J Neurophysiol 121: 1222–1231, 2019.
First published March 12, 2019; doi:10.1152/jn.00062.2019.

RESEARCH ARTICLE | *Higher Neural Functions and Behavior*

Overlapping and distinct contributions of stimulus location and of spatial context to nonspatial visual short-term memory

Ying Cai,[1,2] Andrew D. Sheldon,[3] Qing Yu,[2] and Bradley R. Postle[2,4]

[1]*State Key Laboratory of Cognitive Neuroscience and Learning, and IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, People's Republic of China;* [2]*Department of Psychiatry, University of Wisconsin–Madison, Madison, Wisconsin;* [3]*Medical Scientist Training Program and Neuroscience Training Program, University of Wisconsin–Madison, Madison, Wisconsin; and* [4]*Department of Psychology, University of Wisconsin–Madison, Madison, Wisconsin*

Submitted 29 January 2019; accepted in final form 31 January 2019

Cai Y, Sheldon AD, Yu Q, Postle BR. Overlapping and distinct contributions of stimulus location and of spatial context to nonspatial visual short-term memory. *J Neurophysiol* 121: 1222–1231, 2019. First published March 12, 2019; doi:10.1152/jn.00062.2019.—Stimulus location is not always informative during visual short-term memory (VSTM) for nonspatial features. Nevertheless, there is considerable evidence for the automatic encoding and retention of location information, regardless of its task relevance. To explore the functional and neural bases of the representation of spatial context in VSTM for nonspatial information, functional magnetic resonance imaging was performed while subjects performed delayed recall for the orientation of individual stimuli. Stimulus location varied across trials, and although this information was irrelevant for task performance, multivariate pattern analysis decoding of stimulus location sustained across trials, and also the decoding strength, predicted the precision of the recall of orientation. The influence of spatial context on the representation of orientation was operationalized by comparing the orientation reconstructions with multivariate inverted encoding models (IEM) trained in location context-dependent vs. -independent data. Although orientation reconstructions were robust for both location-dependent and location-independent IEMs, they were markedly stronger for the former. Furthermore, the functional relevance of location context was demonstrated by the fact that only the location-dependent neural representations of stimulus orientation predicted recall precision.

NEW & NOTEWORTHY Neural representation strength of stimulus location predicts the precision of visual short-term memory (VSTM) recall of nonspatial stimulus, even when this information is task irrelevant. Neural representations of nonspatial stimuli that incorporate location context are stronger than those that do not, and only the former representations are strongly linked to behavior. The contributions to nonspatial VSTM performance of the representation of location context are at least partly distinct from those of the representation of stimulus content.

inverted encoding model; location context; visual short-term memory

INTRODUCTION

In the physical world, every object that we perceive and interact with occupies a location in space. Sometimes this spatial context is critical for stimulus individuation and/or for guiding our behavior, and sometimes it is irrelevant. However, the evidence has been mixed about whether, when spatial context is not needed to the task, this information is nonetheless encoded and retained in visual short-term memory (VSTM). Consistent with the idea that spatial context may be obligatorily encoded into VSTM is the finding from delayed recognition of a single item that response time (RT) is shorter when the location of the probe is congruent vs. incongruent with that of the sample (Olson and Marshuetz 2005). Similarly, subjects respond more quickly to a color probe when it appears at the same location as had the critical item from the sample array (Theeuwes et al. 2011). Other studies, however, suggest that the sensitivity of VSTM to spatial context may vary with testing conditions. One study demonstrated that spatial congruity effects were apparent only when the sample-to-probe interval was less than 1 s (Logie et al. 2011). A second study argued the spatial context only influenced VSTM performance on tasks in which it was implicitly emphasized by the experimental procedure, such as when a cue box surrounded one of the items in the test array, indicating the item that might have changed (Woodman et al. 2012). Finally, a third study reported spatial congruity effects were only seen for the most recently presented item in a serially presented memory array (Postle et al. 2013).

At the theoretical level, two recent computational models have posited that context at encoding may be obligatorily incorporated into mnemonic representations. One is a connectionist model in which stimuli are represented in each of two layers, a feature layer that represents object identity and a context layer that represents spatial and/or temporal context. It is the strength of the bindings between the two layers that determines how strongly the stimulus is held in VSTM (Oberauer and Lin 2017). A second is a neural field model in which different populations of neurons represent the conjunctions of one object feature and that feature's location. This model generalizes beyond VSTM by positing that the core object-representation function of feature binding is accomplished via these mappings of discrete features to a common set of location coordinates (Schneegans and Bays 2017).

Recent advances in the application of "information-based" multivariate data analysis methods have made it possible to

Address for reprint requests and other correspondence: B. R. Postle, Dept. of Psychology, Univ. of Wisconsin–Madison, 1202 West Johnson St., Madison, WI 53706 (e-mail: postle@wisc.edu).

evaluate multiple dimensions of stimulus representation. For example, one study applying multivariate pattern analysis (MVPA) to functional magnetic resonance imaging (fMRI) data has provided evidence for the spontaneous neural representation of visually presented objects with visual, phonological, and semantic codes (Lewis-Peacock et al. 2015), consistent with theoretical models of "multiple encoding" in human memory (Wickens 1973). For the neural representation of object location, two sets of studies are particularly germane. First, Ester et al. (2009) have demonstrated that, with MVPA of fMRI data, the orientation of an attended object could be decoded from primary visual cortex in the hemisphere ipsilateral to the visual field in which it had been presented, thereby revealing the representation of object information independent of that object's spatial context. Second, Foster et al. (2017) used inverted encoding models (IEM) to reconstruct, from electroencephalography (EEG) data, persistent representation of stimulus location during a VSTM task in which location was task irrelevant. However, the functional role of this "irrelevant" location context information is still unclear.

In the present study, we investigated ways in which the VSTM representation of a nonspatial feature is influenced by the representation of its (task irrelevant) spatial context. fMRI data were acquired while subjects performed delayed recall of the orientation of a bar, in which the stimulus location was task irrelevant. The strength of the neural representation of location context was estimated with MVPA, and the location-dependent and location-independent neural representations of orientations were estimated separately by reconstructions of the sample orientation with IEMs.

MATERIALS AND METHODS

Subjects

Sixteen right-handed volunteers, 8 women, aged 18–25 yr [mean (SD) = 20.50 (1.78)], from the University of Wisconsin–Madison community participated in the study for remuneration ($20/h). All subjects provided written informed consent according to the procedures approved by the Health Sciences Institutional Review Board at the University of Wisconsin–Madison. Subjects had normal or corrected-to-normal vision, no contraindications for MRI, and no reported history of neurological or psychiatric disease.

Behavioral Procedures

Task. Each trial of the delayed-recall task began with the presentation of a black bar (duration 4 s, length 4°, width 0.08°; rendered as the diameter of a white circular patch), equiprobably and unpredictably at one of four possible locations, each in one quadrant of the screen, and each with horizontal and vertical eccentricities from fixation of ±5°. Also equiprobable and unpredictable, and varying independently of location, was the orientation of the sample bar, which could appear in one of nine orientations (ranging from 0 to 160° in 20° increments, with jitter of ±1–5°, determined randomly on each trial). The 8-s delay period began with a mask (200 ms:18 black 0.08° × 4° bars, intersecting at their midpoints and each differing in orientation from its neighbors by 10°) presented at the same location as the sample. Recall was prompted with the onset of a display comprising a circular patch (reappearing in the same location as the sample) and a response wheel (also centered on fixation, with a radius to the outer edge of 9.2° and a width of 2°) that contained 20 equally spaced black bars (0.05° × 1.8°), ranging in orientation from 0 to 171° in 9° increments. The location on the wheel of the 0° bar varied unpredictably from trial to trial, to prevent action planning during the delay. Subjects were instructed to treat the ring as a continuous orientation space, and to indicate recall of the sample orientation by selection of the appropriate point along the response wheel with an MR-compatible trackball, within a 4-s response window. As soon as the trackball began to move, a bar appeared within the circular patch with an orientation, updating in real time, matching the orientation on the wheel that was highlighted by the cursor. RT was computed as the latency between response wheel onset and mouse click. Feedback indicating the error between the response and the sample orientation (in degrees) was presented immediately after the response until the end of the 4-s response window. The intertrial interval lasted for 8 s. Throughout the trial, a black fixation cross appeared in the center of the screen, and subjects were instructed to maintain fixation throughout each trial (Fig. 1).

