

Decoding the content of visual short-term memory under distraction in occipital and parietal areas

Katherine C Bettencourt & Yaoda Xu

Recent studies have provided conflicting accounts regarding where in the human brain visual short-term memory (VSTM) content is stored, with strong univariate fMRI responses being reported in superior intraparietal sulcus (IPS), but robust multivariate decoding being reported in occipital cortex. Given the continuous influx of information in everyday vision, VSTM storage under distraction is often required. We found that neither distractor presence nor predictability during the memory delay affected behavioral performance. Similarly, superior IPS exhibited consistent decoding of VSTM content across all distractor manipulations and had multivariate responses that closely tracked behavioral VSTM performance. However, occipital decoding of VSTM content was substantially modulated by distractor presence and predictability. Furthermore, we found no effect of target–distractor similarity on VSTM behavioral performance, further challenging the role of sensory regions in VSTM storage. Overall, consistent with previous univariate findings, our results indicate that superior IPS, but not occipital cortex, has a central role in VSTM storage.

VSTM is a short-term memory buffer that has a vital role in temporarily maintaining visual information that is critical to guiding our thoughts and actions. It is an important gateway to information integration and high-level cognition. Research in non-human primates has consistently shown evidence for VSTM maintenance in parietal and prefrontal cortices¹. Similarly, in humans, strong univariate functional magnetic resonance imaging (fMRI) responses during the memory delay period in parietal cortex have highlighted the importance of this region in VSTM information storage. A region expanding across the superior IPS (referred to as superior IPS for simplicity), in particular, has been shown to track the amount of task-relevant information stored in VSTM^{2–7}. Consistent with fMRI findings, transcranial magnetic stimulation (TMS) to parietal regions has also been shown to affect VSTM processing and maintenance^{8,9}.

In more recent studies using fMRI multivariate pattern analysis (MVPA), however, human occipital cortex has been shown to exhibit strong and consistent decoding of VSTM contents^{10–19}. Despite the presence of strong univariate VSTM responses in human parietal cortex and its ability to represent task-relevant visual features^{20–22}, MVPA studies have produced mixed decoding results regarding the role of this brain region in VSTM information representation^{11,15,16,23,24}. Together, these results have been used to argue that occipital cortex, rather than parietal cortex, has a central role in the storage of VSTM in the human brain.

Although findings from occipital cortex are robust, they are also puzzling. First, given the almost unlimited representational capacity of the primary visual cortex in sensory processing, it is unclear how this brain region would give rise to a highly capacity-limited VSTM system. Second, why would a brain region that is primarily involved in perception be recruited for VSTM storage? Given the continuous influx of visual information in everyday visual perception, it is often necessary to hold information in VSTM while

concurrently processing incoming visual stimuli. How can VSTM representations be maintained in the face of such distraction? Previous psychophysical work has shown that distractors that are similar to targets can interfere with VSTM performance¹. Although this has been taken as evidence supporting the sensory nature of VSTM representation, it also highlights the need to separate memory and incoming sensory representations to reduce interference. Furthermore, as both distractor and VSTM processing engage other brain regions such as parietal and prefrontal cortices, distractor interference could occur in any of these regions. Thus, the behavioral interference results alone do not pinpoint occipital cortex as the primary VSTM storage site.

Although previous MVPA studies have produced mixed results regarding the role of the parietal cortex in VSTM information representation^{11,15,16,23,24}, none of them specifically targeted superior IPS, a key parietal region whose activity tracks VSTM storage^{2–6}. Thus, the role of the human parietal cortex in VSTM representation has not been adequately evaluated with MVPA.

In non-human primates, conflicting results have implicated both parietal and prefrontal regions in the representation of VSTM information under distraction^{25–27}. However, to our knowledge, in humans, no brain region has been shown to represent VSTM information during the delay regardless of distraction, and it therefore remains unclear whether or how occipital and parietal cortices would contribute to real-world VSTM processing, where distraction is constant.

Thus, despite substantial research on the neural basis of VSTM, the fundamental question of where in the brain the content of VSTM is stored has not been answered. We found that MVPA decoding in superior IPS, but not occipital cortex, closely tracked behavioral measures of information storage in VSTM across distractor presence and predictability. This suggests that superior IPS, but not occipital cortex, has a central role in VSTM storage in the human brain.

Department of Psychology, Harvard University, Cambridge, Massachusetts, USA. Correspondence should be addressed to Y.X. (yaodaxu@fas.harvard.edu).

Received 17 July; accepted 26 October; published online 23 November 2015; doi:10.1038/nn.4174

RESULTS

Decoding VSTM content with predictable distractors

To assess the role of both occipital and parietal cortices in VSTM storage under visual distraction (experiment 1), we adapted an oriented grating VSTM task used previously¹³, which was shown to elicit robust VSTM decoding in occipital cortex, and manipulated whether or not distractors were present during the delay period (Fig. 1). Ten participants were shown two gratings (~25° or ~115°) sequentially at fixation and then retroactively cued as to which orientation to remember. After an extended delay (11 s), a third grating appeared at fixation and participants reported whether this grating was jittered clockwise or counterclockwise from the remembered grating. During the delay, either a blank screen (trials without distractors) or a series of face or gazebo stimuli (trials with distractors) were presented. In an effort to replicate the previous study's findings¹³ and to minimize any changes in VSTM strategy brought on by the distractors, we had all participants complete all eight blocks of trials without distractors before switching to trials with distractors. Participants were thus able to anticipate, with 100% accuracy, whether a given block of trials would contain distractors.

Behaviorally, we obtained very similar performance accuracy to seen previously¹³, with an average of 76.3% correct across all trials. Notably, there were no differences in performance ($t(9) = 0.8$, $P = 0.43$) between trials with distractors (77.3%) and trials without (75.3%). Moreover, there was no difference between the first and second half of each trial type (trials without distractors: $t(9) = 1.3$, $P = 0.23$; trials with distractors: $t(9) = 0.4$, $P = 0.68$), and no difference between the two trial types when we only examined the first half of trials in each ($t(9) = 0.9$, $P = 0.4$). This suggests that trials with distractors were never more difficult than trials without distractors and that VSTM storage is resistant to the kind of visual distraction introduced here.

MVPA decoding accuracy for the remembered stimulus during the delay period was then examined in our occipital and parietal regions of interest (ROIs; Fig. 2 and Online Methods) after responses were z-scored in a given ROI to remove any response amplitude differences among the different brain regions. In occipital cortex, when decoding performance was examined in areas V1 through V4 individually, we found no significant interaction between ROI and trial type ($F(3,9) = 0.8$, $P = 0.48$). As such, following what was done previously¹³, we combined these regions into a single ROI, V1–V4. Replicating the previous study¹³, decoding accuracy for the average delay period in V1–V4 (Fig. 3a) in trials without distractors was significantly above chance

($t(9) = 7.1$, $P < 0.0001$). However, for trials with distractors, decoding accuracy dropped significantly compared with trials without distractors ($t(9) = 5.6$, $P = 0.0004$) and no longer differed from chance performance ($t(9) = 0.8$, $P = 0.44$), even though there was no significant behavioral difference between the two trial types. Although chance level decoding does not necessarily imply the absence of VSTM representation, as limitations of fMRI MVPA could have prevented the readout of weak VSTM representations, the significant drop in the decoding performance, however, unambiguously shows that distractor presence significantly modulated the strength of VSTM representation in occipital cortex.

In superior IPS, decoding accuracy across the delay period was above chance for both trials without distractors ($t(9) = 3.0$, $P = 0.02$) and those with distractors ($t(9) = 4.6$, $P = 0.001$), with no difference between these two trial types ($t(9) = 1.7$, $P = 0.13$) (Fig. 3b). Although the overall decoding accuracy was lower in this region than in V1–V4 in trials without distractors, this was likely a result of differences in ROI size and signal-to-noise ratios that are unrelated to the actual strength of the memory representations. Notably, the interaction between trial type and ROI was significant ($F(1,9) = 9.5$, $P = 0.01$), indicating that the effect of distractors on VSTM decoding differed between occipital and parietal regions, with distractors affecting VSTM representation in occipital cortex, but not superior IPS.

Although the face and the gazebo distractors were task irrelevant, they could nevertheless be decoded with high accuracy in both V1–V4 (accuracy = 0.98, $t(9) = 39.0$, $P < 0.0001$) and superior IPS (accuracy = 0.92, $t(9) = 18.3$, $P < 0.0001$). This suggests that parietal cortex is capable of maintaining the memory item while concurrently processing incoming visual stimuli. Occipital cortex, on the other hand, was significantly affected by the presence of additional visual stimuli, and appeared to favor incoming visual stimuli over memory representations.

