parameter models to one another by computing the Bayesian Information Criterion separately for each participant.

All behavioral data analyses used 10 000 resampling iterations.

Magnetic Resonance Imaging

We obtained all structural and functional MR images from participants using a GE 3 T MR750 scanner at the University of California, San Diego. We collected all functional images (19.2 cm² FOV, 64 × 64 acquisition matrix, 35 interleaved slices, 3 mm³ voxels with 0-mm slice gap, 263 volumes per memory run, and 179 volumes per mapping run) using a gradient echo planar pulse sequence (2000 ms repetition [TR], 30 ms echo [TE], 90° flip angle) and a 32-channel head coil (Nova Medical). Five dummy scans preceded each functional run. A high-resolution structural image was acquired at the end of each session using a FSPGR T1-weighted pulse sequence (25.6 cm² FOV, 256 × 192 acquisition matrix, 8136/3172 ms TR/TE, 9° flip angle, 1 mm³ voxels, and 172 volumes). All functional scans were coregistered to the anatomical images acquired during the same session, and this anatomical was in turn coregistered to the anatomical acquired during the retinotopy scan.

EPI images were unwrapped with a custom script from UCSD's Center for Functional Magnetic Resonance Imaging that calls the FSL PRELUDE and FUGUE functions. All subsequent preprocessing was performed in BrainVoyager 2.6.1, including slice-time correction, affine motion correction, and temporal high-pass filtering to remove slow signal drifts over the course of each session. Data were then transformed into Talairach space and resampled to a 3 × 3 × 3 mm voxel size. Finally, the BOLD signal in each voxel was Z-transformed on a scan-by-scan basis. All subsequent analyses were performed in MATLAB using custom scripts (available at <https://osf.io/fcp5y/>).

For the encoding and decoding analyses, we balanced the data by equating the number of runs for each memory task. Some data were discarded due to excessive movement (>3 mm in one run/scan). In these cases, we dropped runs until there were an equal number of WM and LTM runs.

ROI Definition

The retinotopic ROIs were first defined for each participant following a previously published procedure (Sprague and Serences 2013). Briefly, we collected fMRI data in a separate 1.5-h session that mapped the visual field. While one task was a passive viewing task, the other task required participants to attend to the portion of the visual field which contained the checkerboard stimulus and to report an occasional dimming in one portion of the stimulus. This attention task allowed us to reliably define parietal ROIs. In total, we defined eight retinotopic regions: V1, V2, V3, V4, V3A/B, IPS0, IPS1, and IPS2.

We then applied a mask to the retinotopically defined occipital and parietal regions to isolate voxels that show some response to the area where the spatial positions were presented.

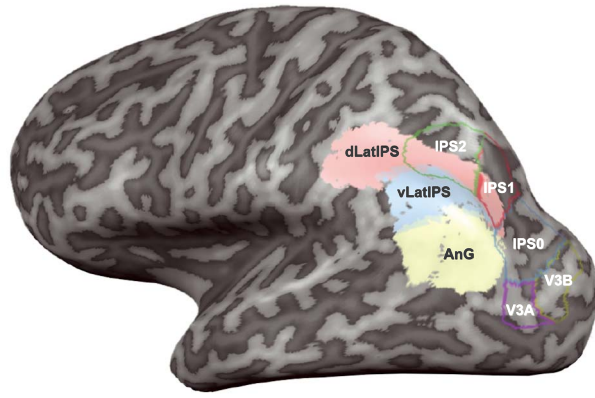


Figure 2. Retinotopic parietal ROIs (outlines) and lateral parietal ROIs (colored) for the left hemisphere of an example participant.

This functional localizer relied on defining different conditions for the circular checkerboard and ring aperture checkerboard trials during the mapping task. We estimated BOLD response activation to each condition by convolving the trial events with a canonical two-gamma HRF (peak at 5 s, undershoot peak at 15 s, response undershoot ratio 6, response dispersion 1, and undershoot dispersion 1). We then solved a generalized linear model (GLM) on all runs of this mapping task for a single participant. This produced a statistical parametric map of voxels with a significant BOLD response change attributable to each condition (FDR $q = 0.05$, Benjamini and Yekutieli 2001). The localizer mask only included voxels that responded with a significant BOLD response increase during the ring aperture condition.

In addition to retinotopic regions, we also defined three lateral parietal regions: dorsal lateral intraparietal sulcus (dLatIPs), ventral lateral intraparietal sulcus (vLatIPs), and AnG. We defined these regions following a previous method used in a study of episodic memory (Favila et al. 2018). First, we used the Freesurfer parcellation to specify the LPC as any area in the superior parietal, inferior parietal, and supramarginal parcels. We then subdivided this portion of cortex using a 17-network atlas defined from resting-state functional connectivity data (Yeo et al. 2011). dLatIPs was part of network 12 in the atlas, and vLatIPs was part of network 13. Both of these networks are part of the frontoparietal control network. The AnG was defined by combining the small parietal nodes from networks 15, 16, and 17, which are part of the default mode network.

Note that these atlas-defined regions were not defined using any functional data. As a result, they exhibit some overlap with the retinotopically defined parietal regions (see Fig. 2).

IEM for Spatial Position

Following previously published methods, we first estimated an encoding model for spatial position using z-scored BOLD responses obtained from the independent perceptual task (Brouwer and Heeger 2009; Sprague and Serences 2013; Sprague et al. 2018). This allowed us to estimate the response of each voxel to any arbitrary spatial position covered by the visual stimuli. Importantly, the encoding model approach allows us to test more specific hypotheses about the format of mnemonic information in a given region. This distinguishes it from other multivariate pattern analysis methods that have been used

to measure encoding-retrieval similarity. Below, we detail a model that assumes each voxel has a spatial tuning profile along a continuous axis such that responses to nearby positions should be more similar than responses to far-apart positions. This assumption is known to be true in retinotopic regions, and prior work has shown that spatial encoding models are able to measure expected properties of perceptual codes, such as broadening of spatial representations with increasing eccentricity (Sprague and Serences 2013). If these models can recover accurate representations of remembered positions in nonretinotopic regions, we can infer that the activation pattern associated with memory recall in each voxel covaries with spatial position. This does not imply retinotopy in the region or even stable spatial tuning across many tasks—it only implies that it is stable across the training data and the test data. As a test of reinstatement from perception, this model is a more exact hypothesis test than correlative methods or general-purpose classifiers which do not instantiate the hypothesis of the encoding format in the algorithm itself. Finally, this approach only accounts for signal correlations between voxels and ignores noise correlations unlike other previously used approaches (see discussion in Kriegeskorte and Douglas 2019). Hence, any inferences made from the IEM analyses are only about the spatial encoding properties in these regions, given single-voxel, spatially specific responses, and not about other information that could be simultaneously encoded.

We defined nine evenly spaced spatial channels, or basis functions, along the polar angle axis (Fig. 1). This was the only axis we varied in our mapping runs. Each spatial channel was defined by equation (1), assuming a discretized polar axis (i.e., $0-2\pi$ in steps of 0.0175 radians, or 1°). The width and center of the basis functions are modified by the term s/FWHM , where s is the circular distance between the channel center and the polar axis, and FWHM is the desired full-width half-maximum of the response of the spatial channels. Our basis functions had a FWHM of 60° . Finally, to keep the encoding model weights within an interpretable range, we want each spatial channel to have a baseline of 0 and a maximum of 1. To achieve this, we used a positive, half-wave rectified cosine.

$$f(s) = \left(0.5 * \left(1 + \cos\left(\frac{s}{\text{FWHM}}\right)\right)\right)^7 \text{ for } s \leq r. \quad (1)$$

Finally, although cosine functions are periodic, each spatial channel should only have a peak response at its center, and nowhere else. We mask the function by setting all values beyond the radius r to 0 (here, $r = 151$).

The basis set $\mathbf{S}$ is the nine spatial channels in the discretized polar axis (c channels $\times$ p points along the polar axis, here, 9×360). For every training trial, we used the basis set to specify a modeled activation on each of the c channels for each presented mapping stimulus, resulting in a design matrix $\mathbf{C}$ (c channels $\times$ n trials).