Trials were blocked into 18-trial blocks (each orientation appearing twice per block; 7 min 12 s), and subjects performed 20 scanned blocks across two separate fMRI scanning sessions, each scanning session lasting ~1.5 h and the two were separated by 2–28 days. The data reported herein were collected as part of a larger experiment in which subjects also performed load-of-3 trials in one of two conditions: "homogeneous" trials, in which three orientation stimuli were presented simultaneously as samples, and "heterogeneous" trials, in which one orientation, one color, and one luminance stimulus were presented simultaneously as samples. On load-of-3 trials, a stimulus occupied three of the four possible sample-presentation locations. Recall was prompted by a recall display comparable to that illustrated in Fig. 1, with the identity of the stimulus to be recalled indicated both by the appearance of a circular patch in the location that had been occupied by the critical item and, for heterogeneous trials, also by the stimulus domain presented in the response wheel. During the first scanning session, subjects were first scanned while performing 4 blocks of 3-item trials and then 8 blocks of 1-item trials. In the second fMRI scanning session, they performed an additional 12 blocks of 1-item trials. Although the results from 3-item trials will be presented elsewhere, the relevance for this report is that because they preceded scanning of 1-item trials, they may have influenced strategy on 1-item trials.

All the experimental stimuli were controlled by the Psychophysics Toolbox (http://psychtoolbox.org; Brainard 1997) running in MATLAB (MathWorks), presented using a 60-Hz projector (Silent Vision 6011;

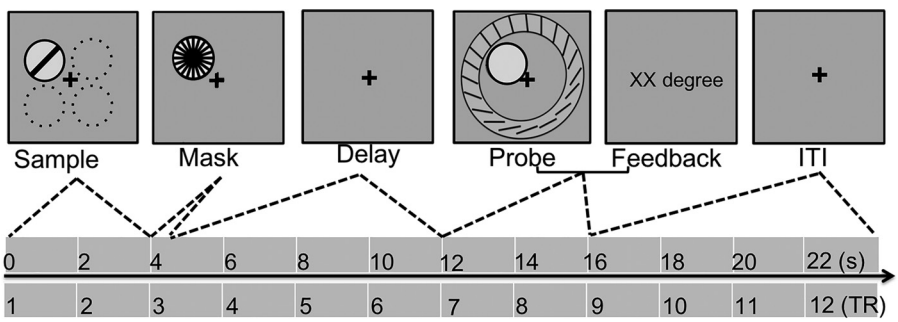

Fig. 1. Schematic illustration of the behavioral task. Dotted circles indicate the other possible stimulus presentation locations, but they were not presented during the experiment. TR, repetition time.

Avotec), and viewed through a coil-mounted mirror. The viewing distance was 68.58 cm and screen width was 33.02 cm.

Data analysis. Behavioral performance was assessed using a two-factor mixture model that estimated the proportion of responses made to the sample [i.e., the probability of a target response (P_T); the balance classified as "guess" responses] and the precision of these target responses (the variability of recall error) (Zhang and Luck 2008). Parameter estimates were obtained using maximum-likelihood estimation (expectation maximization) using MATLAB routines available at https://www.bayslab.com. Responses were entered for each trial, and separate estimates were obtained for each subject. Trials lacking a response, and those with raw error distances larger than 3 SD from the mean were excluded from the model-fitting process. On average, 4.62 trials (SD 2.07), out of the 360 trials performed, were excluded per subject.

Overview of hypothesis tests. To investigate the role of location context in VSTM, we carried out several analyses. First, to address whether stimulus location might influence nonspatial VSTM performance, we estimated the strength of the neural representation of sample location with MVPA and, for time points with successful decoding, regressed location MVPA performance against the precision of the behavioral response. This first analysis was also planned as a hypothesis-generating analysis, the results of which we would use to identify a priori time points at which to carry out the analyses that directly assessed location context binding. Second, we investigated the neural representation of orientations and the role of location context in these representations. IEM reconstructions were used for all these analyses to quantify the strength of neural representations of orientation in different brain areas and in different conditions. The first step was to establish that IEM reconstructions of stimulus orientation were related to the recall of stimulus orientation. Assuming success, the next step would be to compare the reconstruction of sample orientation by IEMs that were trained with stimuli presented in the same location on the screen ("location dependent") vs. IEMs that were trained with stimuli presented at different locations on the screen ("location independent"). The logic of this approach was that the mapping between high-dimensional patterns of blood oxygen level-dependent (BOLD) activity and the span of orientations varying from 0° to 179° that is learned by a location-dependent model may implicitly reflect the location context that was common to each of the trials in that model's training set. That is, the location-dependent models would be expected to be biased by stimulus location. A location-independent model, however, could not contain such information, because it would only be able to successfully reconstruct test orientations if, during training, it had learned to ignore any structure in the data that was location specific. To train location-independent IEMs, we randomly divided all trials into four groups and then trained IEMs on the data from just one of the four groups, to equate the number of trials used to train location-independent and location-dependent IEMs, and averaged the reconstruction results across the four sub-IEMs. Additionally, we addressed whether neural representations that contained congruent or no location context were differently predictive of behavioral performance. All of these analyses were carried out separately in occipital, parietal, and frontal regions of interest (ROIs).

fMRI Methods

Data acquisition. Whole brain images were acquired with a 3-T MRI scanner (Discovery MR750; GE Healthcare) at the Lane Neuroimaging Laboratory at the University of Wisconsin–Madison. For all subjects, a high-resolution T1-weighted image was acquired with a fast spoiled gradient-recalled echo sequence [repetition time (TR) = 8.2 ms, echo time (TE) = 3.2 ms, flip angle = 12°, 160 axial slices, 256 × 256 in-plane, 1.0 mm isotropic]. A T2*-weighted gradient echo pulse sequence was used to acquire data sensitive to the BOLD signal while subjects performed the VSTM task (TR = 2,000 ms, TE = 25 ms, flip angle = 60°, within a 64 × 64 matrix, 39 sagittal slices, 3.5 mm isotropic). Each of the 20 fMRI scans generated 213 volumes.

Preprocessing. fMRI data were preprocessed using the Analysis of Functional Neuroimages (AFNI) software package (https://afni.nimh.nih.gov; Cox 1996). All tasks runs were preceded by 4 s of dummy pulses to achieve a steady state of tissue magnetization. All volumes were spatially aligned to the first volume of the first run using a rigid-body realignment and were then aligned to the T1 volume. Volumes were corrected for slice-time acquisition, and linear, quadratic, and cubic trends were removed from each run to reduce the influence of scanner drift. For univariate analyses, data were spatially smoothed with a 4-mm full-width at half-maximum Gaussian and z-scored separately within run for each voxel. For MVPA and IEM analyses, data were z-scored separately within run for each voxel, but were not smoothed. All analyses were carried out in each subject's native space.