Decoding overlapping visual stimuli in occipital cortex

It is possible that distractor processing obscured our ability to decode the memory representation in occipital cortex resulting from limitations of fMRI MVPA. Although decoding in superior IPS argues against the idea of such limitations, given the greater distractor-induced response amplitude change in occipital cortex than in superior IPS (Supplementary Fig. 1), it is important to directly assess whether overlapping visual stimuli can be successfully decoded in occipital cortex.

Figure 1 Main experimental task from experiments 1 and 3. Participants were shown two orientated gratings and then cued as to which to remember. The cue presented here is enlarged for clarity. After a long delay, a third grating appeared and participants were asked to judge whether this grating was jittered clockwise or counterclockwise to the remembered grating. During the delay, participants either saw a blank screen with a fixation dot (trials without distractors) or a sequential presentation of task irrelevant faces or gazebos (trials with distractors). In experiment 1, trials without distractors were presented in the first half of the experiment while those with distractors were presented in the second half, making distractor presence/absence predictable. In experiment 3, the two types of trials were randomly intermixed within a run, making distractor presence/absence unpredictable.

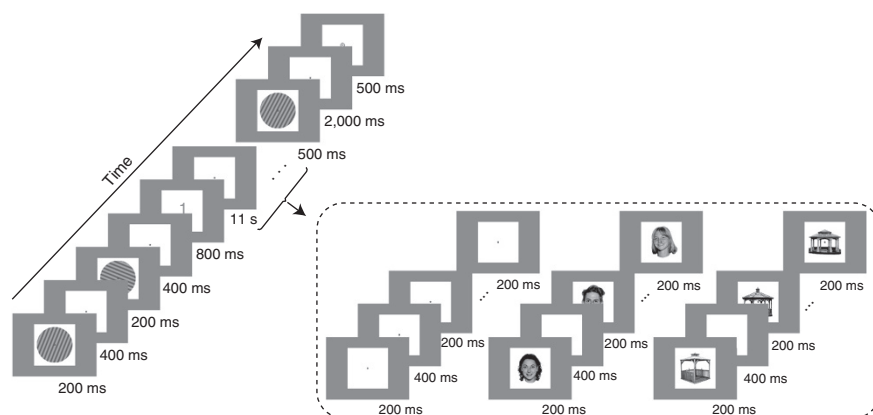
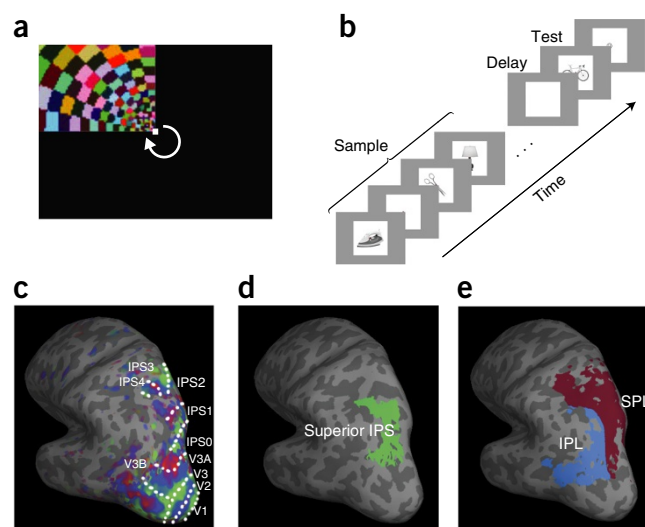


Figure 2 ROIs and the localizer tasks. (a–d) A moving, flashing, colored checkerboard wedge (a) and an object-based VSTM task (b) were used to define occipital and parietal topographic regions (c) and superior IPS (d), respectively. In the VSTM task, participants were shown a sequential presentation of either 1, 2, 3, 4 or 6 real-world objects at fixation and, after a brief delay, reported whether the test object shown at fixation matched or did not match one of the remembered objects. Superior IPS was defined as a region that tracked the behavioral VSTM capacity measures in this task. (e) IPL and SPL were anatomically defined. Each ROI was further refined to select voxels that respond to the task stimuli. All ROIs are shown here on the inflated left hemisphere of an example participant.



Previous fMRI MVPA studies have shown that VSTM representations in occipital cortex are highly similar in pattern to those produced by perceptual stimulation^{13,16,28}. We were able to replicate this finding in our VSTM experiment in the trials without distractors. Specifically, when we trained a classifier using the probe stimuli at test as our perceptual stimulus, we still found significant cross-decoding during the VSTM delay (accuracy = 0.62, $t(9) = 3.1$, $P = 0.01$). The sensory nature of VSTM representations in occipital cortex therefore allowed us to remove VSTM-related processing and directly test whether or not orientated gratings presented perceptually could be decoded in occipital cortex with and without overlapping distractors.

In experiment 2, eight of the ten participants who took part in experiment 1 were shown the same grating stimuli ($\sim 25^\circ$ or $\sim 115^\circ$) as in experiment 1, but at a much lower contrast (25% opacity) to simulate the reduced strength of the VSTM representations (Fig. 4). As shown in both the univariate (Supplementary Fig. 2) and multivariate (Fig. 5) results below, this level of contrast produced very comparable, if not weaker, representations in occipital cortex than the same memory representation in experiment 1 (Fig. 3a and Supplementary Fig. 1). During the experiment, the grating was presented either alone (trials without distractors) or overlapped by the same face or gazebo distractor stimuli from experiment 1 (trials with distractors). The timing of the grating and distractor presentations mirrored that of the delay period in experiment 1 (Online Methods and Fig. 4). To make decoding more challenging, instead of asking participants to attend the gratings, as they would do during the delay period of the VSTM task, we asked them to perform a one-back letter repetition detection task on a letter stream presented at fixation. This fixation task mirrors a perceptual task used previously¹³, and, as it does not require participants to attend or encoding the grating stimulus in any way, it removes all VSTM processing related to the grating stimuli that may be automatically engaged when participants attend to a stimulus.

Using the same ROI defined in experiment 1, we found that orientation decoding accuracy in V1–V4 was significantly above chance for the presented orientation (Fig. 5) whether or not distractors were present (trials without distractors: $t(7) = 3.5$, $P = 0.01$; trials with distractors: $t(7) = 4.2$, $P = 0.004$), with no difference between the two trial types ($t(7) = 0.4$, $P = 0.72$). Decoding accuracy for the distractor category was also above chance (accuracy = 0.98, $t(7) = 58.1$, $P < 0.0001$).

In this task, in trials with distractors, the grating was perceptually present throughout the entire delay, whereas distractors flickered on and off, creating some portion of the time in which only the grating was present during the delay. We modeled this task off the assumption that, in experiment 1, the VSTM representations during the delay were also constant, whereas the appearance of distractors was transient. Thus, any boost to decoding in this experiment created by the grating being presented ‘alone’ during a portion of the delay in trials with distractors would also exist in experiment 1. Moreover, as a result of adaptation, the prolonged presentation of the gratings during the perceptual task should have actually weakened the orientation representation compared to if we had flickered the grating with the distractors, as that type of stimulation would have more optimally driven responses in occipital cortex. Although this experiment might not provide a definitive answer to whether or not completely overlapping stimuli could both be successfully decoded in occipital cortex, it nevertheless created a comparable decoding situation to VSTM to help us better understand the nature of the decoding drop in experiment 1 when distractors were present.