The forward encoding model is then described in equation (2), where $\mathbf{B}$ = BOLD responses (v voxels $\times$ n trials). As long as $n > c$, it is possible to solve for $\mathbf{W}$ using the pseudoinverse (eq. 3). We solved equation (3) using “mldivide” in MATLAB.

$$\mathbf{B} = \mathbf{W}\mathbf{C}, \quad (2)$$

$$\mathbf{W} = \mathbf{C}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T, \quad (3)$$

$$\mathbf{C}_2 = (\mathbf{W}^T\mathbf{W})^{-1}\mathbf{W}^T\mathbf{B}_2. \quad (4)$$

After training the model by solving for $\mathbf{W}$, we inverted the encoding model (eq. 3) to estimate the activation of each spatial channel, given the BOLD responses on a separate test dataset and the independently trained model from the same participant (eq. 4). The spatial channel activations allowed us to estimate the spatial position of the stimulus that was presented in each trial. The test dataset could be the BOLD responses of voxels during the memory delay of either the WM or LTM task (Sprague et al. 2014; Ester et al. 2015), or the BOLD responses to a held out portion of the independent mapping task. Notably, we attempted to recover spatially specific information in the patterns of voxel responses during each of two memory tasks when “no visual stimulus was shown on the screen.” When directly comparing the fidelity of WM and LTM representations, we used a “fixed” encoding model across conditions. That is, the encoding model was estimated once (here, on data from a separate task) such that the trained model is “fixed” across conditions. Then to test the model, reconstructions were obtained for each condition separately. As noted previously, it is essential to use a fixed encoding model scheme to directly compare the model-based representations across different conditions (Sprague et al. 2018, 2019).

We estimated channel responses $\mathbf{C}_2$ from two different types of BOLD data $\mathbf{B}_2$. In our first analysis, $\mathbf{B}_2$ consisted of the BOLD response at each TR (after motion-correction and z-scoring within each run), for a total of eight TRs. In other analyses described below, we averaged the data across TRs for use in the IEM.

To generate a continuous representation of spatial position, we multiplied $\mathbf{C}_2$ by the basis set $\mathbf{S}$. We recentered all of the model-based representations such that the memory stimulus is at 0° and average across all trials.

We varied the dataset used for training and testing the IEM to test different hypotheses about how information is encoded in a given region. To test whether a region encoded information in a perception-like code, we trained on the perceptual mapping task and tested on the memory task. To identify which regions contained significant representations of the sensory stimuli (i.e., the spatial position), we trained and tested within the mapping dataset with leave-one-run-out cross-validation. To characterize whether a region contained significant spatial representations within a memory task, we trained and tested within the memory task, only using TRs from the late delay period. To ensure that the datasets are balanced across spatial position, we randomly resampled from each of 24 position bins and held out those 24 trials for testing. This resampling procedure was repeated 1000 times, and the results were averaged across these iterations. Finally, to test whether a region encoded spatial representations similarly across memory tasks, we trained on one memory task and tested on the other, again only using TRs from the late delay period.

Representational Fidelity Metric

To quantify the model-based memory representations, we compute a metric which describes the fidelity of the spatial representation. This is similar to previously published metrics but is modified so that it monotonically increases as the representation decreases in width (Sprague et al. 2016; Wolff et al. 2017).

Consider each point of the spatial IEM-based representation in polar co-ordinates as $[r_p, \theta_p]$, where r is the value of the representation in arbitrary units (y -axis in Fig. 1B) and θ is the

polar angle (x -axis in Fig. 1B). The representational fidelity is directly proportional to the average directional energy of the spatial representation pointing in the expected direction (0°). We computed this by finding the mean vector for across all n points using two metrics of circular data, the average preferred direction and the dispersion (Jammalamadaka and SenGupta 2001). Since these metrics only exhibit the desired behavior when r is positive, and because we were focused on fidelity as opposed to baseline offsets, we forced the minimum of the model-based representation to be 0 by adding an offset to every representation.

First, we found the mean direction of the spatial IEM-based representation (i.e., the angle of the vector) using the “circ_mean” function in the CircStats MATLAB toolbox (Berens 2009). We then computed how closely the vector points to the expected value of 0 by taking the cosine of this value.

$$\text{fidelity} = \max(r) * |\text{circ}_r(\theta, r)| * \cos(\text{circ_mean}(\theta, r)).$$

We then calculated the dispersion of the spatial representation (i.e., the vector length), which scales monotonically with the width of the function. As the spatial representation becomes narrower and more precise, the vector length increases. The dispersion is computed with “circ_r” (Berens 2009).

Finally, we multiplied the mean direction by the dispersion and scaled the result by the maximum value of the spatial representation, yielding a fidelity metric which monotonically increases as a given spatial representation increases in amplitude and in precision (Fig. 1B). The metric is negative when the mean vector points in the opposite direction (180°), and it is close to 0 if the spatial representation is flat and contains no information about the remembered location. Another way to understand this metric is to compare it to methods for measuring behavioral recall and precision based on error distributions. Here, we are instead quantifying the error of the reconstruction, with wider and lower reconstructions representing higher error, or worse fidelity. However, it is critical to note that a model-based reconstruction computed with the IEM approach is not analogous to a probability distribution.

As noted in the Statistical Analysis, we resampled across participants to generate CIs for the sampled population. To generate single-participant CIs, we bootstrapped the data by resampling across trials. If the 95% CIs overlapped with 0, then there was no significant memory representation for that participant or for that region. When we compared fidelity across memory tasks, we subtracted the two bootstrapped distributions (e.g., WM – LTM) and compared this difference distribution against 0.

We also computed a post-hoc power analysis to examine the ability of our sample to detect significant changes in representational fidelity. Details are in the Supplementary Material (Table S1, Fig. S2).

Support Vector Machine Classification

The analyses described above allowed us to generate continuous estimates of remembered stimuli to precisely compare the quality of spatial memory representations across ROIs and memory tasks. We were also interested in whether activation patterns in these ROIs could discriminate between the two memory tasks. To test this, we trained a support vector machine (SVM) with the default linear kernel to discriminate between the WM trials and LTM trials. This was done using MATLAB’s “fitclinear” routine.

We first divided the data into two temporal epochs during the memory delay period and averaged across the TRs in those epochs (early epoch: TRs 2–4, late epoch: TRs 5–7). While we equated the WM and LTM tasks as closely as possible, the WM task displayed an additional sensory stimulus that may increase the likelihood that a given ROI can classify a trial into either task. We hypothesized that this should only occur in the early epoch when the visually evoked hemodynamic response is at its peak. Classifier training and testing were performed separately for each epoch, ROI, and participant. To ensure that classification performance was not driven by a global difference in the univariate response between the memory tasks, we computed the global mean across all trials and voxels for each memory task. We then subtracted this constant from the data before we performed any of the SVM analyses below. This mean subtraction amounts to a baseline shift and thus has no effect on the variance of the data in each condition.

Mean classifier accuracy was determined by k -fold cross-validation ($k = 12$, which is the typical number of runs in each task condition) using MATLAB's "crossval" function. We then assessed statistical significance by randomizing the labels associated with each task and generated a null distribution of classifier accuracies over 1000 iterations.

To combine the data across participants, we took all classifier accuracies for a given ROI and temporal epoch (10 values, one for each participant) and calculated the mean. We then performed a similar operation on the permuted null distributions, concatenating all the distributions (10 participants $\times$ 1000 iterations) and calculating the mean across participants to generate a single null distribution for the mean.

We computed empirical one-tailed P values by comparing when the mean classifier performance was higher than the permuted null values [$1 - \text{mean}(\text{classAcc} > \text{nullDistr})$]. Note that we used one-tailed tests to reflect the fact that the classifier accuracies were not predicted to drop significantly below chance. These P values were subject to FDR correction across temporal epochs and ROIs. We used a stricter FDR $q = 0.025$ to reflect the fact that we only used one-tailed P values.

Code Accessibility

The code to analyze the behavioral data and the fMRI data (i.e., the IEM, SVM, and associated analyses) is available at <https://osf.io/fcp5y/>. This repository also includes the code for the training and scanner tasks.