ROI generation. To operationalize the construct of location context dependence, we generated functionally defined ROIs from voxels that were specifically responsive to sample stimuli presented in each of the four locations and from voxels that were not. First, we solved a modified general linear model (GLM; implemented in AFNI) with regressors modeling the sample-presentation and delay epochs as 4- and 8-s boxcars, respectively, convolved with the canonical hemodynamic response function supplied with AFNI. "Location-responsive" voxels were then defined with the contrasts [$\text{Sample}_{\text{upper left}}$ − baseline], [$\text{Sample}_{\text{upper right}}$ − baseline], [$\text{Sample}_{\text{lower left}}$ − baseline], and [$\text{Sample}_{\text{lower right}}$ − baseline]. "Location-general" voxels (needed for spatial MVPA analyses) were identified with the contrast [($\text{Sample}_{\text{upper left}}$ + $\text{Sample}_{\text{upper right}}$ + $\text{Sample}_{\text{lower left}}$ + $\text{Sample}_{\text{lower right}}$) − baseline]. In parallel, the standard anatomical masks for occipital, parietal, and frontal cortex were obtained from the MNI152_T1_1mm template, transformed to each subject's individual structural image via affine transformations (Jenkinson and Smith 2001), and further refined via nonlinear interpolation (Andersson et al. 2007). Finally, anatomically constrained functional ROIs were generated for each subject by selecting the 400 voxels with the highest t values for the relevant contrast that were located within each of the anatomical masks (Fig. 2*A*). The number of voxels per ROI was determined after [Sample – baseline] contrasts from the GLMs, when thresholded at $z > 2.58$ and $P < 0.01$, and yielded averages of 459, 361, and 328 voxels per subject in occipital, parietal, and frontal cortex, respectively. The pattern of results reported did not change appreciably when the MVPA and IEM analyses were repeated with ROIs of 600 and 2,000 voxels. The number of voxels overlapping between location-responsive and location-general ROIs was 150 (79), 287 (50), and 231 (69) [mean (SD)] in occipital, parietal, and frontal cortex, respectively.

Pattern classification analysis for locations. To examine the neural representation of sample location within the location-general ROIs, which can also be construed as covert spatial attention, we trained MVPA classifiers to discriminate among the four locations and examined classifier sensitivity for each location using a leave-one-run-out cross-validation approach. Classification was performed within the location-general ROIs in occipital, parietal, and frontal cortex using L2-regularized logistic regression with a lambda penalty term of 25, implemented with the Princeton Multi-Voxel Pattern Analysis toolbox (www.pni.princeton.edu/mvpa/) and custom routines in MATLAB.

MVPA cross-validation was achieved by training a classifier on data from all runs but one, testing on the trials from the held-out run, and rotating through all possible permutations. For each TR of fMRI data, the classifier produced a probability estimate (from 0 to 1) of the extent to which the observed pattern on the tested trial matched the trained pattern for each of the trained locations. Significance of classifier performance was determined using one-tailed, one-sample t-tests, testing against chance performance of 0.25. To examine the dynamical nature of the neural representation of location, each clas-

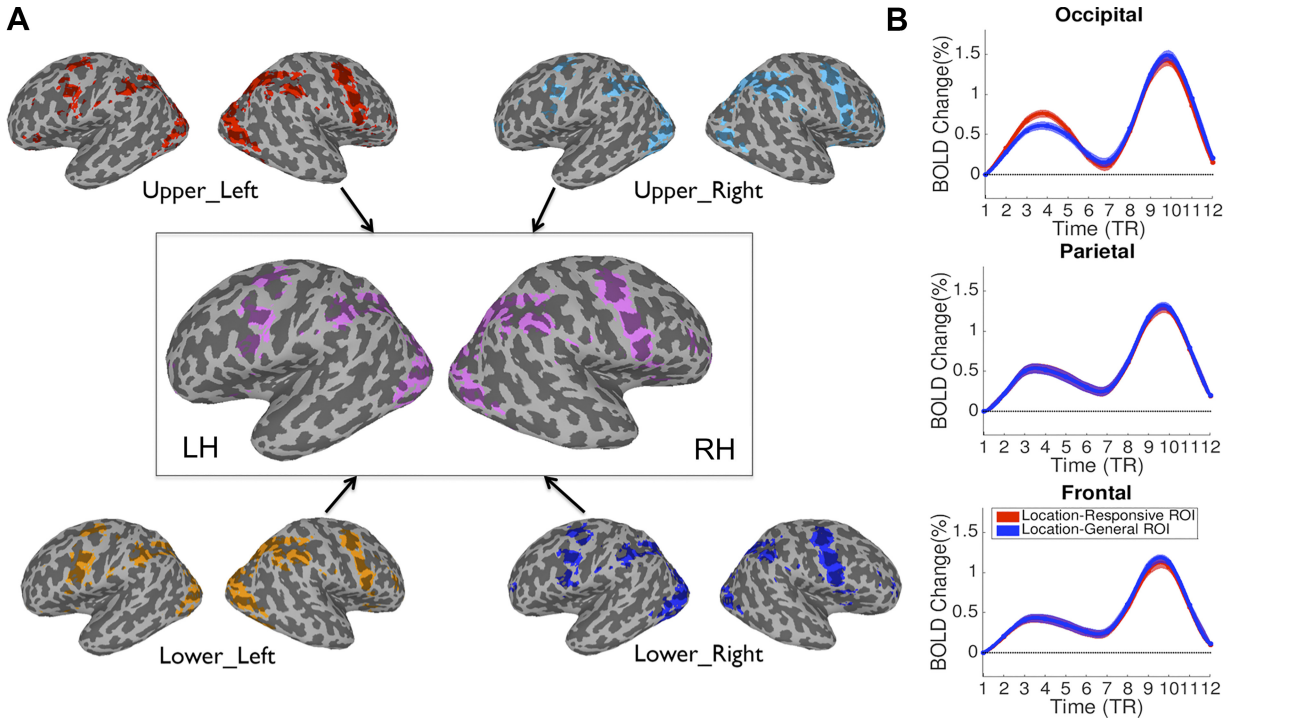

Fig. 2. *A*: location-responsive regions of interest (ROIs), displayed on smaller images and labeled by sample location. Location-general ROIs, displayed on larger images (*middle*), were generated by collapsing across these 4 trial types. LH, left hemisphere; RH, right hemisphere. *B*: blood oxygen level-dependent (BOLD) signal intensity changes in 3 ROIs across the trial. TR, repetition time.

sifier trained on data from one time point in the trial was tested on all time points across the trial, thereby generating a "temporal generalization matrix" (Cichy et al. 2014; Riggall and Postle 2012).

Finally, we explored whether variation in strength of this neural index of covert spatial attention related to (nonspatial) orientation recall performance. This was carried out, at each cell in the temporal generalization matrix at which location decoding accuracy was statistically reliable at the group level, by regressing location MVPA decoding accuracy against the behavioral precision of recall. These analyses were planned as a hypothesis-generating step, to define a priori the TRs at which we would perform direct tests of location context binding (i.e., for analyses carried with IEM reconstructions of orientation; described below).

Inverted encoding model analysis for orientations. Inverted encoding modeling (Serences and Saproo 2012) allowed us to reconstruct response profiles (termed channel tuning functions, or CTFs) that tracked the perceived and remembered orientations from multivoxel patterns of activity in each of the three ROIs. First, we extracted the normalized responses of each voxel in each ROI for each time point. Trials were sorted into one of nine bins based on stimulus orientation. We next divided the data into "training" and "test" sets, using an iterative leave-one-run-out approach, and modeled the measured responses of each voxel in the training set as a linear sum of nine orientation "channels," each with an idealized response function. Following the terminology of Brouwer and Heeger (2009), we estimated the weight matrix mapping each voxel's response to each orientation channel from the training data set and then inverted this matrix to estimate channel responses on each test trial. The average response output for each channel across trials was obtained by circularly shifting each response to a common center of 0°. To generate smooth, 180-point CTFs, we repeated the encoding model analysis 180 times and shifted the centers of the orientation channels by 1° on each iteration (Brouwer and Heeger 2009). The CTFs were averaged across permutations.