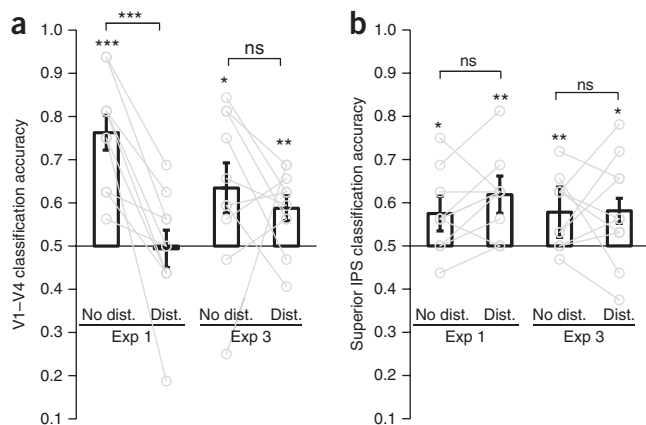


Figure 3 MVPA decoding accuracy for experiments 1 and 3. (a,b) MVPA decoding accuracy for the average VSTM delay period activity in V1–V4 (a) and superior IPS (b) in experiments 1 (with predictable distractors) and 3 (with unpredictable distractors). The same ten participants took part in both experiments. Although the presence and predictability of distractors did not affect behavioral performance, V1–V4 showed successful VSTM decoding when distractors were absent and the presence of distractors was predictable (experiment 1), but showed a significant drop to chance-level decoding when distractors were present. However, when the presence of distractors was unpredictable (experiment 3), V1–V4 showed weaker, but significant and comparable, VSTM decoding for both distractor present and absent conditions. Unlike V1–V4, superior IPS mirrored behavioral performance and showed consistent and significant VSTM decoding irrespective of distractor presence and predictability. Error bars indicate s.e.m. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$; ns, non-significant; No dist., trials without distractors; Dist., trials with distractors.

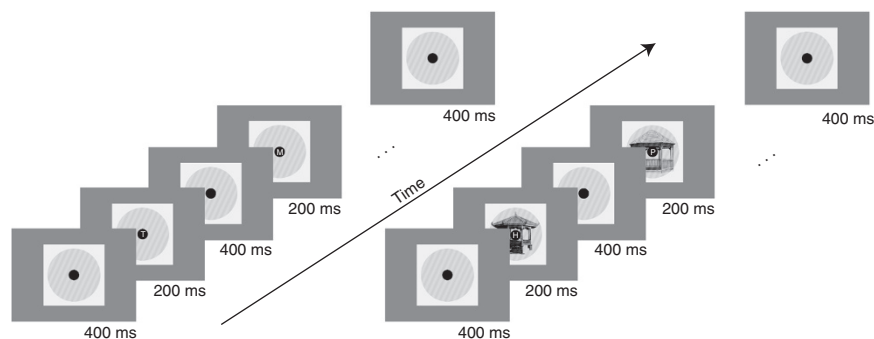


Figure 4 Stimuli and task for experiment 2. Participants were continuously shown a low contrast oriented grating that was either presented alone (trials without distractors) or overlaid with high-contrast distractor stimuli that flickered on and off following the distractor presentation timing during the delay period in experiment 1 (trials with distractors). Participants performed a one-back letter repetition detection task at fixation. Both the fixation dot and the letter have been enlarged in this figure for clarity.

Because the decoding accuracy for trials without distractors in the perceptual task was equivalent, if not lower, than that seen for the VSTM trials without distractors in experiment 1, our contrast manipulation replicated the strength of the memory representation fairly well. However, unlike in experiment 1, we saw no effect of distractors on decoding accuracy. Thus, occipital cortex was capable of simultaneously representing the contents of both gratings and distractors robustly, even though participants were never explicitly asked to encode either. These results indicate that the drop in VSTM decoding accuracy in occipital cortex in trials with distractors in experiment 1 was not a result of limitations in fMRI MVPA in decoding a weak memory stimulus among a much stronger distractor stimulus, but rather of a lack of robust VSTM representation when other visual stimuli had to be processed. Given the ubiquitous nature of distractors in our everyday visual environment, this vulnerability to distraction suggests that occipital cortex cannot be the primary storage region for VSTM.

Decoding of VSTM content with unpredictable distractors

If occipital cortex is capable of representing perceptually presented grating information while processing additional incoming visual stimuli, as shown in experiment 2, then what causes the drop in VSTM representation in this brain region in the face of distraction? One possibility is that the processing of incoming visual stimuli automatically weakens any VSTM representation present in occipital cortex. However, as the presence of distractors was fully predictable in experiment 1, it is also possible that, when participants know distractors will be present during the delay period, they can strategically choose not to engage occipital cortex in VSTM representation.

To test this idea, in experiment 3, we brought back the participants from experiment 1 and had them complete the exact same task, but we removed their ability to anticipate the upcoming trial type by randomly intermixing trials with and without distractors in each run. If the representation of VSTM information in occipital cortex reflects a particular task strategy, we should no longer see a difference in VSTM decoding accuracy in this brain region for the two trial types. Depending on whether or not participants still choose to engage occipital cortex in VSTM representation, VSTM decoding in occipital cortex could be either above or at chance level for both trial types. On the other hand, if VSTM representation in occipital cortex is always negatively affected by the presence of distractors, then, as in experiment 1, we would expect to see a significant difference in

decoding performance between the two trial types, with higher decoding accuracy seen in trials without distractors than in those with distractors.

Behavioral performance in this experiment was similar to that of experiment 1 ($t(9) = 1.1$, $P = 0.34$), with an average of 77.9% correct across all trials. As in experiment 1, there were no differences in performance ($t(9) = 1.1$, $P = 0.29$) between trials with (78.1%) and without distractors (77.7%).

However, unlike experiment 1, when we examined decoding accuracy for the remembered orientation during the delay period, we found above chance decoding in both V1–V4 and superior IPS for trials with (V1–V4: $t(9) = 3.5$, $P = 0.007$; superior IPS: $t(9) = 2.5$, $P = 0.03$) and without distractors (V1–V4: $t(9) = 2.7$, $P = 0.02$; superior IPS: $t(9) = 3.4$,

$P = 0.008$) and no significant differences between trial types (V1–V4: $t(9) = 0.7$, $P = 0.52$; superior IPS: $t(9) = 0.1$, $P = 0.94$) (Fig. 3). There was also no interaction between brain region and trial type ($F(1,9) = 0.3$, $P = 0.52$). As in experiment 1, decoding accuracy for the distractor category was significant in both V1–V4 (accuracy = 0.98, $t(9) = 100.6$, $P < 0.0001$) and superior IPS (accuracy = 0.88, $t(9) = 17.04$, $P < 0.0001$).

Our ability to decode VSTM contents in occipital cortex in trials with distractors here further supports our results from experiment 2. Combined together, these two experiments strongly suggest that the drop in decoding seen in trials with distractors in experiment 1 was not a result of a failure of fMRI MVPA to decode a memory stimulus amongst a stronger distractor stimulus, but rather a decrease in the memory representation.

A direct comparison between experiments 1 and 3 revealed an interaction between experiment, ROI and trial type ($F(1,9) = 6.13$, $P = 0.02$), indicating that, although VSTM decoding accuracy was consistently above chance in superior IPS across experiments and trial types, decoding accuracy in V1–V4 varied on the basis of the presence and predictability of distractors. This variability exists despite the fact that the task and trials were identical in the two experiments. Across the two experiments, as the presence of distractors became more predictable, VSTM decoding accuracy in V1–V4 decreased such that decoding accuracy of both trial types in experiment 3 were lower than that of trials without distractors in experiment 1 (experiment 3 trials without distractors: $t(9) = 2.1$, $P = 0.07$; experiment 3 trials with distractors: $t(9) = 4.5$, $P = 0.002$), but higher than that of trials with distractors in experiment 1 (experiment 3 trials without distractors: $t(9) = 2.0$, $P = 0.07$; experiment 3 trials with distractors: $t(9) = 2.1$, $P = 0.06$). These results suggest that the predictability of distractor presence governs whether or not participants choose to engage occipital cortex in VSTM representation. Given that behavioral VSTM performance in both experiments was unaffected by the presence and the predictability of distractors, these results suggest that superior IPS has a central role in VSTM storage, whereas VSTM representations seen in occipital cortex are unlikely to be essential.

VSTM decoding in other parietal regions

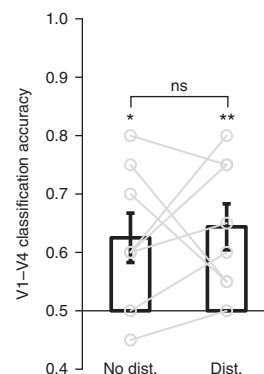
Unlike in superior IPS, none of the topographic IPS regions or anatomically defined superior and inferior parietal lobules (SPL and IPL, respectively) showed consistent decoding of memory information across distractor presence, absence or level of predictability

Figure 5 MVPA decoding results for experiment 2. Eight of the ten participants from experiments 1 and 3 took part in this experiment. Decoding accuracy for the presented grating was significantly above chance in both trials with and without distractors, with no differences seen between trial types, suggesting that MVPA can decode two simultaneously represented stimuli in occipital cortex. Error bars indicate s.e.m. ns, non-significant; * $P < 0.05$, ** $P < 0.01$; No dist., trials without distractors; Dist., trials with distractors.