Results

Behavioral Data

Outside of the scanner, participants were trained on both the LTM and WM tasks for at least 5 days, with the goal that mean recall error on the LTM retrieval task was stable at the end of training (mean number [95% CIs] of training days 7.39 [5.78, 9.22]). This training target did not enforce similar performance between tasks (participant averaged error on the last day: WM: 4.29° [3.58°, 5.12°]; LTM: 6.54° [4.69°, 8.52°]). To characterize each participant's recall success and precision, we plotted a histogram of their recall errors and fit the data with an additive mixture of a circular Gaussian and a uniform distribution (Zhang and Luck 2008; Harlow and Yonelinas 2016). The SD of the Gaussian is a measure of the spatial precision of their recall, with larger SDs corresponding to imprecise recall. The height of

the uniform distribution is a measure of their rate of forgetting. Here, we report the complement of this number, the probability of recall ($P(\text{recall})$). The interval training procedure combined with precise feedback significantly increased both the precision and probability of recall in the LTM task (difference between first and last set of trials: SD: 14.12° [2.88°, 37.23°], $P < 0.001$; $P(\text{recall})$: 0.05 [0.01, 0.10], $P = 0.01$). Although our training procedure did not require that performance was matched between the WM and LTM tasks, we nevertheless found that mean response error was comparable between the tasks during the sessions in the scanner (participant averaged WM error: 7.40 [5.67, 9.53]; LTM error: 7.78 [4.54, 11.88], $P = 0.78$; Bayes factor in favor of null $\text{BF}_{01} = 3.24$). The probability of recall was high ($P(\text{recall})$: WM 0.98 [0.96, 1.00], LTM: 0.97 [0.94, 1.00]; Fig. 3B, right) and similar between both memory tasks ($P = 0.35$; $\text{BF}_{01} = 2.40$). As a group, participants also had similar recall precision between both tasks (mixture model SD: WM 7.61 [6.56, 8.69], LTM: 7.91 [5.38, 10.81]; $P = 0.85$; $\text{BF}_{01} = 3.23$). However, recall precision varied more on the LTM task: Some participants had worse precision in the LTM task than the WM task, while others had better precision (Fig. 3B, left). We obtained similar results when we fit a simpler single parameter model (one-parameter model that only estimates circular Gaussian SD: WM: 11.73 [8.20, 15.96], LTM: 11.37 [6.22, 17.83]; $P = 0.78$; $\text{BF}_{01} = 3.06$). Model comparison with the Bayesian Information Criterion revealed that neither model was consistently better across the 10 participants (WM: mixture model better in four participants; LTM: mixture model better in five participants). Thus, the number of parameters in the models does not substantively alter our main conclusions.

Sensory-like Representations of Remembered Spatial Position

To assess how spatial information was encoded during the delay period of both memory tasks, we trained an encoding model using data from an independent perceptual task, which allowed us to map the spatial selectivity of visually responsive voxels in several retinotopic ROIs. Having trained the model of the independent perceptual task, we inverted the encoding model to reconstruct spatial representations of the remembered location from the pattern of activity across voxels in the WM or LTM tasks. If this reconstruction is highly accurate, then we can conclude that the pattern of voxel activations during the perceptual task is similar to the pattern of activations during the delay period of the memory tasks (i.e., it is sensory-like). This is a strong test of the cortical reinstatement hypothesis, which posits that the neural representation of an item during perception is reinstated when that item is recalled from memory. Furthermore, using an encoding model trained on a single independent task allows us to directly compare between spatial representations across WM and LTM (Sprague et al. 2018, 2019).

The spatial reconstructions show activation at each modeled location, and peaks within reconstructions can be interpreted as visual spatial memory representations that can be parameterized as a curve in circular (polar) co-ordinate space since the eccentricity of the stimuli was fixed (Fig. 1B). To average across trials with different remembered positions, we recentered the reconstruction on each trial so that the remembered position was at 0° . In Figure 4A, we plot these averaged spatial reconstructions from area V1 at several timepoints during the memory delay of both tasks. As expected, spatial information

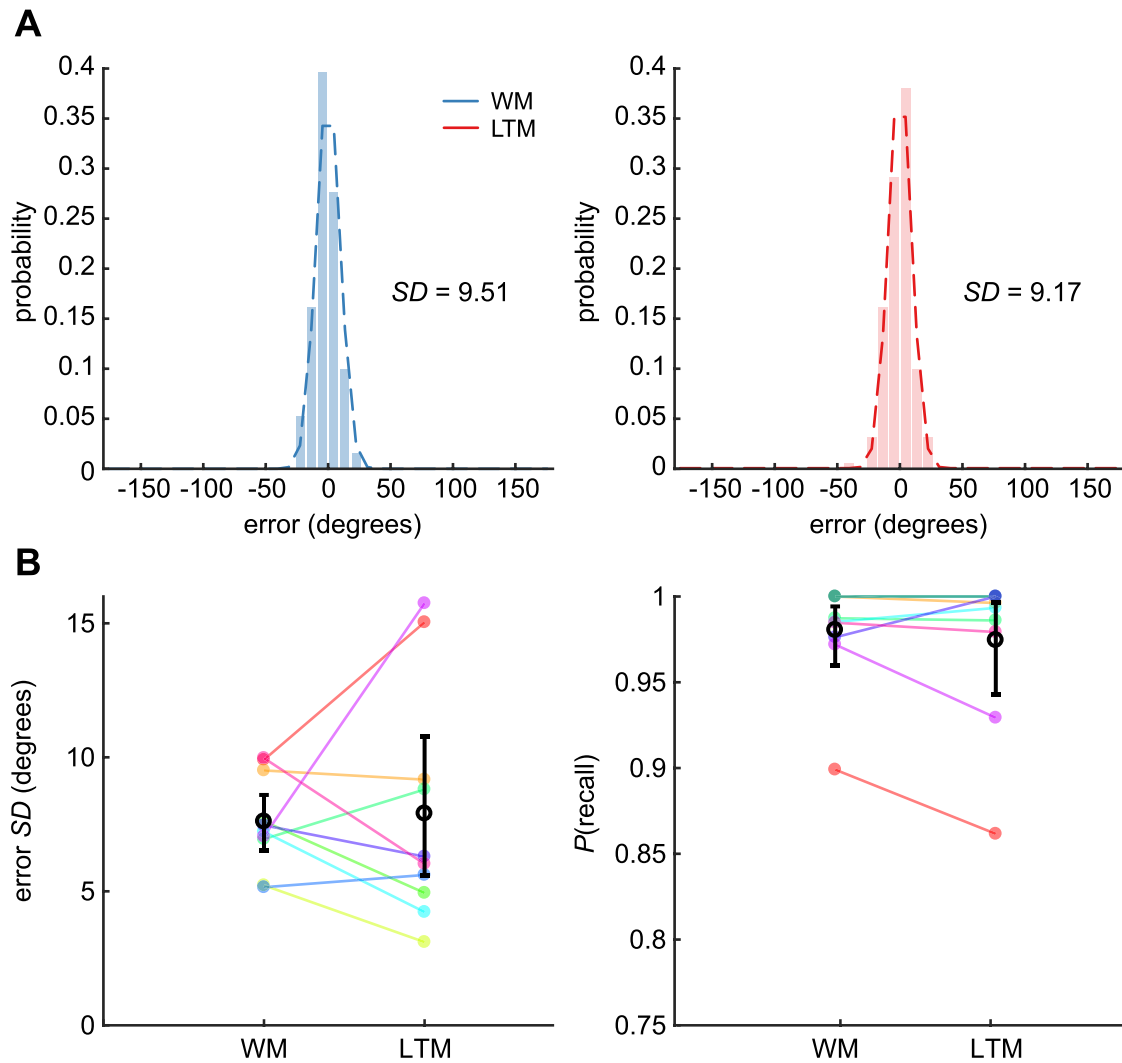


Figure 3. Memory recall across both scanner tasks is similar. (A) The distribution of memory recall errors for an example participant. We fit a mixture model to this distribution for every participant, where the Gaussian distribution characterized the variability of recall (SD) and the uniform distribution characterized the likelihood of recall ($P(\text{recall})$). (B) The mixture model fit parameters for each subject and task. The mean across participants and 95% CIs are shown in black.

during the delay period did not emerge until 4–6 s after stimulus onset, when the BOLD response evoked by a stimulus is expected to peak. Additionally, the spatial reconstruction was higher amplitude in the WM task than in the LTM task early in the delay period. This early difference is likely due to fact that the remembered position was marked by a dot in the WM task but not in the LTM task (Fig. 1A), which should evoke a sensory response. However, retrieval from LTM also occurs in the early TRs and may also contribute to this difference. Regardless, we found that late in the delay period (~10 s after stimulus onset), when lingering sensory activation should have decayed, we found that model-based spatial representations were similar between the WM and LTM tasks. For comparison to the univariate timecourses for each task, see Figure S1.