To quantify the neural reconstructions, we used linear regression to estimate CTF slope, with slope values >0 interpreted as evidence for an active neural representation (Foster 2017). Within each ROI, statistical significance of CTF slope was assessed with a bootstrapping method (Ester et al. 2015, 2016). For CTF slope estimates, within each ROI, we randomly selected (with replacement) and averaged the reconstructions across our 16 participants. This step was repeated 2,500 times, yielding 2,500 unique stimulus reconstructions. We then estimated the CTF slope of each reconstruction, and a *P* value was computed as the proportion of permutations for which the slope estimates less than (or equal to) 0 were obtained, and *P* values reported in RESULTS were false discovery rate (FDR)-corrected across TRs and ROIs. In a manner comparable to our procedure with the spatial MVPA analyses, IEMs were trained on data from each time point in the trial, and then reconstructions were attempted at the same time point ("diagonal reconstruction") as well as all the other time points of the trial ("off-diagonal reconstruction").

Location-dependent vs. -independent IEM reconstructions of orientations. To test whether location context modulated the neural representation of orientation, we compared the neural reconstructions of orientation with location-dependent vs. with location-independent IEMs. More specifically, location-dependent reconstructions were carried out within location-responsive ROIs and generated with data from trials with the same sample location as that with which the IEM was trained (90 trials in each location, results collapsed across the 4 sample locations). Location-independent reconstructions, in contrast, were carried out within location-general ROIs and generated with IEMs trained on trials drawn from all four locations. To match the number of trials used to train and test for each location-dependent IEM, this procedure was carried out four times, each time with an IEM trained on a different set of 90 trials (with an average of 22.5 trials drawn from each of the 4 stimulus locations), and reconstructions were carried out on data drawn from just a single stimulus location, with the results then collapsed across the reconstructions of data from the four locations. This operationalization of location congruency allowed us to adjudicate among three possible outcomes. First, orientation reconstructions might not differ between location-dependent and location-independent IEMs. This would be consistent

with the idea that VSTM representations of orientation are independent of location context. Second, stimulus reconstruction might only be successful for location-dependent IEMs. This would be consistent with the idea that location context is integral to the VSTM representations of orientation. The third possibility is the intermediate outcome, that location-independent reconstructions might be successful, but quantitatively shallower in slope (and therefore of lower strength and/or precision), than location-dependent reconstructions. This would suggest an interactive influence of location context on the working memory representation of orientation (see Fig. 4*A*). The location-dependency analyses were conducted on diagonal reconstructions.

We approached the statistical assessment of the influence of location context on the representation of stimulus orientation by computing both frequentist and Bayesian statistics (Dienes and McLatchie 2017; Ziori and Dienes 2015). *P* values were generated with a bootstrapping procedure whereby 2,500 permutations were applied to CTF slope estimates for location-dependent and location-independent reconstructions separately, and the results from each permutation were subtracted. For these comparisons, because values were invariably higher for location-dependent than location-independent reconstructions, one-tailed tests were conducted and *P* refers to the proportion of the 2,500 subtractions for which the parameter of the location-dependent reconstruction was greater than that of the location-independent reconstruction. All the *P* values reported in RESULTS were FDR-corrected across TRs and ROIs. Bayes factors (BF), which can be understood as the ratio of the likelihood of the alternative hypothesis compared with the null hypothesis, were computed to facilitate the TR-by-TR comparison of spatial congruency across IEM parameters, for each brain region. For these analyses, BFs with values >1 would indicate greater evidence in favor of the parameter in question differing between the two conditions, and values <1 would indicate greater evidence in favor of the two values not differing. BFs were computed using the calculator described by Dienes (2014), assuming a uniform distribution of the difference between location-dependent and location-independent reconstructions, and ranges for slope difference based on estimates from the model fitting (see RESULTS for details).

If any differences in slope estimates were observed across the location-dependent and location-independent IEMs, we further tested whether these were due to differences in the strength and/or the precision of the CTF. To estimate the reconstruction strength and precision, the CTF for each subject in each ROI was fit with an exponentiated cosine function of the following form: $f(x) = \alpha\{e^{\kappa[\cos(\mu - x) - 1]}\} + \beta$, where α and β control the vertical scaling (i.e., signal over baseline) and baseline of the function, respectively, and κ and μ control the concentration (the inverse of dispersion) and center of the function, respectively. No biases in reconstruction centers were expected or observed, so we fixed μ at 0. Fitting was performed by combining a GLM with a grid search procedure. We first defined a range of plausible κ values (from 1 to 30 in 0.1 increments). For each possible value of κ, we generated a response function using the fitting equation after setting α to 1 and β to 0. Next, we generated a design matrix containing the predicted response function and a constant term (i.e., a vector of 1 s) and used ordinary least-squares regression to obtain estimates of α and β (defined by the regression coefficients for the response function and constant term, respectively). We then selected the combination of κ, α, and β that minimized the sum of squared errors between the observed and predicted reconstructions. We approached the statistical assessment of the influence of location context on the reconstruction strength and precision by using a bootstrapping procedure similar to that used to test the slope difference. We also calculated BFs to facilitate interpretation of differences in the strength and precision of IEM reconstructions in different conditions.

Finally, the functional relevance of location context in the neural representation of orientation was assessed with across-subject regression of IEM reconstruction slope measure (for location-dependent and for location-independent reconstructions) against behavioral recall precision, at the time periods identified by the MVPA analyses of spatial selective attention. All the *P* values reported in RESULTS reflect FDR correction across regressions.

RESULTS

Behavioral Results

The RT was 2.825 (0.247) s, and the mean raw error distance for orientation recall was 7.76° (4.22°). The mixture model analysis revealed the mean P_T was 0.946 (0.00089) ($P_T > 0.99$ for 10 of 16 subjects), and the mean precision estimate was 13.749 (6.881) rad^{-1}.

fMRI Results

BOLD signal intensity. Within the occipital ROI, the trial-averaged time course from the top 400 location-general sample-responsive voxels showed a sample-related increase in BOLD signal that returned to baseline before the end of the delay period and then increased again during the response epoch. In the frontal and parietal ROIs, this task-related activity remained elevated across the delay period. The BOLD signal intensity patterns were quite similar in location-responsive and location-general ROIs, except in occipital cortex, where BOLD signal intensity in location-responsive ROIs was significantly larger than that in location-general ROIs (TRs 3 and 4, *P* values <0.001; Fig. 2*B*).

Neural representation of location. In occipital and parietal cortex, MVPA decoding of sample location was robust across the trial, and there was also considerable evidence for temporal generalization, suggesting temporally stable neural representations of stimulus location (Fig. 3, *A* and *B*). In the frontal ROI, MVPA decoding of location was much less reliable across the trial (Fig. 3*C*). Across-subject correlations of MVPA decoding of sample location against the precision of behavioral recall of orientations (i.e., nonspatial VSTM performance) revealed significant positive relations between these measures, in the occipital ROI, at TRs 8 and 9, corresponding to the late delay/early response period of the trial, along the diagonal of the matrix [TR 8: $r(15) = 0.589$, $P = 0.012$; TR 9: $r(15) = 0.531$, $P = 0.034$; averaged across TRs 8 and 9, $r = 0.583$, $P = 0.018$]. Additionally, at an off-diagonal location of the temporal generalization matrix, the accuracy of MVPA decoding at TR 9 with a classifier that had been trained on data from TR 3 was positively related to recall precision [$r(15) = 0.519$, $P = 0.039$; this can be understood as behavioral performance relating to the fidelity with which the neural representation of location at recall (i.e., TR 9) reinstated the neural representation of location during stimulus encoding (i.e., TR 3)]. No correlations were found between MVPA decoding of location in the parietal or frontal ROIs and behavioral performance (*P* values >0.15). (Note that because these correlations were to be used to identify time points at which to test hypotheses about location context, correction for multiple comparisons was not applied here.)