(Supplementary Figs. 3–6). This suggests that VSTM storage may not be a general function of the parietal cortex, but rather may be specific to superior IPS. This also underscores the importance of appropriate ROI selection in understanding the role of parietal cortex in these types of higher order processes. If regions that are involved in VSTM, such as superior IPS, are combined with regions specialized for other processes, as in the large anatomically defined IPL and SPL regions, then our ability to detect VSTM representations in parietal cortex would be substantially hampered, resulting in an inaccurate depiction of the role of parietal cortex in VSTM.

Behavioral and neural VSTM correlations

The results from experiment 1 clearly show that orientation representations in occipital cortex are unrelated to behavioral performance on the task, as a sharp decrease in decoding was seen in trials with distractors with no concurrent disruption of behavioral performance. However, although we have established a similar null distractor effect for both decoding in superior IPS and behavioral performance, it remains unclear whether orientation representations in this brain region are directly related to behavioral VSTM performance. In experiment 4, we brought back a subset of the original participants to more directly examine this relationship. Each participant completed two experimental sessions, an MRI session and a behavioral session outside the MRI scanner. In both sessions, participants were shown and asked to remember a single orientated grating, followed by a mask to disrupt any lingering perceptual representation. Target orientations were drawn from a set of six orientations (10°, 40°, 70°, 100°, 130° and 160°). In the MRI experiment, after a delay, participants were asked to report the direction of a small rotation in the test stimulus relative to the remembered orientation, similar to what was done in experiments 1 and 3. We obtained VSTM decoding accuracies during the delay period for each possible pair of orientations in both our V1–V4 and superior IPS ROIs, creating a neural orientation representation similarity matrix for each ROI. In the behavioral task, after a delay, participants were asked to report whether there was an orientation change, which occurred in half of the trials. The test orientation was drawn from the same set of six orientations as the target orientation and, in change trials, it came equally often from the five remaining orientations. The larger the angular difference between the remembered and the test orientations, the faster participants are



able to respond. Using these reaction time measures, we constructed a behavioral orientation representation similarity matrix. We then calculated the correlation between the neural and behavioral representation similarity matrices^{29,30}. If VSTM representations in a brain region were directly related to VSTM behavioral performance, then the distinctiveness of a pair of orientation representations in that brain region should directly correlate with how fast participants could tell them apart in the behavioral change detection task, resulting in a negative correlation between the two measures (that is, the bigger the neural representational difference, the shorter the reaction time).

Indeed, we found strong negative correlations between decoding and behavioral performance for both V1–V4 and superior IPS (V1–V4: $r = -0.7$, $P = 0.002$; superior IPS: $r = -0.59$, $P = 0.009$, permutation tests for both; Fig. 6). These results held even if we removed the first time point used for the average delay activity (V1–V4: $r = -0.68$, $P = 0.004$; superior IPS: $r = -0.51$, $P = 0.03$, permutation tests for both), suggesting that these results are not driven by any lingering encoding period activity. Thus, as the VSTM representations in superior IPS and V1–V4 become harder to distinguish, behavioral reaction time increases. Combined with the results of our other experiments, this strongly supports the idea that superior IPS has a central role in the storage of information into VSTM. Being a VSTM region, superior IPS is unlikely to be involved in the initial computation and representations of the orientation information. Rather, such information must be processed elsewhere (for example, V1–V4) and uploaded into superior IPS when it needs to be retained in VSTM. It is therefore not surprising that delay representations in V1–V4 also correlated with behavioral performance. However, experiment 1 clearly shows that such representations cannot reliably support successful information retention in VSTM.

Target-distractor similarity and behavioral performance

If delay period representations in occipital cortex are important in VSTM, then, because occipital cortex must also process incoming distractor

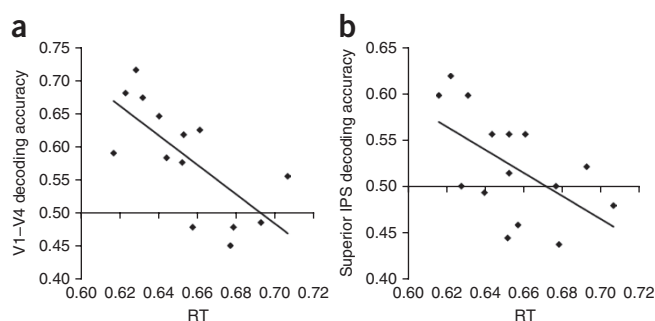
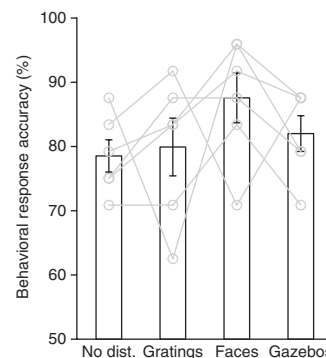


Figure 6 Correlation of neural and behavioral VSTM representations from experiment 4. Six participants from experiment 1 took part in this experiment. (a,b) Both V1–V4 (a) and superior IPS (b) showed strong negative correlations between behavioral (RT) and neural (decoding accuracy) measures of VSTM representation similarity across the six orientations tested, showing that the more similar a pair of orientation representations are in these brain regions during the VSTM delay period, the harder it is to discriminate them behaviorally in a change-detection task. In V1–V4, two pairs of orientation representations (40–160° and 130–160°) had identical RTs and decoding accuracies, and so both points occupy the same place in the graph. These results establish a link between VSTM representations in both brain regions and behavioral VSTM performance when distractors were absent during the delay period.

Figure 7 Accuracy results for experiment 5. Six participants from experiment 1 took part in this experiment. In this behavioral experiment, the presence and absence of distractors during the VSTM delay period, as well as the similarity between the target and distractors, were varied. There was no difference in accuracy, as measured by percent correct for any distractor condition, nor did any distractor condition differ from the no distractor condition, indicating that neither distractor presence/absence nor target-distractor similarity affected performance. Error bars indicate s.e.m. No dist., trials without distractors; faces, trials with face distractors during the delay; gazebos, trials with gazebo distractors; gratings, trials with oriented grating distractors.

information, the more similar the distractors are to the targets, the more they should share similar neural processing substrates and compete for representation. However, if delay period representations in occipital cortex are not a central component of neural VSTM representation, then the competition caused by target-distractor similarity would minimally affect VSTM performance. To test these predictions, in experiment 5, we asked six participants from experiments 1 and 3 to complete a behavioral version of the oriented grating VSTM task in experiment 3 and cued them to remember one of two sequentially presented grating stimuli ($\sim 25^\circ$ or $\sim 115^\circ$). Following the parameters used in experiment 3, during the extended delay period (400 ms after the offset of the target stimulus), in addition to viewing a series of faces, a series of gazebos or simply a fixation dot (no distractor trials), the subjects also viewed a series of oriented gratings. Trials containing the different distractor conditions were randomly intermixed in a given run, just as in experiment 3. The oriented grating distractors differed from the to-be-remembered target gratings only in orientation and were drawn from a set of six orientations covering the entire visual field in 30° increments (0° , 30° , 60° , 90° , 120° and 150°). During the delay period, the entire set of grating distractors was presented roughly three times. The presentation of the grating distractors would therefore activate similar neural processing substrates as the target orientation, maximally masking the activation of the remembered orientation. Among the three types of distractors shown, the grating distractors imposed the greatest representation competition in occipital cortex. However, we found no significant differences in VSTM performance between any of the distractor conditions (no distractor versus grating distractor trials: $t(5) = 0.3$, $P = 0.81$; no distractor versus face distractor trials: $t(5) = 2.0$, $P = 0.10$; no distractor versus gazebo distractor trials: $t(5) = 1.8$, $P = 0.14$; Fig. 7). Thus, the type of distractors present during the delay did not affect behavioral VSTM performance. These behavioral results further suggest that delay period representations in occipital cortex cannot be a central component of neural VSTM representation, reaffirming our fMRI decoding results.

Although previous behavioral studies have reported that passive viewing of a distractor similar to the remembered target negatively affected VSTM performance, a close examination of these results reveal that distractor interference is mainly present during the early delay period, when sensory information is still being consolidated into VSTM, whereas late delay period activity seems to be resilient to distractor interference^{1,31}. Thus, prior work actually argues against the idea that sensory areas are recruited and are necessary for VSTM maintenance. Instead, they suggest that once encoding is complete, stored information in VSTM is resilient to distraction. As our distractor stimuli were shown after the consolidation processing was completed³¹, our results further argue that consolidated VSTM representations are protected against incoming perceptual interference. Given that distracting visual information is ubiquitous in the real world, this protection against interference is an essential feature of VSTM if it is to have any substantial role in real world vision.