To quantify the model-based spatial representations, we computed a metric to measure the representational fidelity of each recentered spatial representation (see Methods). We illustrate some properties of this fidelity metric and its computation in Figure 1B (left panel). First, each point in the

spatial representation is replotted in polar space. We then find the mean vector of all of these points, which describes the average directional energy of the spatial representation. The fidelity, then, is proportional to the length of the vector along the expected direction (i.e., 0° , or the x-component of the vector; bold italic numbers in Fig. 1B, left panel). Thus, fidelity increases as a spatial representation increases in amplitude, and increases as dispersion—the width of the spatial representation—decreases.

We found that spatial position was represented with high fidelity during the delay period of both memory tasks in several retinotopic ROIs. Representational fidelity tended to peak earlier for the WM task (mean timepoint across ROIs: WM 6.89 s after stimulus presentation; LTM: 8.89 s). Again, this is likely due to sensory activity evoked by the small dot stimulus marking the to-be-remembered position in the WM task. However, the spatial representation persisted well into the delay period, after the initial sensory response would decay (Fig. 4B, V1–IPS2). In the lateral parietal ROIs (dLatIPS, vLatIPS, and AnG), there appeared

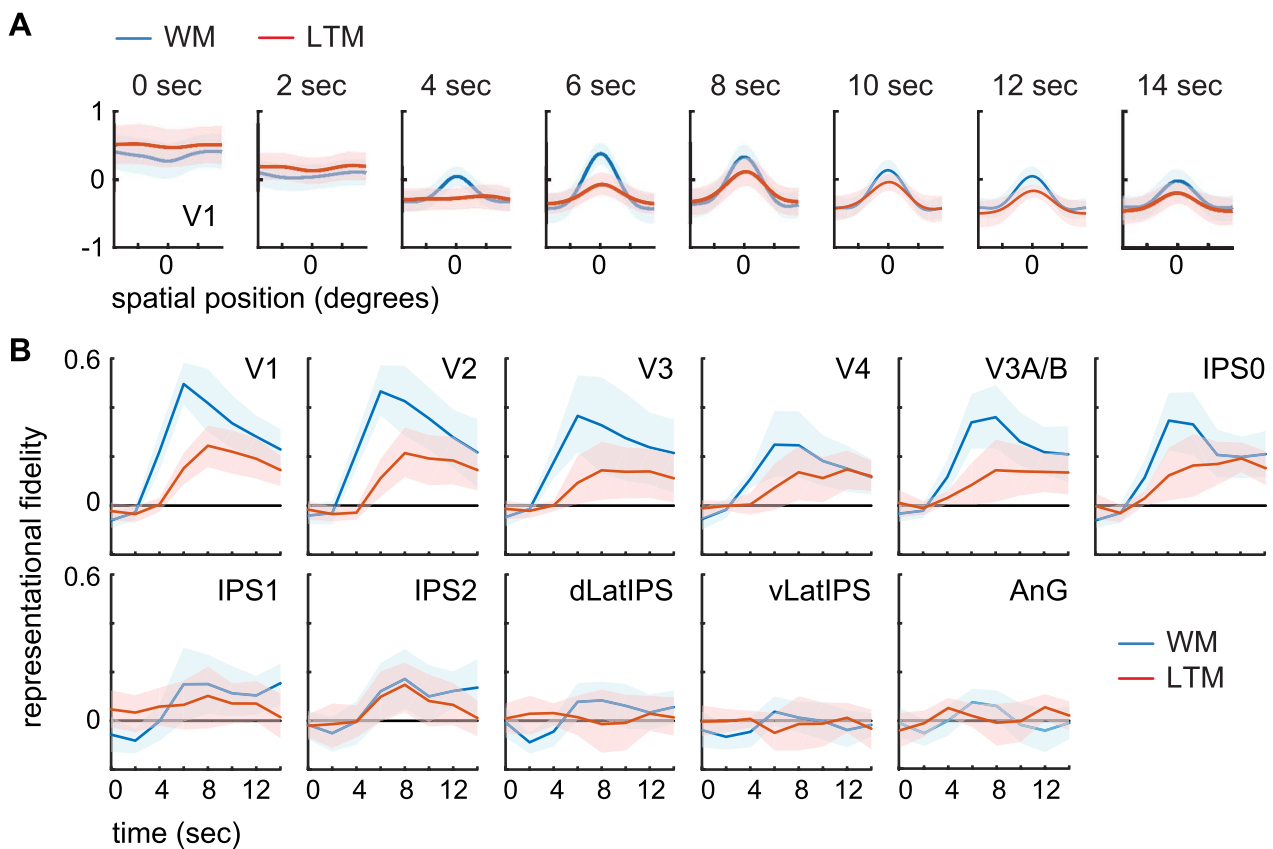


Figure 4. Model-based representations of remembered spatial positions over time for both memory tasks. Error bars are participant-resampled 95% CIs. (A) The timecourse of model-based spatial representations using V1 data, averaged across participants (0 s is stimulus onset). Remembered position was represented similarly for WM and LTM late in the delay period (8–14 s). (B) Representational fidelity timecourses for all retinotopic areas we analyzed. CIs that intersect with 0 are not significant. dLatIPS, dorsolateral IPS; vLatIPS, ventrolateral IPS; AnG, AnG.

to be no significant evidence of sensory reinstatement across the entire delay.

In the remainder of our analyses, we report results from data averaged over TRs either early in the delay period (TRs 2–4, or 2–6 s after stimulus presentation) or late in the delay period (TRs 5–7, or 8–12 s). **Figure 5A** summarizes representational fidelity late in the delay period (comparable to **Fig. 4B**). As we saw previously, the retinotopic ROIs show evidence of sensory reinstatement, while the LPC ROIs do not. To test this, we first ran a two-way repeated-measures ANOVA with ROI and memory task as factors and determined statistical significance by comparing the F-scores to a permuted null distribution. We found a main effect of ROI ($P < 0.001$) and a significant interaction ($P = 0.003$) but no significant difference between memory tasks ($P = 0.191$). This interaction suggests that some ROIs have a larger difference between memory tasks than others (e.g., higher WM fidelity in V1 – V3/V3AB, **Fig. 5A**) even though the average difference between tasks is insignificant.

Bayesian t-tests on the data from the late delay period revealed that most ROIs had weak-to-moderate evidence for the hypothesis that there was no difference between WM and LTM (BF_{01} for V3: 1.47, V4: 2.25, IPS0: 1.80, IPS1 2.18, IPS2 2.89, dLatIPS 2.16, vLatIPS 3.24, AnG 2.92), while in the remaining ROIs, there was weak evidence for the alternative hypothesis that one condition has higher fidelity (BF_{01} for V1: 0.35, V2: 0.77, V3AB: 0.45; [Jeffreys 1961](#); [Schönbrodt and Wagenmakers 2018](#)).

Overall, we do not find compelling evidence of a difference in sensory reinstatement between model-based representations during WM and LTM.

One possible explanation for this null result in LPC ROIs is that there was poor signal in the sensory mapping task, and this led to unstable estimates of the spatial selectivity of each voxel that did not generalize well to the memory tasks. To evaluate this possibility, we used a leave-one-run-out cross-validation procedure to train and test models based only on data from the sensory mapping task. In this analysis, as shown in **Figure 5B**, we found that activation patterns in all retinotopic and non-retinotopic ROIs except vLatIPS contained information about the viewed position of the mapping stimulus (all significant P s < 0.05). Thus, even the nonretinotopically organized areas dLatIPS and AnG encoded information about the location of a continuously visible stimulus, suggesting that a failure to find information about remembered positions was not due to poor SNR in the mapping task. Note that this is contrast to prior work using a different method to measure feature-level information—using correlative methods, they failed to find significant feature-level information in any lateral parietal region during their perceptual task ([Favila et al. 2018](#)).