Neural representation of sample orientation. For location-dependent IEMs, reconstruction of the sample orientation was significant across the entire duration of the trial, in all three anatomical regions. Furthermore, training IEMs on data from one time point and testing on others revealed considerable

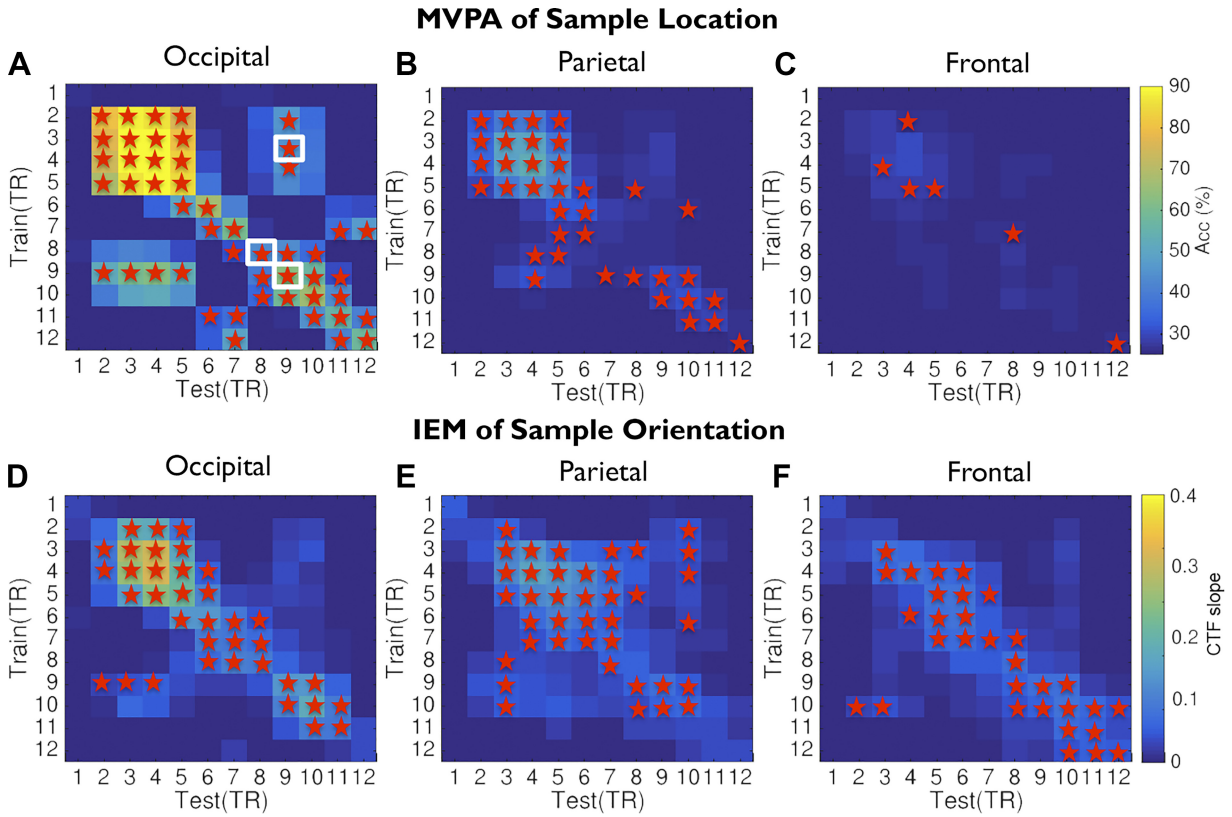

Fig. 3. Temporal generalization matrices of multivariate pattern analysis (MVPA) decoding of sample location (*A–C*) and of inverted encoding models (IEM) reconstruction of sample orientation (*D–F*). Red stars indicate significant decoding (*A–C*) or reconstruction (*D–F*) where the false discovery rate (FDR) corrected *P* values <0.05. White squares in *A* indicate cells at which successful location decoding also correlated with behavioral precision (without FDR correction to maximize the sensitivity). Color scales (*far right*) indicate accuracy (Acc) of MVPA decoding (*A–C*) or channel tuning function (CTF) slope (*D–F*). TR, repetition time.

temporal stability in the neural representation of orientation (Fig. 3, *D–F*).

Sensitivity of the Neural Representation of Orientation to Location Context

The analyses presented up to this point have considered the factors of stimulus location and stimulus identity in isolation. To directly examine the influence of location context on the neural representation of orientation, we carried out TR-by-TR comparisons of location-dependent IEM reconstructions of sample orientation vs. location-independent IEM reconstructions of sample orientation.

In occipital cortex, although both the location-dependent and location-independent reconstructions were significant across the trial, the slope of the reconstructions was reliably larger for location-dependent than for location-independent reconstructions during both the sample (at TRs 3 and 4) and probe (TR 10) epochs (Fig. 4*B*; Table 1). At each of these three TRs, the follow-up fits with an exponentiated cosine function suggested that this difference was due to the location-dependent reconstructions being greater in magnitude than the location-independent reconstructions (*P* values < 0.05, BFs > 3.582) and that there was no compelling evidence for differences in precision (BFs < 0.098).

In parietal and frontal cortex, although reconstructions across the trial appeared to be less robust for location-independent analyses, there was little compelling statistical evidence for a reliable influence of spatial context on the representation of orientation in these regions. Although the bootstrapping analyses suggested differences at one TR in parietal cortex and at two TRs in frontal cortex, the corresponding BFs suggested only very weak evidence in favor of the alternative hypothesis (Fig. 4, *C* and *D*; Table 1).

Brain-behavior relations. Correlations of IEM reconstruction with behavioral precision were carried out for the portion of the trial spanning TRs 8 and 9, as identified in the MVPA of stimulus location. This correlation was significant and positive for location-dependent IEM reconstructions [$r(15) = 0.567$, $P = 0.050$; Fig. 5*A*] but not significant for location-independent reconstructions [$r(15) = 0.139$, $P = 0.608$; Fig. 5*B*]. However, these effects did not differ statistically from each other ($z = 1.190$; not significant).

Although our analyses operationalized the neural representation of stimulus location and the influence of location context on the neural representation of stimulus orientation in very different ways, it is possible that these two measures tapped into the same underlying mechanism. To address this question, we first correlated spatial MVPA performance against the CTF slope of location-dependent IEM reconstruction of stimulus orientation (our operationalization of location context) and found no evidence of a link between these two measures [$r(15) = 0.245$; not significant]. Next, we entered these two neural measures, together with behavioral precision, into a stepwise regression model, which indicated that although the neural representations of stimulus location and of location context jointly explained 11.9% of the variance in behavioral