DISCUSSION

Using fMRI response amplitude measures, previous reports have highlighted the role of superior IPS in maintaining VSTM representations^{2–6}. In contrast, using fMRI MVPA measures, recent studies have revealed VSTM representations in occipital cortex^{10,11,13,16,17,19}. In this set of fMRI MVPA studies, we critically evaluated the contribution of both occipital and parietal cortices to VSTM representation and storage by varying the presence and predictability of distractors during the delay period of an oriented grating VSTM task. Although distractor presence predictably did not affect behavioral performance, it significantly affected VSTM decoding in occipital cortex. Specifically, when the presence of distractors was predictable in experiment 1, we found strong VSTM decoding during the delay period in occipital cortex when distractors were absent, but we observed a significant drop, to chance level decoding, when distractors were present. This drop in VSTM decoding was not simply a failure of MVPA to resolve a weak VSTM pattern among a stronger distractor pattern, as distractor processing had no effect on the successful decoding of perceptually presented weak oriented gratings in occipital cortex in experiment 2. Moreover, when distractor presence was no longer predictable in experiment 3, equal VSTM decoding was seen in occipital cortex in trials with and without distractors. Decoding accuracy was lower in experiment 3 than the trials without distractors in experiment 1, but still significantly above chance. Thus, processing incoming task-irrelevant visual stimuli does not automatically degrade VSTM representations in occipital cortex. Rather, the predictability of distractor presence allows participants to strategically decide whether or not to engage occipital cortex in VSTM representation. In the decoding analyses for experiments 1 and 3, we only tested whether a brain region represented left-tilted gratings differently than right-tilted gratings; however, in our task, we required participants to perform the much harder classification of a 3° or 6° orientation change. Given the high level of precision required in VSTM representation to perform this task, if a region shows poor decoding for the gross left versus right orientation discrimination, as occipital cortex does in experiment 1 when distractor presence is predictable, then it would be unlikely for it to support robust behavioral performance on the much harder VSTM task we gave our participants. In contrast, superior IPS mirrored behavioral performance and showed equally strong VSTM decoding independent of distractor presence and predictability. These results indicate that superior IPS, and not occipital cortex, may have a central role in supporting VSTM storage.

Using fMRI representation similarity measures^{29,30}, experiment 4 further showed that VSTM representation (in the absence of distraction) in superior IPS closely tracked behavioral performance on a VSTM task, thereby establishing a link between neural representation and behavior in this brain region. This neural-behavioral link is a necessary feature of any region that has a central role in VSTM

information maintenance and this finding therefore underscores the importance of the role of superior IPS in VSTM storage. Delay period representations in occipital cortex also reflected behavioral VSTM measures in the absence of distractors, likely as a result of its role in the initial processing of the orientation information. However, as experiment 1 clearly revealed, such representations cannot reliably support successful information retention in VSTM.

By manipulating the similarity between the target and distractor stimuli in a behavioral study, we also found that the type of distractors present during the delay period did not affect behavioral VSTM performance. This is consistent with previous studies showing that consolidated VSTM representations, which were what we tested, are largely immune to incoming perceptual interference, regardless of the target-distractor similarity^{1,31}. This behavioral evidence further speaks against the role of occipital cortex in VSTM information representation, as occipital cortex is necessarily involved in the processing of incoming distractor information and the similarity between the target and distractor stimuli should negatively affect VSTM representation in this brain region. Overall, our findings reestablish the important contribution of superior IPS to VSTM representation and argue against the notion that occipital cortex has a central role in maintaining VSTM information.

Our findings echo a recent neurophysiological finding in macaques³². Using a VSTM task with motion stimuli, the authors found that the spiking activity in direction-selective neurons in the middle temporal (MT) area did not reflect the memorized motion direction. Similar to V1–V4, MT is a primary sensory region. Instead, the authors found that VSTM information was only present in the spiking activity of higher order, multimodal areas. However, VSTM-related local field potential (LFP) activity was present in both MT and higher order areas. As previous work has linked fMRI activity primarily to LFPs³³, the authors³² reasoned that LFP activity was likely the source of VSTM decoding in occipital cortex in fMRI studies. However, given the lack of single-unit activity in these same regions, the authors argued that such LFP activity, and the fMRI decoding findings based on that activity, did not reflect VSTM storage, but instead reflected another process, which they suggested could be an attentional priority map. Despite the limitations in the types of neural activity fMRI can measure, our stimulus manipulation let us reach the same conclusions as those of the previous study³²: that occipital VSTM representations do not reflect the primary storage of VSTM information. Thus, with appropriate stimulus manipulations and experimental design, fMRI is capable of revealing the nature of visual information processing in the human brain. In parietal cortex, both single-unit and LFP activity related to VSTM representation has been reported¹, suggesting that this brain region is important for VSTM representation. This is again consistent with our results. Together, our fMRI findings and those from neurophysiological recording studies³² provide converging evidence indicating that higher order multimodal areas, and not primary sensory regions, are critical for the storage of VSTM information.

fMRI MVPA operates on the assumption that neurons selective for the different features are distributed differentially across different voxels. As such, an inappropriate fMRI voxel resolution may result in the lack of heterogeneity among the voxels and null results from a brain region that would be otherwise important in a neural process. Although this is an important limitation of fMRI MVPA, in each of our ROIs, we were able to obtain robust VSTM decoding in at least one condition across the experiments. Given that feature distribution in a brain region would not change rapidly enough with our experimental

manipulations to create the differences that we observed, a change in decoding accuracy must then be a result of how the brain region participates in the task under the different experimental conditions. Thus, at least for the regions that we examined in the present study, the resolution of our MRI voxels does not appear to impede our ability to decode VSTM representations.

Previous work has shown that mental imagery activates occipital cortex^{34,35}, and that imagery, perception and VSTM all share similar representations in occipital cortex²⁸. Notably, individuals with poor mental imagery skills show lower VSTM decoding in occipital cortex than those who excel at imagery²⁸. Thus, it has been argued that individuals with strong mental imagery may rely on imagery to support VSTM performance, whereas those with poor imagery may rely on different strategies³⁶. In our VSTM tasks, it is likely that mental imagery-based visual rehearsal was deployed in memory delay periods when distractors were known to be absent, less so when distractor presence was unpredictable, and minimally when distractors were known to be present. However, this strategy ultimately produced no noticeable behavioral benefit and therefore does not seem to be a necessary component of VSTM.

Although previous MVPA studies have produced mixed results regarding the role of the parietal cortex in VSTM information representation^{10,11,15,16,19,23,24}, none of them specifically targeted the superior IPS, a key parietal region whose response amplitude tracks VSTM storage^{2–6}. We found that, mirroring behavioral performance, VSTM representations could be consistently decoded from superior IPS regardless of the presence and the predictability of distractors. No other parietal regions showed such reliable VSTM decoding, including parietal topographic maps in IPS and anatomically defined IPL and SPL. This may explain why previous attempts have failed to reveal consistent VSTM decoding in parietal cortex when superior IPS was not targeted and highlights the importance of appropriate ROI selection in understanding the role of parietal cortex in visual cognition.

Parietal cortex has also long been associated with attention-related processing^{37–41}. Our results suggest that one way parietal cortex may participate in attention-related information processing is by directly representing task-relevant VSTM information in superior IPS. One can argue that parietal cortex may simply contain an attentional template that tracks what is behaviorally relevant. However, as such an attentional template has to be distinct for the different orientation gratings shown and has to be maintained for a prolonged period in the absence of any visual stimulation, it is unclear how such an attentional template would differ fundamentally from a VSTM representation.

To conclude, we found that MVPA decoding in superior IPS, but not occipital cortex, closely tracked behavioral measures of information storage in VSTM across distractor presence and predictability. This suggests that superior IPS, and not occipital cortex, has a central role in VSTM storage in the human brain.

METHODS

Methods and any associated references are available in the [online version of the paper](#).

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We would like to thank J. Swisher for his retinotopy code and members of the Harvard Vision Lab for their valuable comments on this study. This research was supported by US National Institutes of Health (NIH) grant 1R01EY022355 to Y.X. and NIH grant F32-EY022874 to K.C.B.