In summary, we assessed the fidelity of memory representations using an encoding model that was trained on sensory-evoked responses from an independent task. We found that activation patterns in retinotopic ROIs contained information

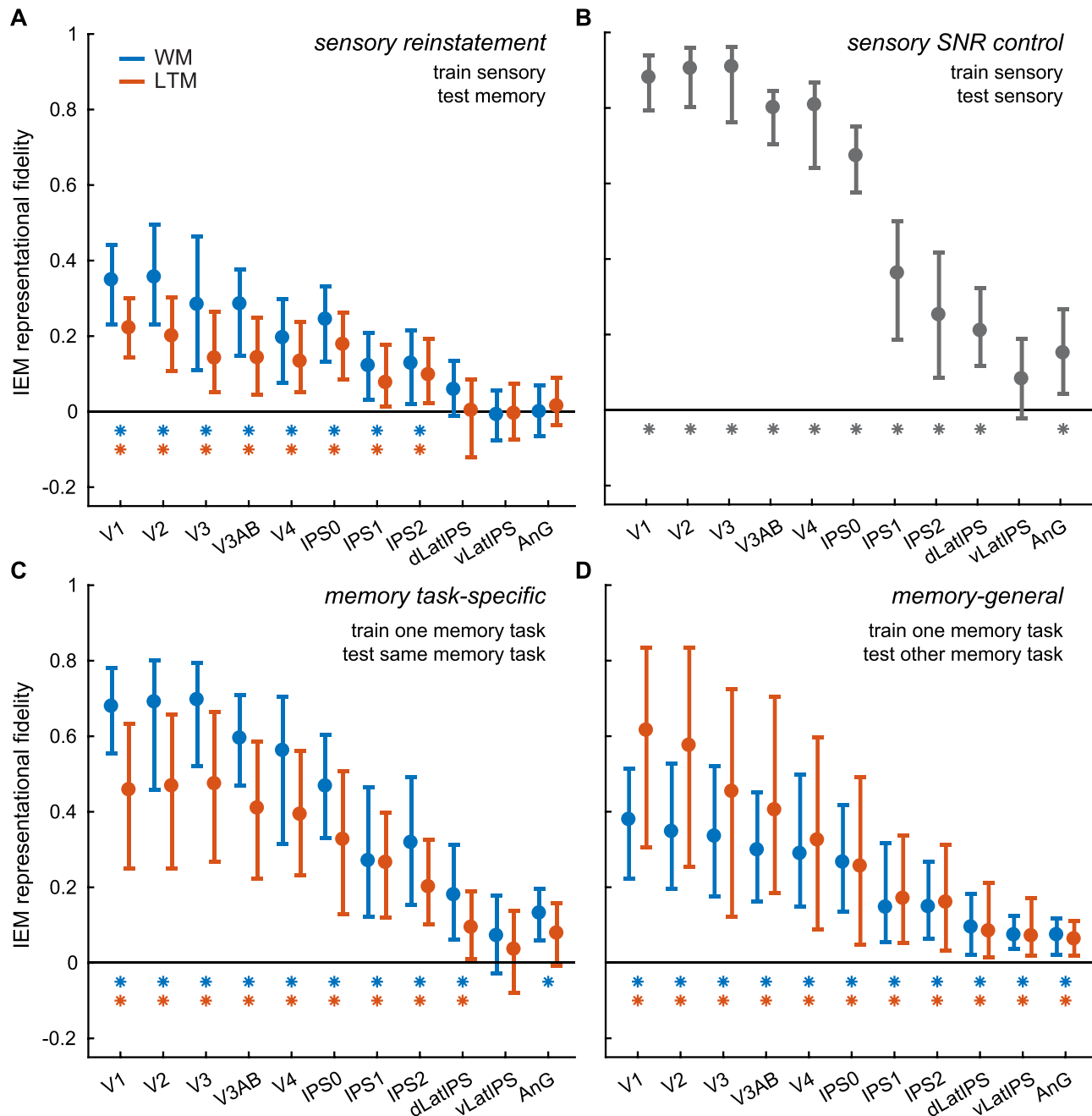


Figure 5. Model-based representations for different combinations of training and testing data, taken from the end of the delay period (8–12 s). For each plot, we ran a permuted two-way ANOVA of ROI by memory task (or for the sensory control, a one-way ANOVA across ROIs). (A) Train on the sensory task and test memory, as a test of sensory reinstatement. (B) As a control, train and test within the sensory task. (C) Train and test within each memory task, to examine task-specific representations. (D) Cross-train memory tasks, e.g. train WM and test LTM, to examine memory representations that are general across tasks. In all four cases, there was a significant main effect of ROI ($P < 0.001$). We found a significant two-way interaction for both the independent training set ($P = 0.003$) and the cross-training procedure ($P = 0.001$), but no main effect of task. Training within memory tasks yielded a borderline significant effect of memory task ($P = 0.052$), as the WM task generally resulted in higher fidelity than the LTM task. Asterisks indicate significant information after FDR correction across all ROIs and both memory tasks, $q = 0.05$.

about the remembered spatial position in both the WM and LTM tasks, and there was little evidence for a difference in representational fidelity between the memory tasks after ~8 s into the delay period. This is a compelling test of the reinstatement hypothesis because it demonstrates fMRI activation patterns, evoked by viewing and attending a spatial position, have a strong

overlap with activation patterns elicited by remembering the same spatial position. By contrast, nonretinotopically organized ROIs in parietal cortex did not contain information about remembered positions in either memory task when the model was trained on the sensory localizer and tested on data from each memory task. Taken together, these data imply that

information reinstated from both WM and LTM is encoded in a sensory-like code, but only in retinotopically organized regions of occipital and parietal cortex.

Model Cross-Generalization to Compare Sensory and Mnemonic Representations in Retinotopic and Nonretinotopic ROIs

The presence of spatial information in the LPC ROIs during a perceptual task raises the possibility that they do encode spatial information in the memory tasks but in a different format than the sensory evoked code.

To evaluate this possibility, we used a balanced cross-validation procedure to train and test the spatial IEM separately within each memory task (Fig. 5C). A permuted two-way ANOVA of ROI by task only showed a significant effect of ROI (ROI: $P < 0.001$; task: $P = 0.052$; interaction: $P = 0.071$). We then tested each ROI and task individually as before (FDR correcting over all comparisons, $q = 0.05$). All of the retinotopically organized ROIs contained information about the remembered spatial position in both the WM and the LTM tasks (all P s < 0.005). However, we also found evidence that activation patterns in dLatIPS and AnG also represented the remembered spatial position in both tasks (dLatIPS WM: $P = 0.002$, LTM: $P = 0.027$; AnG WM: $P < 0.001$, LTM: $P = 0.072$). By contrast, vLatIPS did not contain position-specific representations of the remembered position (WM: $P = 0.165$, LTM: $P = 0.549$). The observation of information about remembered position when training and testing was carried out separately for each memory task—but not when we trained on the sensory localizer—suggests that dLatIPS and AnG encode mnemonic information in a format that differs from the sensory evoked response patterns.

Finally, we also performed cross-training between the memory tasks (i.e., train WM, test LTM; train LTM, test WM) to more directly estimate the extent to which memory-related activation patterns were shared across the two tasks (Fig. 5D). The two-way ANOVA showed a main effect of ROI and a significant interaction (ROI: $P < 0.001$; task: $P = 0.211$; interaction: $P < 0.001$). Similar to the interaction effect when training on the independent sensory task, the interaction appeared to be driven by higher fidelity when training on the WM task and testing on the LTM task in the early visual areas (red dots; Fig. 5D). When we tested each ROI and task individually, we found significant generalization between the WM and LTM tasks in all of the retinotopic ROIs and in dLatIPS and AnG (all P s < 0.05). Interestingly, we also found significant information in the cross-training analysis in vLatIPS for both memory tasks (WM: $P < 0.001$, LTM: $P = 0.003$). This was surprising given that vLatIPS did not contain information about either task when training was carried out on the sensory mapping task or when training and testing was done entirely within task. It is possible that this is due to noisy data, especially because training and testing within-task means that the model is trained on fewer trials. It is also possible that vLatIPS activation patterns may be representing other information, such as features of the clip art item (which could be task-relevant, in the LTM task, or task-irrelevant). Some have reported evidence that cues may have a significant effect on activation patterns during memory retrieval (Xiao et al. 2017). Furthermore, there is other evidence that vLatIPS represents feature information that is irrelevant to a memory judgment (Favila et al. 2018). While these studies relied on

correlation methods to compute information scores, it is still possible that such factors would impact our IEM analysis. In particular, if some cue features are not well randomized with respect to spatial position, the IEM may overfit to feature-level responses.

Overall, these findings suggest that dLatIPS and AnG encode mnemonic information for spatial position during WM and LTM in a different format than the original perceptual representation. Moreover, the WM/LTM cross-training analysis suggests that the representational format of remembered information in dLatIPS, vLatIPS, and AnG is at least partially common to both WM and LTM.