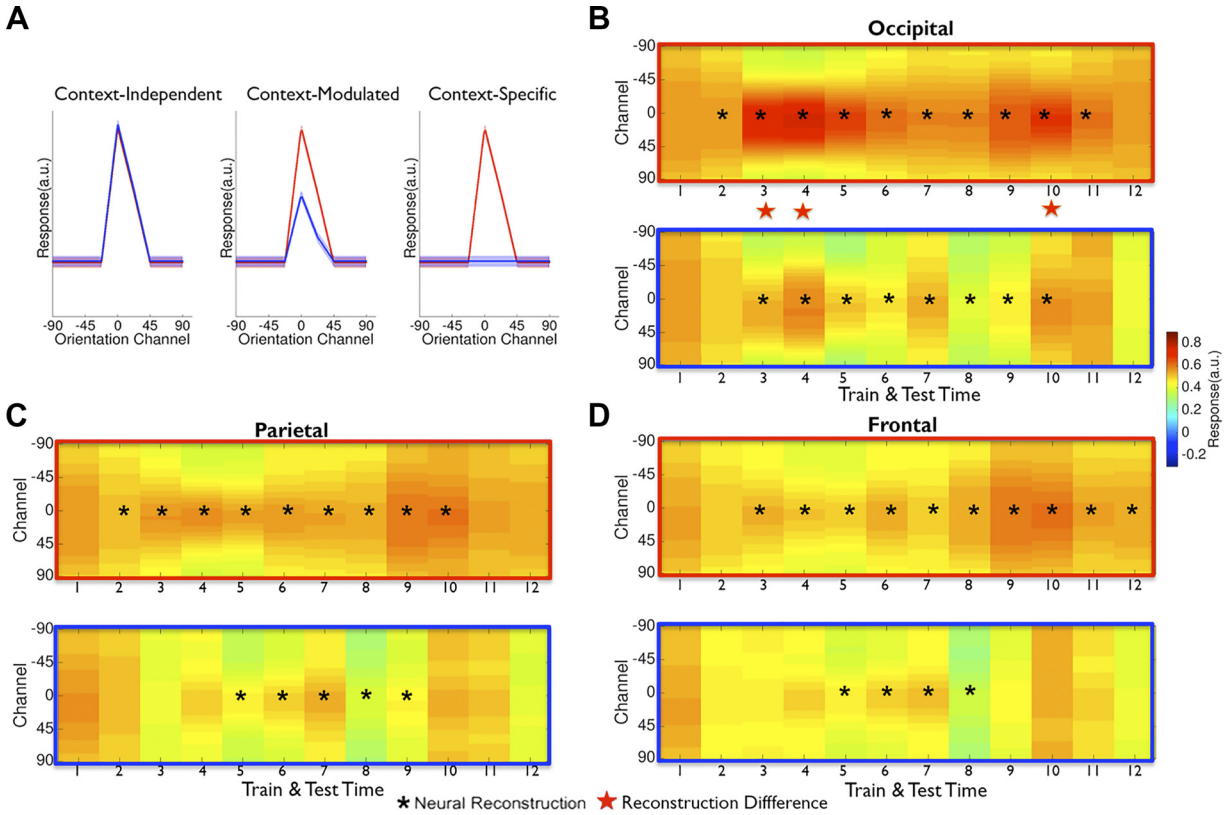

Fig. 4. *A*: 3 possible effects of location-context dependency. The neural reconstructions were quantified via a channel tuning function (CTF) slope estimate. *B–D*: inverted encoding model reconstruction time courses for location-dependent (*top*) and location-independent (*bottom*) inverted encoding models in each of 3 cortical regions, respectively. Black and red stars indicate significant comparisons where $P < 0.05$ after false discovery rate correction across repetition times and regions of interest.

performance, the representation of stimulus location uniquely explained an additional 22.0% of the behavioral variance, and the representation of location context uniquely explained an additional 15.5% of the behavioral variance.

The preceding analyses were carried out in the TRs of the hypothesis-generating MVPA analyses. We were additionally prompted by carrying out post hoc analyses on TRs 3 and 4 of the sample epoch, and on TR 10 of the probe epoch, after discovering the effects of location context at these TRs (Table 1). Analyses at the sample-epoch TRs also address the intuition that the initial encoding of location context must necessarily be constructed from information only present during that portion of the trial. The analyses of the sample-epoch TRs indicated that the precision of behavioral performance is correlated with CTF slope of location-dependent IEM reconstructions [$r(15) = 0.543$, $P = 0.050$; Fig. 5*C*], but not with the CTF slope of

Table 1. *Bayes factors for CTF slope difference across IEMs*

	TR 1	TR 2	TR 3	TR 4	TR 5	TR 6	TR 7	TR 8	TR 9	TR 10	TR 11	TR 12
Location-dependent reconstruction within location-responsive ROIs vs. location-independent reconstruction within location-general ROIs												
Occipital	0.035	0.066	5.199	24.412	0.087	0.253	0.05	0.081	0.087	3.856	0.495	0.042
Parietal	0.039	0.046	0.753	0.396	0.084	0.026	0.031	0.028	0.049	0.286	0.046	0.103
Frontal	0.037	0.041	0.375	0.097	0.043	0.03	0.035	0.046	0.086	2.06	0.642	0.091
Location-dependent reconstruction within location-responsive ROIs vs. location-dependent reconstruction within location-general ROIs												
Occipital	0.013	0.019	0.248	0.122	0.015	0.074	0.017	0.018	0.014	0.03	0.012	0.008
Parietal	0.009	0.013	0.017	0.008	0.038	0.01	0.007	0.013	0.087	0.005	0.015	0.048
Frontal	0.011	0.034	0.047	0.013	0.041	0.007	0.016	0.011	0.016	0.031	0.01	0.006
Location-dependent reconstruction within location-general ROIs vs. location-independent reconstruction within location-general ROIs												
Occipital	0.039	0.053	0.384	3.239	0.116	1.374	0.075	0.139	0.1	3.921	0.346	0.038
Parietal	0.04	0.05	1.334	0.497	0.142	0.028	0.03	0.028	0.029	0.334	0.123	0.048
Frontal	0.042	0.049	0.838	0.124	0.053	0.032	0.055	0.04	0.055	0.318	1.818	0.186

Values are Bayes factors (BFs) for channel tuning function (CTF) slope differences across the inverted encoding models (IEMs). To calculate the BF values, change limitations for slope differences were set at −1 to 1 [slope changes ranged from −0.376 to 0.428 across all 3 sets of comparisons, repetition times (TRs), regions of interest (ROIs), and individuals; SD = 0.113].

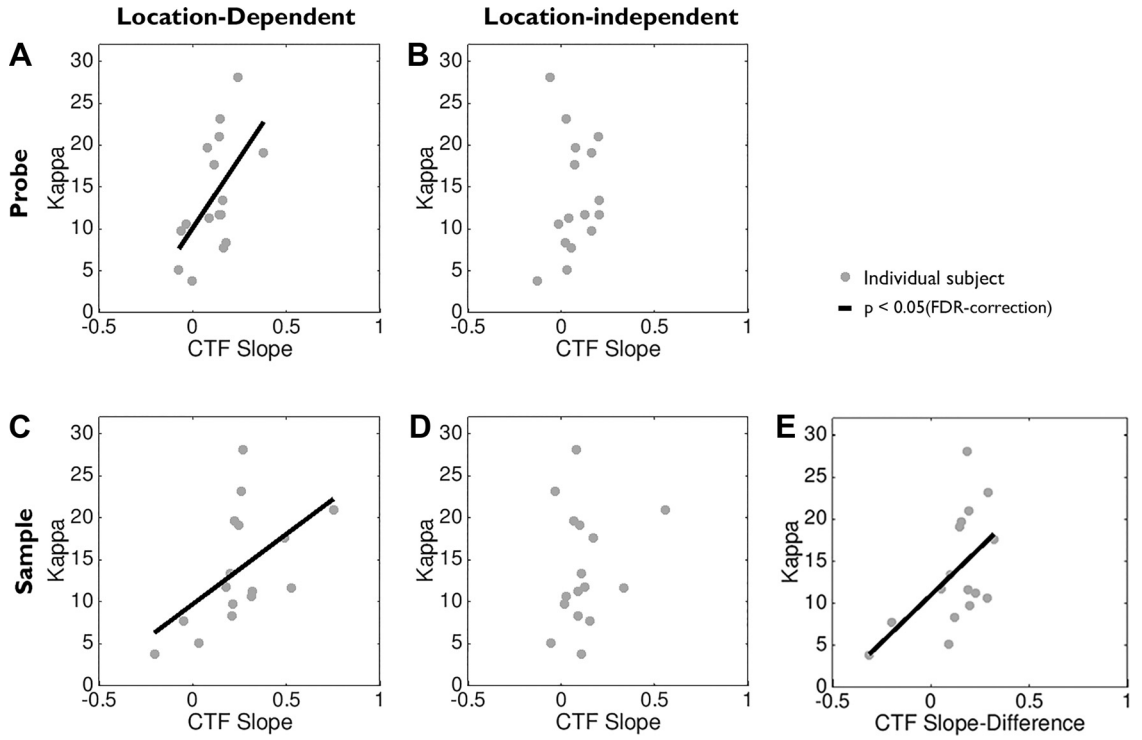

Fig. 5. Nonspatial visual short-term memory performance as a function of location dependency of inverted encoding model (IEM) reconstructions in occipital regions of interest. *A* and *B*: at probe epoch [averaged from repetition times (TRs) 8 and 9), the channel tuning function (CTF) slope of the location-dependent IEM reconstructions predicted the precision of behavioral recall (*A*), but the slope of the location-independent IEM reconstructions did not (*B*). *C–E*: at encoding epoch (averaged from TRs 3 and 4), the CTF slope of the location-dependent IEM reconstructions predicted the precision of behavioral recall (*C*), but the slope of the location-independent IEM reconstructions did not (*D*). *E*: the slope difference between location-dependent and -independent IEMs can also predict the precision of behavioral recall.