AUTHOR CONTRIBUTIONS

K.C.B. and Y.X. designed the experiments. K.C.B. conducted the experiments and analyzed the data. K.C.B. and Y.X. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

- Pasternak, T. & Greenlee, M.W. Working memory in primate sensory systems. *Nat. Rev. Neurosci.* **6**, 97–107 (2005).
- Todd, J.J. & Marois, R. Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature* **428**, 751–754 (2004).
- Xu, Y. Distinctive neural mechanisms supporting visual object individuation and identification. *J. Cogn. Neurosci.* **21**, 511–518 (2009).
- Xu, Y. & Chun, M.M. Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature* **440**, 91–95 (2006).
- Jeong, S.K. & Xu, Y. Neural representation of targets and distractors during object individuation and identification. *J. Cogn. Neurosci.* **25**, 117–126 (2013).
- Xu, Y. The neural fate of task-irrelevant features in object-based processing. *J. Neurosci.* **30**, 14020–14028 (2010).
- Todd, J.J. & Marois, R. Posterior parietal cortex activity predicts individual differences in visual short-term memory capacity. *Cogn. Affect. Behav. Neurosci.* **5**, 144–155 (2005).
- Beck, D.M., Muggleton, N., Walsh, V. & Lavie, N. Right parietal cortex plays a critical role in change blindness. *Cereb. Cortex* **16**, 712–717 (2006).
- Tseng, P. *et al.* Posterior parietal cortex mediates encoding and maintenance processes in change blindness. *Neuropsychologia* **48**, 1063–1070 (2010).
- Ester, E.F., Sprague, T.C. & Serences, J.T. Parietal and frontal cortex encode stimulus-specific mnemonic representations during visual working memory. *Neuron* **87**, 893–905 (2015).
- Emrich, S.M., Riggall, A.C., LaRocque, J.J. & Postle, B.R. Distributed patterns of activity in sensory cortex reflect the precision of multiple items maintained in visual short-term memory. *J. Neurosci.* **33**, 6516–6523 (2013).
- Ester, E.F., Serences, J.T. & Awh, E. Spatially global representations in human primary visual cortex during working memory maintenance. *J. Neurosci.* **29**, 15258–15265 (2009).
- Harrison, S.A. & Tong, F. Decoding reveals the contents of visual working memory in early visual areas. *Nature* **458**, 632–635 (2009).
- Lee, S.H., Kravitz, D.J. & Baker, C.I. Goal-dependent dissociation of visual and prefrontal cortices during working memory. *Nat. Neurosci.* **16**, 997–999 (2013).
- Linden, D.E., Oosterhof, N.N., Klein, C. & Downing, P.E. Mapping brain activation and information during category-specific visual working memory. *J. Neurophysiol.* **107**, 628–639 (2012).
- Riggall, A.C. & Postle, B.R. The relationship between working memory storage and elevated activity as measured with functional magnetic resonance imaging. *J. Neurosci.* **32**, 12990–12998 (2012).
- Serences, J.T., Ester, E.F., Vogel, E.K. & Awh, E. Stimulus-specific delay activity in human primary visual cortex. *Psychol. Sci.* **20**, 207–214 (2009).
- Sneve, M.H., Alnaes, D., Endestad, T., Greenlee, M.W. & Magnussen, S. Visual short-term memory: activity supporting encoding and maintenance in retinotopic visual cortex. *Neuroimage* **63**, 166–178 (2012).
- Sprague, T.C., Ester, E.F. & Serences, J.T. Reconstructions of information in visual spatial working memory degrade with memory load. *Curr. Biol.* **24**, 2174–2180 (2014).
- Hou, Y. & Liu, T. Neural correlates of object-based attentional selection in human cortex. *Neuropsychologia* **50**, 2916–2925 (2012).
- Liu, T., Hospadaruk, L., Zhu, D.C. & Gardner, J.L. Feature-specific attentional priority signals in human cortex. *J. Neurosci.* **31**, 4484–4495 (2011).
- Xu, Y. & Jeong, S.K. The contribution of human superior intraparietal sulcus to visual short-term memory and perception. in *Mechanisms of Sensory Working Memory: Attention and Performance XXV* (eds. Jolicoeur, P., Lefebvre, C. & Martinez-Trujillo, J.) 33 (Academic Press, 2015).
- Christophel, T.B., Hebart, M.N. & Haynes, J.-D. Decoding the contents of visual short-term memory from human visual and parietal cortex. *J. Neurosci.* **32**, 12983–12989 (2012).
- Nelissen, N., Stokes, M., Nobre, A.C. & Rushworth, M.F. Frontal and parietal cortical interactions with distributed visual representations during selective attention and action selection. *J. Neurosci.* **33**, 16443–16458 (2013).
- Jacob, S.N. & Nieder, A. Complementary roles for primate frontal and parietal cortex in guarding working memory from distractor stimuli. *Neuron* **83**, 226–237 (2014).
- Suzuki, M. & Gottlieb, J. Distinct neural mechanisms of distractor suppression in the frontal and parietal lobe. *Nat. Neurosci.* **16**, 98–104 (2013).
- Qi, X.L., Elworthy, A.C., Lambert, B.C. & Constantinidis, C. Representation of remembered stimuli and task information in the monkey dorsolateral prefrontal and posterior parietal cortex. *J. Neurophysiol.* **113**, 44–57 (2015).
- Albers, A.M., Kok, P., Toni, I., Dijkerman, H.C. & de Lange, F.P. Shared representations for working memory and mental imagery in early visual cortex. *Curr. Biol.* **23**, 1427–1431 (2013).
- Kriegeskorte, N. *et al.* Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron* **60**, 1126–1141 (2008).
- Kriegeskorte, N. & Kievit, R.A. Representational geometry: integrating cognition, computation, and the brain. *Trends Cogn. Sci.* **17**, 401–412 (2013).
- Vogel, E.K., Woodman, G.F. & Luck, S.J. The time course of consolidation in visual working memory. *J. Exp. Psychol. Hum. Percept. Perform.* **32**, 1436–1451 (2006).
- Mendoza-Halliday, D., Torres, S. & Martinez-Trujillo, J.C. Sharp emergence of feature-selective sustained activity along the dorsal visual pathway. *Nat. Neurosci.* **17**, 1255–1262 (2014).
- Logothetis, N.K., Pauls, J., Augath, M., Trinath, T. & Oeltermann, A. Neurophysiological investigation of the basis of the fMRI signal. *Nature* **412**, 150–157 (2001).
- Kosslyn, S.M., Ganis, G. & Thompson, W.L. Neural foundations of imagery. *Nat. Rev. Neurosci.* **2**, 635–642 (2001).
- Slotnick, S.D., Thompson, W.L. & Kosslyn, S.M. Visual mental imagery induces retinotopically organized activation of early visual areas. *Cereb. Cortex* **15**, 1570–1583 (2005).
- Keogh, R. & Pearson, J. Mental imagery and visual working memory. *PLoS One* **6**, e29221 (2011).
- Wojciliuk, E. & Kanwisher, N. The generality of parietal involvement in visual attention. *Neuron* **23**, 747–764 (1999).
- Yantis, S. *et al.* Transient neural activity in human parietal cortex during spatial attention shifts. *Nat. Neurosci.* **5**, 995–1002 (2002).
- Behrmann, M., Geng, J.J. & Shomstein, S. Parietal cortex and attention. *Curr. Opin. Neurobiol.* **14**, 212–217 (2004).
- Colby, C.L. & Goldberg, M.E. Space and attention in parietal cortex. *Annu. Rev. Neurosci.* **22**, 319–349 (1999).
- Culham, J.C. & Kanwisher, N.G. Neuroimaging of cognitive functions in human parietal cortex. *Curr. Opin. Neurobiol.* **11**, 157–163 (2001).

ONLINE METHODS

Participants. Ten paid participants (seven female) from the Harvard University community were recruited to participate in experiments 1 and 3. Six (five female) of those also completed experiments 4 and 5. Finally, eight (six female) of the ten also completed experiment 2. All participants gave informed consent in accordance with the Institutional Review Board of Harvard University. Participants were between 23 and 36 years old (mean age = 29.5). All had normal or corrected-to-normal visual acuity, all were right-handed, and received payments for their participation.