Task Classification

In the analyses above, we used different IEM training and testing schemes to test whether the format of reinstated memories was shared between WM and LTM. Next, we examined a related but distinct question: Is there a discriminable difference between the activation patterns elicited by WM recall and LTM recall? ROIs that share representational codes may still contain encode some information unique to the memory task. To test this possibility, we trained a binary SVM to classify each memory trial as belonging to the WM or LTM task. The SVM was trained using data balanced across spatial position and task.

We ran the SVM classifier for both the early and late temporal epochs as defined above (early epoch: TRs 2–4, late epoch: TRs 5–7). We found that we could discriminate between the two tasks early in the delay period across all retinotopic ROIs and in both dLatIPS and vLatIPS (Fig. 6). We believe this effect is largely driven by the presence of a sensory stimulus in the WM task but not in the LTM task. Note that successful SVM discrimination cannot be driven by differences in univariate delay period activity because we removed the mean within each memory task before training and testing the classifier (see Methods). However, we also found that early visual areas (V1–V4) failed to discriminate between tasks in the late delay period. Instead, only activation patterns in retinotopic parietal regions discriminated between the WM and LTM tasks during the late temporal epoch. Activation patterns in the dorsal and ventral lateral IPS also discriminated between tasks during the late temporal epoch. Lastly, we found that activation patterns in AnG were unable to discriminate between tasks entirely.

To evaluate the differential classification accuracy across ROIs, we ran a two-way repeated-measures ANOVA of ROI by temporal epoch on classifier accuracy and calculated P values as previously described. Both main effects of ROI and epoch were significant (P s < 0.001), and there was a significant interaction between the two ($P = 0.001$). Since AnG did not show any ability to discriminate between task, we reran the ANOVA without this ROI. While there was no main effect of ROI ($P = 0.076$), the effect of epoch ($P < 0.001$) and the interaction ($P = 0.01$) remained significant. To test the hypothesis that the interaction was driven by the difference in discriminability between epochs, we ran post hoc one-way ANOVAs of ROI on classifier accuracy separately within each epoch (main effect of ROI on early epoch: $P = 0.065$; late epoch: $P = 0.016$). Since all ROIs were able to discriminate task in the early epoch, we saw no significant effect of ROI there. These data suggest that although the representation of the remembered location is similar between the WM and LTM

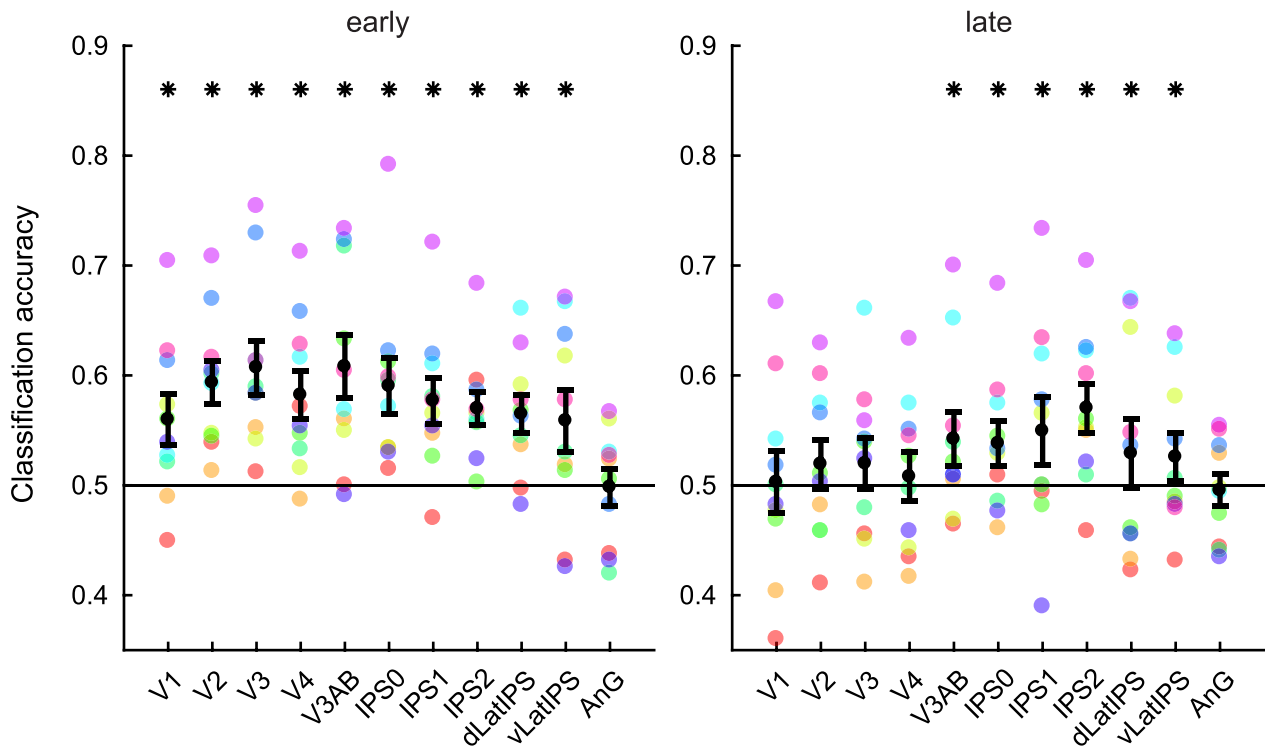


Figure 6. Task decoding in the early and late epochs of the delay period in each ROI. Colored dots show data from each individual subject, and error bars are 95% CIs estimated with bootstrapping. Classifier accuracy is significantly above chance (*) if it passes a one-tailed a permutation test, FDR-corrected across ROI and temporal epoch ($q = 0.025$). Unlike the occipital ROIs, the activation patterns in regions of parietal cortex (V3AB, IPS0, etc.) resulted in above chance decoding accuracy in the late epoch.

tasks, in most ROIs that we examined, information that differentiates the two tasks is selectively present in both retinotopic and lateral parietal regions.

Discussion

The present study, in which we examined WM and LTM using an experimental design that closely matched behavioral precision, makes three main contributions. First, we demonstrate that the fidelity of sensory-like representations is similar between both WM and LTM in both occipital (V1–V4) and parietal (V3AB, IPS0–2) retinotopic regions. Second, we provide evidence that while lateral parietal regions do not reinstate information in a sensory-like format, they do encode spatial memories in a format that is partially shared between the WM and LTM tasks. Third, we show that activity patterns in parietal regions, but not early visual areas, contained information about whether subjects were maintaining information encoded into WM or retrieved from LTM. We propose that these parietal areas may be jointly coding information about the content and the source of visual memories.

The results that directly compare between WM and LTM representations have implications for theories about how memory systems interact with one another. Many theories posit that LTM retrieval places remembered content directly into WM. That is, both WM and actively retrieved LTMs are maintained via the same processes, a capacity-limited state called the “focus of attention” (Atkinson and Shiffrin 1968; Cowan 1995; LaRocque et al. 2014; D’Esposito and Postle 2015; Fukuda and Woodman 2017). Our finding that WM and LTM representations

are both represented in the same retinotopic regions in the same sensory-like format provides compelling evidence for this account.

This theory is also consistent with our finding that lateral parietal regions encode mnemonic information in a format that was shared between WM and LTM tasks. We could robustly reconstruct the remembered position from dLatIPS activity when we trained and tested the model within or across memory tasks. The pattern of results was similar in AnG, although only modest task-specific mnemonic information was observed in the within-task training/testing analysis (see Table 1; Fig. 5C). By contrast, we only observed spatial information in vLatIPS when models were trained on one memory condition and tested on the other. Together, these results support a general role for these regions in representing spatial information in both WM and LTM. However, since the format is not sensory-like, it suggests that the information is somehow reformatted or transformed. This is supported by recent evidence (Xiao et al. 2017; Favila et al. 2018; Bainbridge et al. 2021) that has led some to argue that these mnemonic representations are spatially transformed versions of perceptual representations (Favila et al. 2020). This transformation may be functional, ensuring that information in memory is distinguishable from incoming sensory information (Bettencourt and Xu 2015; Stokes 2015; Xu 2017). We note this is purely speculative in the absence of additional experimental data evaluating memory fidelity in the face of different kinds of competing distractors (i.e., Does the degree to which the codes in LPC are transformed away from sensory codes predicts success in the face of concurrent sensory distractors?).