location-independent IEM reconstructions [$r(15) = 0.189$, $P = 0.604$; Fig. 5*D*]. Furthermore, although these correlations did not differ from each other ($z = 1.06$; $P = 0.145$), the difference between the CTFs slopes from context-dependent vs. context-independent reconstructions of stimulus orientation did correlate with the precision of behavioral performance ($r = 0.554$, $P = 0.050$; Fig. 5*E*). The same analyses, when carried out on TR 10, failed to show any significant correlations (P values >0.646).

Location dependency estimates not confounded by systematic differences in BOLD signal intensity. In occipital cortex, the BOLD signal intensity was significantly higher in location-responsive ROIs than in location-general ROIs during the encoding epoch (Fig. 2*B*). This raises a potential concern for interpretation of the location-dependency comparisons between IEM reconstructions for TRs 3 and 4 in occipital cortex, because differences in BOLD signal intensity have been shown, in some conditions, to explain differences in IEM reconstruction (Liu et al. 2018). We were able to rule out this possibility, however, by demonstrating that when location-dependent reconstructions were performed with the lower signal intensity BOLD data (i.e., in location-general ROIs), the slopes of these reconstructions were comparable to those of the reconstructions with the higher amplitude data (i.e., in location-responsive ROIs). That is, when IEMs were trained in a location-dependent manner (train IEMs with stimuli presented at just one location), the reconstructions obtained from location-responsive ROIs did not differ from the reconstructions obtained from location-general ROIs (P values >0.652; Fig. 6, *A–C*). This conclusion was reinforced when BFs were computed (Table 1).

Additionally, we replicated the main results of the location-dependency analyses by repeating the procedures for generating location-specific vs. location-independent reconstructions but carrying them out within location-general ROIs. In occipital cortex, CTF slopes were higher for location-dependent than for location-independent reconstructions during both the sample (TR 4) and probe (TR 10) epochs. These results further excluded the potential concern that the adaption of different sets of ROIs lead to the differences between location-dependent and location-independent IEMs. The sample period in parietal cortex (TR 3) and probe period in frontal cortex (TR 11) revealed reconstruction differences, but not as reliable as those in occipital cortex (P values <0.05; BF results in Table 1).

DISCUSSION

In this study of VSTM for line orientation, only one sample item was presented as a memorandum, in one of four possible locations, and the recall dial always appeared in the same location, making the encoding and retention of stimulus-location information unnecessary for successful performance. Nonetheless, our analyses suggest important roles for spatial processing on this task. First, the efficacy of maintaining stimulus location (as operationalized by MVPA performance), particularly during recall, was positively related to the precision of recall of stimulus orientation. Second, the neural representation of orientation was higher in amplitude, and

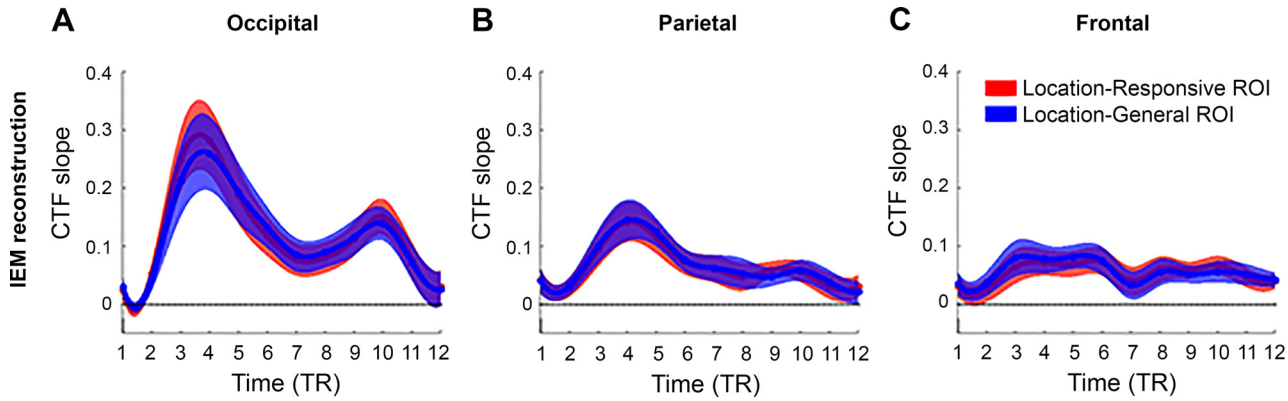

Fig. 6. *A–C*: selectivity of the inverted encoding model (IEM) reconstructions, estimated by channel tuning function (CTF) slope, in 3 location-responsive regions of interest (ROI; red) and 3 location-general ROIs (blue). Both IEMs were trained and tested with the trials sharing the same location context. TR, repetition time.

more closely related to the precision of recall of stimulus orientation, when it included information about location context. Finally, the mechanisms underlying neural representation of stimulus location and location-context binding are at least partly dissociable.

Strongest Effects in Occipital Cortex

Although IEM analyses indicated that neural representation of the remembered orientation was retained throughout the trial in all three regions that we examined (sample-responsive ROIs in occipital, parietal, and frontal cortex), evidence for location processing was strongest in occipital cortex. With regard to the representation of stimulus location, which can be construed as spatial covert attention (Awh et al. 1998), the effects were most robust in the occipital ROI, and only in the occipital ROI (and only during probe TRs) did individual differences in the MVPA decoding performance, a proxy for the efficacy of spatial covert attention, predict individual differences in the behavioral precision of VSTM for line orientation. Future research will be required to determine the extent to which probe-epoch MVPA may be indexing the efficacy of spatial attention, per se, or some more general factor such as effort or arousal.

The influence of location context on the representation of orientation was also strongest in occipital cortex, where the IEM reconstructions from sample- and probe-epoch data were stronger with location-dependent data. It was also at these TRs that the slope estimates of the orientation reconstructions correlated with the precision of VSTM. This result is compatible with computational models of VSTM, suggesting that the binding of context to stimulus features (Oberauer and Lin 2017; Schneegans and Bays 2017) may occur in the overlapping occipital circuits that represent line orientation and that represent retinotopic space. In contrast, the lack of evidence for an influence of location context on stimulus representations in parietal and frontal cortex suggests that these may be more abstract. Future work will need to explore this possibility, together with the possibility that the strength of parietal and frontal representations of the stimulus per se may be less important for guiding behavior. Additionally, a future study in which the probe does not always appear in the same location as the sample will also be needed to assess the extent to which the present results may have been influenced by strategies afforded by this within-trial congruity.