VSTM experiments examining distractor presence and predictability (experiments 1 and 3). The design of experiments 1 and 3 was adapted from the delayed orientation discrimination task (Fig. 1) used previously¹³. In each trial, participants saw a sequential presentation of two centrally presented sine-wave gratings one at ~25° and the other at ~115° (radius, 5° of visual angle; contrast, 20%; spatial frequency, 1 cycle per degree), followed by a numerical cue (1 or 2) that indicated which grating they were to remember, first or second. The presentation order of the two gratings and whether the first or second grating would be cued were counterbalanced within each run. After an extended delay, participants were asked to report the direction of rotation ($\pm 3^\circ$ or $\pm 6^\circ$) of a test grating relative to the cued grating. The grating was rotated equally often to the left or to the right, but the amount of rotation (3° or 6°) was random. The precise timing of each trial was as follows: first sample grating (200 ms), blank (400 ms), second sample grating (200 ms), blank (400 ms), cue (800 ms), delay (11 s), test grating (500 ms), response (2,000 ms), and feedback (500 ms). Feedback was given after every trial as either a happy face (for correct trials) or a sad face (for error trials). Each trial lasted 16 s and was followed by 16 s of fixation to allow the hemodynamic response to return to baseline. To ensure proper fixation, a black fixation dot was present throughout the trial and the inter-trial fixation period (this applied to all the experiments reported here except where noted). To alert and prepare participants for the upcoming stimulus trial, during the last 500 ms of the fixation period, the fixation dot turned from black to red (this design feature was implemented in all the experiments reported here). We used eye tracking to monitor the gaze of each participant to ensure that participants maintained fixation throughout the trial (this applied to all the fMRI experiments reported here). Each run (296 s) consisted of eight trials plus an initial practice trial that was excluded from later analyses.

Half of the trials were trials without distractors in which no stimuli were presented during the delay, identical to what was done previously¹³. The other half of the trials were trials with distractors in which a sequential presentation of 17 faces or gazebos were shown throughout the entire delay. The delay began and ended with 400 ms of blank fixation. Each distractor was presented for 200 ms, followed by a 400-ms blank and subtended $6.9^\circ \times 8.3^\circ$ of visual angle. In experiment 1, trials with and without distractors were shown in separate runs with each participant first completing all eight runs of trials without distractors, followed by all eight runs of trials with distractors, with nine trials (one practice) in each run. This was done so that performance on trials without distractors would not be influenced by any strategy that participants might develop after completing trials with distractors, thus giving us the best chance to replicate previous work¹³. In trials with distractors, trials containing face distractors (50%) and those containing gazebo distractors (50%) were randomly intermixed. In experiment 3, trials with and without distractors were randomly intermixed in each run. The random nature prevented participants from being able to accurately anticipate whether a distractor would be present for any given trial. Half the trials (four) in each run had no distractors, and half had distractors (with two containing face distractors and two containing gazebo distractors). Thus, the total number of trials of each type was matched to that of experiment 1. Within each run the two grating orientations appeared equally often and, when distractors were present, appeared equally often with the two types of distractors. In both experiments 1 and 3, each participant completed a total of 16 runs, each lasting 4 min and 56 s.

Control experiment with overlapping perceptual stimuli (experiment 2). The goal of this experiment was to perceptually recreate the processing that would have occurred during the delay period of experiment 1, to examine whether or not MPVA can decode a weak perceptual stimulus superimposed by a strong perceptual stimulus. Using a block design, with each block lasting 12 s, we

presented a semi-transparent (25% opacity) oriented grating (25° or 115°) for an entire block. This grating was either presented alone (trials without distractors) or overlaid with a sequential presentation of 20 faces or gazebos (75% opacity) timed to match the presentation of the distractor stimuli during the delay period of experiment 1 (on for 200 ms and off for 400 ms) (Fig. 4). The opacity of the grating and distractor stimuli was intended to simulate the weaker representation of the stored grating in memory and any possible suppression of the distractor stimuli. To further challenge the decoding ability of MPVA, we diverted participants' attention away from the gratings and distractors by asking them to attend to a sequential presentation of letters at fovea, timed to match the presentation of the distractor stimuli in experiment 1. Thus, the presentation of the letters coincided with that of the distractors, although the presentation of the grating was kept visible and constant across the entire block to simulate the grating memory representation formed during the delay period of experiment 1. Participants performed a 1-back letter repetition detection task on the letter stream and responded with a button press. As in experiments 1 and 3, within each run the two grating orientations appeared equally often and, when distractors were present, appeared equally often with the two types of distractors. Each run lasted 3 min and 20 s, containing eight 12-s stimulus blocks alternating with eight 12-s fixation blocks in which only the fixation dot was present. In one session, participants completed ten runs of blocks with distractors and, in a separate session, completed ten runs of blocks without distractors.

Correlation between VSTM decoding and behavioral performance experiment (experiment 4). To examine whether VSTM representations formed in each brain region during the delay period are related to VSTM behavioral performance, we had participants complete two separate tasks, one fMRI experiment to measure the neural representation similarity, and one behavioral experiment outside the MRI to measure the behavioral representation similarity. In the fMRI experiment, participants completed a task very similar to what was done in experiments 1 and 3, but with only one grating stimulus presented in each trial, followed by a mask and no distractors during the delay. In each trial, participants saw a brief presentation of one grating in one of six orientations, 10°, 40°, 70°, 100°, 130° and 160°, followed by a briefly presented plaid mask containing two overlapping orientations, 0° and 90° for 200 ms. After a delay of 11.4 s, participants were asked to report the direction of rotation ($\pm 3^\circ$ or $\pm 6^\circ$) of a test grating relative to the remembered grating. Feedback was provided after every trial. The precise timing of each trial was as follows: sample grating (200 ms), blank (200 ms), mask (200 ms), delay (11.4 s), test grating (500 ms), response (2,000 ms) and feedback (500 ms). Each trial lasted 15 s and was followed by 15 s of fixation to allow the hemodynamic response to return to baseline. Neural measures of similarity were calculated based on the decoding accuracy between each pair of orientations (see below).

In the behavioral experiment, participants were also shown a single oriented grating (drawn from the same six orientations, 10°, 40°, 70°, 100°, 130° and 160°), followed by the same plaid mask as in the fMRI experiment. After a short delay, participants performed a same-different judgment on a test grating drawn from the same set of orientations as the sample grating. Feedback was provided after each trial. The precise timing of each trial was as follows: sample grating (200 ms), blank (200 ms), mask (200 ms), delay (1,000 ms), test grating (500 ms), response (2,000 ms) and feedback (300 ms). Each trial lasted 4.4 s and was followed by 1 s of fixation. Reaction time was recorded as a behavioral measure of the similarity between the sample and test orientations. Each run consisted of 60 trials, plus one practice trial at the beginning of the run. Participants completed a total of eight runs. In each run, there were an equal number of change and no change trials, and in the change trials each orientation was paired equally often with all the other orientations. Anytime participants made an incorrect response, a red unhappy face flickered on and off for 5 s and the trial was repeated at the end, until correct responses were obtained for all trials in the run. Reaction time measures were calculated from all the correct trials, and only reaction time, and not accuracy, was included in further analysis.

Behavioral control experiment using multiple different distractor types (experiment 5). The goal of this experiment was to examine how different types of distractors might affect behavioral performance in the main VSTM task. Trial structure and timing of this experiment was identical to experiments 1 and 3. Specifically, participants were shown sequential, brief presentations of

two centrally presented sine-wave gratings at $\sim 25^\circ$ or $\sim 115^\circ$ (radius, 5° of visual angle; contrast, 20%; spatial frequency, 1 cycle per degree) in a randomized order, followed by a numerical cue that indicated which grating they were to remember, first or second. After an extended delay, participants were asked to report the direction of rotation ($\pm 3^\circ$ or $\pm 6^\circ$) of a test grating relative to the cued grating. Feedback was provided after every trial. As in experiments 1 and 3, the precise timing of each trial was as follows: first sample grating (200 ms), blank (400 ms), second sample grating (200 ms), blank (400 ms), cue (800 ms), delay (11 s), test grating (500 ms), response (2,000 ms), and feedback (500 ms). Unlike in experiments 1 and 3, however, here the inter-trial interval was only 1 s, instead of 16 s, as we did not need to wait for the hemodynamic response to return to baseline in this behavioral experiment. During the delay, participants either saw a blank screen with a fixation dot (no distractor trials), or a sequential presentation of either 17 faces, 17 gazebos, or 17 oriented gratings. The face and gazebo stimuli were drawn from the same set as in experiments 1 and 3. The oriented distractor gratings differed from the to-be-remember target gratings only in orientation and they were drawn from a set of six covering the entire visual field in 30° increments (0° , 30° , 60° , 90° , 120° and 150°). As in experiments 1 and 3, each distractor was presented for 200 ms, followed by a 400-ms blank and subtended $6.9^\circ \times 8.3^\circ$ of visual angle. As in experiment 3, trials with and without distractors were randomly intermixed in each run. Each run contained a total of 33 trials, 8 for each trial type (no distractor, face distractor, gazebo distractor and grating distractor), plus one practice trial at the beginning of the run. Participants completed three runs each, each lasting 9 min and 23 s.