Table 1 Summary of our findings, separated by the major region groupings (retinotopic occipital, retinotopic parietal, and lateral parietal)

Method	Test of representation	Occipital	Retinotopic parietal	Lateral parietal		
		V1–V4	V3A/B, IPS0-2	dLatIPS	vLatIPS	AnG
IEM	Sensory info (control)	1	1	1	0	1
IEM	Sensory-like reinstatement	1	1	0	0	0
IEM	Task-specific	1	1	1	0	~1
IEM	Task-general (cross-training)	1	1	1	1	1
SVM	Task discriminable	0	1	1	1	0

Note: A “1” indicates a significant effect and a “0” indicates a nonsignificant effect. A tilde indicates that the effect was not consistent across the WM and LTM tasks.

We also found that activity patterns in almost all parietal ROIs, but not occipital ROIs, could decode whether a trial belonged to the WM or LTM task (Fig. 6). The one exception was AnG, which could not discriminate between the two memory tasks. We discuss three possible reasons we observed this finding in the remaining parietal ROIs. One possibility is that parietal regions encode the task set (i.e., higher-level representations that govern the execution of a specific task). Indeed, past work has shown that parietal regions encode higher-order task representations, such as the task rule that the subject is executing (Bode and Haynes, 2009; Woolgar et al., 2011a, 2011b). The second possibility is that the WM and LTM tasks require different attentional processes. Parietal regions have been reported to be engaged during attention tasks, WM tasks, and LTM tasks (Awh and Jonides 2001; Awh et al. 2006; Cabeza et al. 2008; Hutchinson et al. 2014; Sestieri et al. 2017). Researchers have separately proposed that the parietal mechanisms involved in spatial attention also support spatial WM (Awh and Jonides 2001; Awh et al. 2006) and that parietal regions mediate attention to long-term memories during retrieval (Cabeza et al. 2008; Sestieri et al. 2017). The ability of the SVM decoder to distinguish between WM and LTM tasks in parietal regions raises the possibility that these attention-mediated memory processes may not be identical across tasks. The last possibility is that parietal regions encode the “source” of memory representations (perception for WM, and an internal store for LTM). One issue with relying solely on cortical reinstatement for memory recall is that multiple sensory representations in one cortical region, such as area V1, could be attributable to ongoing perception, recent perception (WM), or distant perception (LTM). In order to maintain separable representations within sensory areas, parietal regions may additionally represent the source of sensory information. This could help protect reinstated sensory memories from interference by concurrent sensory input (Bettencourt and Xu 2015; Rademaker et al. 2019). Further work is needed to test these hypotheses.

Finally, we note that our investigation of memory representations is limited to visual and parietal cortex and does not provide any evidence about the differential role of other areas, such as the medial temporal lobe (MTL), which is thought to play a role in both LTM and WM (Squire and Wixted 2011). Several studies have shown that the hippocampus and surrounding cortical regions exhibit encoding-retrieval reinstatement, both for stimulus categories (Polyn et al. 2005; Liang and Preston 2017; Schultz et al. 2019; Bainbridge et al. 2021) and for individual episodic events (Staresina et al. 2012; Tompary et al. 2016) across a variety of different recall tasks that may involve both WM and LTM to varying degrees. Although the MTL has long been linked

with LTM (Scoville and Milner 1957; Mishkin 1978; Squire and Zola-Morgan 1978), a review of existing evidence suggests that the hippocampus is necessary for some types of WM tasks—perhaps specifically when they are complex (Yonelinas 2013) or spatial tasks (Nadel and Hardt 2011). Future studies that are optimized for imaging and segmenting the MTL could use similar techniques as the current study to compare memory reinstatement and retrieval across WM and LTM in MTL regions.

In conclusion, our findings suggest that remembered spatial positions encoded into WM or retrieved from LTM are represented similarly in occipital and parietal cortex. In retinotopic regions, activity patterns recapitulated those that were observed during an independent sensory stimulation condition, showing that a sensory recruitment model provides a useful perspective for how visual details are maintained regardless of the specific memory system. In nonretinotopic areas of LPC, mnemonic representations appear to share a sensory format across WM and LTM that is distinct from a sensory-like code. Moreover, parietal regions—both retinotopic and lateral—also represented the memory task being performed, unlike retinotopic occipital regions.

Supplementary Material

Supplementary material can be found at Cerebral Cortex online.

Funding

National Institutes of Health MH087214 (to E.A.), EY025872 (to J.T.S.), and F32-EY028438 (to T.C.S.); National Science Foundation GRFP (to V.A.V.); Sloan Research Fellowship (to T.C.S.).

Notes

We thank Megan deBettencourt and Nuttida Rungratsameeta-weemana for helpful discussions. Conflict of Interest: None declared.

References

Atkinson RC, Shrifin RM. 1968. Human memory: a proposed system and its control processes. In: Spence KW, Spence JT, editors. *Psychology of learning and motivation*. New York: Academic Press, pp. 89–195.

Awh E, Jonides J. 2001. Overlapping mechanisms of attention and spatial working memory. *Trends Cogn Sci*. 5: 119–126.

Awh E, Vogel EK, Oh SH. 2006. Interactions between attention and working memory. *Neuroscience*. 139:201–208.