Nature of the Neural Implementation of Location Context

In occipital and parietal cortex, MVPA decoding indicated that a representation of the location of the sample was actively maintained throughout the trial, including across the delay period. Neural evidence for an influence of location context on the representation of stimulus orientation diverged from this pattern along two dimensions: it was only observed in occipital cortex, and it was not observed during the delay period. This indicates that the neural implementation of location context consists of more than "just" the modulation of an object representation with spatial attention, or some other spatial signal. Further evidence for a difference between these two constructs comes from the individual-differences analyses. Spatial MVPA performance was not correlated with location-dependent IEM reconstruction, and these two measures each accounted for unique variance when being related to VSTM performance.

One noteworthy aspect of our results was that the influence of location context, as operationalized in our analyses, was only expressed when a stimulus was on the screen. This is consistent with results from a previous study in which a neural correlate of the spatial congruity effect was only observed in the probe-evoked response. This fMRI finding, together with behavioral results from a series of related experiments, led the authors to propose that, for working memory for nonspatial information, "position-related information is not actively stored in VSTM, but may be retained in a passive tag that marks the most recent site of selection" (Postle et al. 2013). Along these lines, it will be interesting to assess, in future work, whether, and if so, how, location context is actively represented during the retention period, when no stimulus-related information is visible. Furthermore, although we have noted the conceptual overlap between the question studied here and computational models (Oberauer and Lin 2017; Schneegans and Bays 2017), our understanding of these models is that they would predict sustained representation of location context across the delay period. Might the delay-period representation of context, like the working-memory representation of information that is temporarily outside the focus of attention (LaRocque et al. 2017; Rose et al. 2016), be maintained in a format that is different from the information that is of principal relevance for the impending memory probe?

GRANTS

This work was supported by the National Science Foundation of China (31730038), the National Basic Research Program of China (973 Program No. 2014CB846102), and National Institutes of Health Grant R01MH064498 (to B. R. Postle).

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

Y.C., A.D.S., and B.R.P. conceived and designed research; Y.C. performed experiments; Y.C. and A.D.S. analyzed data; Y.C., A.D.S., Q.Y., and B.R.P. interpreted results of experiments; Y.C. prepared figures; Y.C. drafted manuscript; Y.C., Q.Y., and B.R.P. edited and revised manuscript; B.R.P. approved final version of manuscript.

REFERENCES

Andersson JL, Jenkinson M, Smith S. *Non-Linear Registration, aka Spatial Normalisation FMRIB Technical Report TR07JA2*. Oxford, UK: FMRIB Analysis Group of the University of Oxford, 2007.

Awh E, Jonides J, Reuter-Lorenz PA. Rehearsal in spatial working memory. *J Exp Psychol Hum Percept Perform* 24: 780–790, 1998. doi:10.1037/0096-1523.24.3.780.

Brainard DH. The Psychophysics Toolbox. *Spat Vis* 10: 433–436, 1997.

Brouwer GJ, Heeger DJ. Decoding and reconstructing color from responses in human visual cortex. *J Neurosci* 29: 13992–14003, 2009. doi:10.1523/JNEUROSCI.3577-09.2009.

Cichy RM, Pantazis D, Oliva A. Resolving human object recognition in space and time. *Nat Neurosci* 17: 455–462, 2014. doi:10.1038/nn.3635.

Cox RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res* 29: 162–173, 1996.

Dienes Z. Using Bayes to get the most out of non-significant results. *Front Psychol* 5: 781, 2014. doi:10.3389/fpsyg.2014.00781.

Dienes Z, McLatchie N. Four reasons to prefer Bayesian analyses over significance testing. *Psychon Bull Rev* 25: 207–218, 2017. doi:10.3758/s13423-017-1266-z.

Ester EF, Serences JT, Awh E. Spatially global representations in human primary visual cortex during working memory maintenance. *J Neurosci* 29: 15258–15265, 2009. doi:10.1523/JNEUROSCI.4388-09.2009.

Ester EF, Sprague TC, Serences JT. Parietal and frontal cortex encode stimulus-specific mnemonic representations during visual working memory. *Neuron* 87: 893–905, 2015. doi:10.1016/j.neuron.2015.07.013.

Ester EF, Sutterer DW, Serences JT, Awh E. Feature-selective attentional modulations in human frontoparietal cortex. *J Neurosci* 36: 8188–8199, 2016. doi:10.1523/JNEUROSCI.3935-15.2016.

Foster JJ, Bsales EM, Jaffe RJ, Awh E. Alpha-band activity reveals spontaneous representations of spatial position in visual working memory. *Curr Biol* 27: 3216–3223.e6, 2017. doi:10.1016/j.cub.2017.09.031.

Jenkinson M, Smith S. A global optimisation method for robust affine registration of brain images. *Med Image Anal* 5: 143–156, 2001. doi:10.1016/S1361-8415(01)00036-6.

LaRocque JJ, Riggall AC, Emrich SM, Postle BR. Within-category decoding of information in different attentional states in short-term memory. *Cereb Cortex* 27: 4881–4890, 2017. doi:10.1093/cercor/bhw283.

Lewis-Peacock JA, Drysdale AT, Postle BR. Neural evidence for the flexible control of mental representations. *Cereb Cortex* 25: 3303–3313, 2015. doi:10.1093/cercor/bhu130.

Liu T, Cable D, Gardner JL. Inverted encoding models of human population response conflate noise and neural tuning width. *J Neurosci* 38: 398–408, 2018. doi:10.1523/JNEUROSCI.2453-17.2017.

Logie RH, Brockmole JR, Jaswal S. Feature binding in visual short-term memory is unaffected by task-irrelevant changes of location, shape, and color. *Mem Cognit* 39: 24–36, 2011. doi:10.3758/s13421-010-0001-z.

Oberauer K, Lin H-Y. An interference model of visual working memory. *Psychol Rev* 124: 21–59, 2017. doi:10.1037/rev0000044.

Olson IR, Marshuetz C. Remembering "what" brings along "where" in visual working memory. *Percept Psychophys* 67: 185–194, 2005. doi:10.3758/BF03206483.

Postle BR, Awh E, Serences JT, Sutterer DW, D'Esposito M. The positional-specificity effect reveals a passive-trace contribution to visual short-term memory. *PLoS One* 8: e83483, 2013. doi:10.1371/journal.pone.0083483.

Riggall AC, Postle BR. The relation between working memory storage and elevated activity, as measured with functional magnetic resonance imaging. *J Neurosci* 32: 12990–12998, 2012. doi:10.1523/JNEUROSCI.1892-12.2012.

Rose NS, LaRocque JJ, Riggall AC, Gosseries O, Starrett MJ, Meyering EE, Postle BR. Reactivation of latent working memories with transcranial magnetic stimulation. *Science* 354: 1136–1139, 2016. doi:10.1126/science.aah7011.

Schneegans S, Bays PM. Neural architecture for feature binding in visual working memory. *J Neurosci* 37: 3913–3925, 2017. doi:10.1523/JNEUROSCI.3493-16.2017.

Serences JT, Saproo S. Computational advances towards linking BOLD and behavior. *Neuropsychologia* 50: 435–446, 2012. doi:10.1016/j.neuropsychologia.2011.07.013.

Theeuwes J, Kramer AF, Irwin DE. Attention on our mind: the role of spatial attention in visual working memory. *Acta Psychol (Amst)* 137: 248–251, 2011. doi:10.1016/j.actpsy.2010.06.011.

Wickens DD. Some characteristics of word encoding. *Mem Cognit* 1: 485–490, 1973. doi:10.3758/BF03208913.

Woodman GF, Vogel EK, Luck SJ. Flexibility in visual working memory: accurate change detection in the face of irrelevant variations in position. *Vis Cogn* 20: 1–28, 2012. doi:10.1080/13506285.2011.630694.

Zhang W, Luck SJ. Discrete fixed-resolution representations in visual working memory. *Nature* 453: 233–235, 2008. doi:10.1038/nature06860.

Ziori E, Dienes Z. Facial beauty affects implicit and explicit learning of men and women differently. *Front Psychol* 6: 1124, 2015. doi:10.3389/fpsyg.2015.01124.