- Baddeley A. 2000. The episodic buffer: A new component of working memory? *Trends Cogn Sci*. 4:417–423.
- Bainbridge WA, Hall EH, Baker CI. 2021. Distinct representational structure and localization for visual encoding and recall during visual imagery. *Cereb Cortex*. 31:1898–1913.
- Benjamini Y, Yekutieli D. 2001. The control of the false discovery rate in multiple testing under dependency. *Ann Stat*. 29:1165–1188.
- Berens P. 2009. CircStat: a MATLAB toolbox for circular statistics. *J Stat Softw*. 31:1–21. <https://doi.org/10.18637/jss.v031.i10>.
- Bettencourt KC, Xu Y. 2015. Decoding the content of visual short-term memory under distraction in occipital and parietal areas. *Nat Neurosci*. 19:150–157.
- Bode S, Haynes J-D. 2009. Decoding sequential stages of task preparation in the human brain. *Neuroimage*. 45:606–613.
- Bosch SE, Jehee JFM, Fernández G, Doeller CF. 2014. Reinstatement of associative memories in early visual cortex is signaled by the hippocampus. *J Neurosci*. 34:7493–7500.
- Brouwer GJ, Heeger DJ. 2009. Decoding and reconstructing color from responses in human visual cortex. *J Neurosci*. 29:13992–14003.
- Cabeza R, Ciaramelli E, Olson IR, Moscovitch M. 2008. The parietal cortex and episodic memory: an attentional account. *Nat Rev Neurosci*. 9:613–625.
- Christophel TB, Jamshchinina P, Yan C, Allefeld C, Haynes J-D. 2018. Cortical specialization for attended versus unattended working memory. *Nat Neurosci*. 21:494–496.
- Cowan N. 1995. *Attention and memory: an integrated framework*. New York: Oxford University Press.
- D'Esposito M, Postle BR. 2015. The cognitive neuroscience of working memory. *Annu Rev Psychol*. 66:115–142.
- Davison AC, Hinkley DV. 1997. *Bootstrap methods and their application*. Cambridge: Cambridge University Press.
- Emrich SM, Riggall AC, Larocque JJ, Postle BR. 2013. Distributed patterns of activity in sensory cortex reflect the precision of multiple items maintained in visual short-term memory. *J Neurosci*. 33:6516–6523.
- Ester EF, Anderson DE, Serences JT, Awh E. 2013. A neural measure of precision in visual working memory. *J Cogn Neurosci*. 25:754–761.
- Ester EF, Sprague TC, Serences JT. 2015. Parietal and frontal cortex encode stimulus-specific mnemonic representations during visual working memory. *Neuron*. 87:893–905.
- Ester EF, Sutterer DW, Serences JT, Awh E. 2016. Feature-selective attentional modulations in human frontoparietal cortex. *J Neurosci*. 36:8188–8199.
- Favila SE, Lee H, Kuhl BA. 2020. Transforming the concept of memory reactivation. *Trends Neurosci*. 43:939–950.
- Favila SE, Samide R, Sweigart SC, Kuhl BA. 2018. Parietal representations of stimulus features are amplified during memory retrieval and flexibly aligned with top-down goals. *J Neurosci*. 38:7809–7821.
- Fukuda K, Woodman GF. 2017. Visual working memory buffers information retrieved from visual long-term memory. *Proc Natl Acad Sci*. 114:5306–5311.
- Harlow IM, Yonelinas AP. 2016. Distinguishing between the success and precision of recollection. *Memory*. 24:114–127.
- Harrison SA, Tong F. 2009. Decoding reveals the contents of visual working memory in early visual areas. *Nature*. 458:632–635.
- Henderson MM, Serences JT. 2019. Human frontoparietal cortex represents behaviorally-relevant target status based on abstract object features. *J Neurophysiol*. 121:1410–1427.
- Hesterberg T. 2011. Bootstrap. *Wiley Interdiscip Rev Comput Stat*. 3:497–526.
- Hesterberg T. 2014. What teachers should know about the bootstrap: resampling in the undergraduate statistics curriculum. arXiv:1411.5279.v1.
- Hutchinson JB, Uncapher MR, Weiner KS, Bressler DW, Silver MA, Preston AR, Wagner AD. 2014. Functional heterogeneity in posterior parietal cortex across attention and episodic memory retrieval. *Cereb Cortex*. 24:49–66.
- Jamshchinina P, Christophel TB, Gayet S, Rademaker RL. 2021. Essential considerations for exploring visual working memory storage in the human brain. *Vis cogn*. 0:1–12.
- Jammalamadaka SR, SenGupta A. 2001. *Topics in circular statistics, series on multivariate analysis*. Singapore: World Scientific Publishing Co.
- Jeffreys H. 1961. *Theory of probability*. 3rd ed. Oxford: Oxford University Press, Clarendon Press.
- Kosslyn SM. 1980. *Image and mind*. Cambridge (MA): Harvard University Press.
- Kriegeskorte N, Douglas PK. 2019. Interpreting encoding and decoding models. *Curr Opin Neurobiol*. 55:167–179.
- Kuhl BA, Chun MM. 2014. Successful remembering elicits event-specific activity patterns in lateral parietal cortex. *J Neurosci*. 34:8051–8060.
- LaRocque JJ, Lewis-Peacock JA, Postle BR. 2014. Multiple neural states of representation in short-term memory? It's a matter of attention. *Front Hum Neurosci*. 8:5.
- Lee H, Chun MM, Kuhl BA. 2017. Lower parietal encoding activation is associated with sharper information and better memory. *Cereb Cortex*. 27:2486–2499.
- Lee S-H, Kravitz DJ, Baker CI. 2013. Goal-dependent dissociation of visual and prefrontal cortices during working memory. *Nat Neurosci*. 16:997–999.
- Lewis-Peacock JA, Postle BR. 2008. Temporary activation of long-term memory supports working memory. *J Neurosci*. 28:8765–8771.
- Liang JC, Preston AR. 2017. Medial temporal lobe reinstatement of content-specific details predicts source memory. *Cortex*. 91:67–78.
- Manly BFJ. 2007. *Randomization, bootstrap, and Monte Carlo methods in biology*. 3rd ed. London: Chapman and Hall.
- Mishkin M. 1978. Memory in monkeys severely impaired by combined but not by separate removal of amygdala and hippocampus. *Nature*. 273:297–298.
- Nadel L, Hardt O. 2011. Update on memory systems and processes. *Neuropsychopharmacol Rev*. 36:251–273.
- Polyn SM, Natu VS, Cohen JD, Norman KA. 2005. Category-specific cortical activity precedes retrieval during memory search. *Science (80-)*. 310:1963–1966.
- Rademaker RL, Chunharas C, Serences JT. 2019. Coexisting representations of sensory and mnemonic information in human visual cortex. *Nat Neurosci*. 22:1336–1344.
- Riggall AC, Postle BR. 2012. The relationship between working memory storage and elevated activity as measured with functional magnetic resonance imaging. *J Neurosci*. 32:12990–12998.
- Ritchey M, Wing EA, LaBar KS, Cabeza R. 2013. Neural similarity between encoding and retrieval is related to memory via hippocampal interactions. *Cereb Cortex*. 23:2818–2828.
- Rouder JN, Speckman PL, Sun D, Morey RD, Iverson G. 2009. Bayesian t tests for accepting and rejecting the null hypothesis. *Psychon Bull Rev*. 16:225–237.

- Schneegans S, Bays PM. 2016. No fixed item limit in visuospatial working memory. *Cortex*. 83:181–193.
- Schönbrodt FD, Wagenmakers EJ. 2018. Bayes factor design analysis: planning for compelling evidence. *Psychon Bull Rev*. 25:128–142.
- Schultz H, Tibon R, Larocque KF, Gagnon SA, Wagner AD, Staesina BP. 2019. Content tuning in the medial temporal lobe cortex: voxels that perceive, retrieve. *eNeuro*. 6:1–11.
- Scoville WB, Milner B. 1957. Loss of recent memory after bilateral hippocampal lesions. *J Neurol Neurosurg Psychiatry*. 20:11–21.
- Serences JT. 2016. Neural mechanisms of information storage in visual short-term memory. *Vision Res*. 128:53–67.
- Serences JT, Ester EF, Vogel EK, Awh E, Serences J. 2009. Stimulus-specific delay activity in human primary visual cortex. *Psychol Sci*. 20:207–214.
- Sestieri C, Shulman GL, Corbetta M. 2017. The contribution of the human posterior parietal cortex to episodic memory. *Nat Rev Neurosci*. 18:183–192.
- Sprague TC, Adam KCS, Foster JJ, Rahmati M, Sutterer DW, Vo VA. 2018. Inverted encoding models assay population-level stimulus representations, not single-unit neural tuning. *eNeuro*. 5:1–5.
- Sprague TC, Boynton GM, Serences JT. 2019. The importance of considering model choices when interpreting results in computational neuroimaging. *eNeuro*. 6:1–11.
- Sprague TC, Ester EF, Serences JT. 2014. Reconstructions of information in visual spatial working memory degrade with memory load. *Curr Biol*. 24:2174–2180.
- Sprague TC, Ester EF, Serences JT. 2016. Restoring latent visual working memory representations in human cortex. *Neuron*. 91:694–707.
- Sprague TC, Serences JT. 2013. Attention modulates spatial priority maps in the human occipital, parietal and frontal cortices. *Nat Neurosci*. 16:1879–1887.
- Squire LR, Zola-Morgan S. 1978. The medial temporal lobe memory system. *Nature*. 253:1380–1386.
- Staesina BP, Henson RNA, Kriegeskorte N, Alink A. 2012. Episodic reinstatement in the medial temporal lobe. *J Neurosci*. 32:18150–18156.
- Stokes MG. 2015. ‘Activity-silent’ working memory in prefrontal cortex: a dynamic coding framework. *Trends Cogn Sci*. 19:394–405.
- Sutterer DW, Awh E. 2015. Retrieval practice enhances the accessibility but not the quality of memory. *Psychon Bull Rev*. 23:831–841.
- Sutterer DW, Foster JJ, Serences JT, Vogel EK, Awh E. 2019. Alpha-band oscillations track the retrieval of precise spatial representations from long-term memory. *J Neurophysiol*. 122:539–551.
- Tompary A, Duncan K, Davachi L. 2016. High-resolution investigation of memory-specific reinstatement in the hippocampus and perirhinal cortex. *Hippocampus*. 26:995–1007.
- Wolff MJ, Jochim J, Akyurek EG, Stokes MG. 2017. Dynamic hidden states underlying working-memory guided behavior. *Nat Neurosci*. 20:864–871.
- Woolgar A, Hampshire A, Thompson R, Duncan J. 2011a. Adaptive coding of task-relevant information in human frontoparietal cortex. *J Neurosci*. 31:14592–14599.
- Woolgar A, Thompson R, Bor D, Duncan J. 2011b. Multi-voxel coding of stimuli, rules, and responses in human frontoparietal cortex. *NeuroImage*. 56:744–752.
- Xiao X, Dong Q, Gao J, Men W, Poldrack RA, Xue G. 2017. Transformed neural pattern reinstatement during episodic memory retrieval. *J Neurosci*. 37:2986–2998.
- Xu Y. 2017. Reevaluating the sensory account of visual working memory storage. *Trends Cogn Sci*. 21:794–815.
- Xue G. 2018. The neural representations underlying human episodic memory. *Trends Cogn Sci*. 22:544–561.
- Yeo TBT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zöllei L, Polimeni JR et al. 2011. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol*. 106:1125–1165.
- Yonelinas AP. 2013. The hippocampus supports high-resolution binding in the service of perception, working memory and long-term memory. *Behav Brain Res*. 254:34–44.
- Zhang W, Luck SJ. 2008. Discrete fixed-resolution representations in visual working memory. *Nature*. 453:233–235.