Localizer experiments. To identify topographic regions in occipital and parietal cortices, we mapped topographic visual field representations of polar angle for each participant with flashing checkerboard stimuli using standard techniques^{42–45}. To reveal maps in parietal cortex, we optimized our parameters following a previous study⁴⁵. Specifically, the colored polar angle wedge swept across the entire screen ($23.4 \times 17.5^\circ$ of visual angle), had an arc of 72° , a sweep period of 55.467 s, flashed at 4 Hz, and swept out 12 cycles per run (Fig. 2a). Each participant completed 4–6 runs (each lasting 11 min 5.6 s). The task varied slightly across participants. All participants were asked to detect a dimming in the visual display, for some participants the dimming occurred only at fixation, for others it occurred only within the polar angle wedge, and for some it could occur in both locations, commiserate with the various methodologies used in the literature^{45,46}. No differences were seen in the maps obtained through each of these methods.

To identify the superior IPS region previously shown to be involved in VSTM storage^{3–5,47}, we followed the procedures used previously⁴. Participants completed a VSTM object experiment, similar to the sequential central presentation shape experiments used previously⁴. In each trial, participants saw 1, 2, 3, 4 or 6 real world objects presented sequentially at central fixation, and after a short delay, judged whether a probe object shown at fixation was present in the original display (Fig. 2b). Eight distinctive objects were used, each subtended $7.9^\circ \times 7.5^\circ$ of visual angle and presented on a light gray background. Each trial lasted 6 s and consisted of fixation (500 ms), a sample display period, consisting of eight possible stimulus presentation slots (100 ms each, followed by 50-ms blank, for a total of 1,150 ms), a blank delay period (1,000 ms), a test display and response period (2,000 ms), and response feedback (1,350 ms). Each run also contained blank fixation trials (6 s). Trial presentation order was pseudorandom and balanced for trial history in a run^{2,4}. Participants completed 2–4 runs, with each run lasting 7 min and 42 s. Those completing four runs were those for whom superior IPS could not be localized reliably with two runs. The additional runs allowed us to obtain comparable number of voxels in superior IPS across all participants. Each run contained 76 trials, 12 per set size, plus 2 practice trials at the beginning of the run and 2 filler trials, one at the beginning and one at the end of the run, for trial history balancing purposes. Practice and filler trials were removed from further data analysis.

MRI methods. Stimuli were generated by a Macintosh MacBook Pro and back projected onto a screen mounted at the rear end of the scanner bore. Topographic mapping stimuli were presented using VisionEgg software⁴⁸, whereas all other stimuli were presented using Matlab with Psychtoolbox extensions⁴⁹. All data were acquired on a Siemens Tim Trio 3T scanner with a 32-channel head coil at the Center for Brain Science at Harvard University. Participants took part in two or three sessions of MRI scanning. In one session, a high resolution ($1.0 \times 1.0 \times 1.3$ mm) anatomical image was collected for surface reconstruction. Before functional

imaging in each session, T1-weighted echo-planar images were collected in the same slice prescription as the functional scans to allow each session to be registered to the participant's high-resolution anatomical scan. Functional data were acquired using T2*-weighted gradient-echo, echo-planar sequences. Each volume of the main experimental data, for all four fMRI experiments (experiments 1–4), contained 28 slices (3 mm thick, 3×3 mm in plane, no skip) oriented just off parallel from the AC-PC line to ensure full occipital and parietal coverage (TR = 2 s, TE = 35 ms, flip angle = 80°). For the topographic IPS localizer, each volume of the topographic data contained 42 slices (3 mm thick, 3.125×3.125 mm in plane, no skip) oriented just off parallel from the AC-PC line to cover the full brain (TR = 2.6 s, TE = 30 ms, flip angle = 90°). For the superior IPS localizer, each volume contained 24 slices (5 mm thick, 3.75×3.75 mm in plane, no skip) parallel to the AC-PC line (TR = 1.5 s, TE = 30 ms, flip angle = 90°). fMRI data were analyzed using the Freesurfer software package^{50–55}. Data preprocessing included motion correction, slice timing correction, linear drift correction and intensity normalization. Computer representations of each cortical hemispheric surface were unfolded and inflated. All data was analyzed in the native space of each participant.

ROI definitions. By following the procedures described previously⁴⁵ and by examining phase reversals in the polar angle maps, we were able to identify topographic areas in occipital and parietal cortices including V1, V2, V3, V4, V3A, V3B, IPS0, IPS1, IPS2, IPS3 and IPS4 in each participant (Fig. 2a,c).

To identify superior IPS, which has previously been shown to track VSTM storage^{3–5,47}, fMRI data from the superior IPS localizer was analyzed using a multiple regression analysis with the regression coefficient for each set size weighted by each participant's corresponding behavioral K score estimate for that set size calculated using Cowan's K (ref. 56). Superior IPS was defined as a region that showed significant activation in the regression analysis overlapping or near the region previously reported in Talairach coordinates^{2,4}. Although we originally set the threshold for activation in this region to $P < 0.05$, this created ROIs that, for some participants, contained too few voxels for the MVPA analysis (see below for more details), and so the threshold was relaxed to $P < 0.1$ to create a larger ROI (Fig. 2d). This ROI was localized individually in each participant.

Besides the above ROIs, we also created anatomically defined ROIs corresponding to the SPL and IPL (Fig. 2e). These regions were defined using Freesurfer's automatic parcellation⁵⁷.

Previously, occipital ROIs were further refined¹³ by using eccentricity data to select the region of each ROI that corresponded to the location of the stimuli to be remembered. As eccentricity data has been shown to be somewhat unreliable in parietal regions⁴⁵, in order to perform the same kind of refinement on our ROIs, we selected a subregion of each ROI (superior IPS, SPL, IPL and parietal and occipital topographic regions) that showed higher activation ($P < 0.05$) during the encoding period relative to fixation in trials without distractors. This contrast allowed us to select voxels that were visually responsive to the location of the grating stimuli without any contamination from distractor stimuli. Separate ROIs were created for experiments 1, 3, and 4 based on the activity in the respective experiments. Experiment 2 used the V1–V4 ROI defined in experiment 1. For one participant, V3A showed no task related activity in experiment 3 in the left hemisphere and this participant was removed from both the univariate and MVPA analyses of V3A.

As is common in other studies using MVPA^{13,58,59}, feature selection was also applied to select the top 120 voxels in each ROI that were the most active during the encoding period in trials without distractors. For combined regions, this number was multiplied by the number of regions in the combined ROI (for example, V1–V4 consists of four regions so a total of 480 voxels were selected). Decoding results were very similar for feature selected subset of voxels and the whole ROI. Unlike occipital regions, superior IPS is a relatively small region. Initially, we defined superior IPS with a threshold for activation at $P < 0.05$. However, for several participants, this produced an ROI that contained too few voxels for clear MVPA decoding, with an average of only 89 voxels across all participants, and a range (after feature selection) of 29–120 voxels. Therefore, we decreased the threshold for activation in the superior IPS localizer to $P < 0.1$ and performed the same subregion selection and feature selection detailed above. This produced a ROI with an average of 106 voxels across all participants, and a range of 45–120 voxels. This larger superior IPS ROI was used for the results presented here. We performed the same analyses using the more stringent ROI threshold (at $P < 0.05$) and saw a very similar pattern in this region. For individuals with less than 120 voxels in their superior IPS ROI, the entirety of the ROI was selected.

MVPA analysis. To assess whether a brain region was involved in VSTM storage (experiments 1 and 3), we used MVPA decoding methods to determine whether activity within each ROI reflected the orientation of the remembered grating for that trial. Similar to a previous study¹³, our methodology isolated memory specific activity by presenting both types of orientations in each trial and by using a cue based on stimulus order (that is, 1 or 2). This ensured that neither stimulus nor cue driven activity could contaminate the fMRI responses used to decode the orientation held in VSTM for any given trial. In addition to decoding the orientation of the remembered stimulus, in trials with distractors, we also performed separate analysis to decode the type of distractor presented in a given trial.

Decoding analysis was performed on the average delay period fMRI response (including time points 6–12 s or TRs 4–8) from each voxel in a given ROI, similar to a previous study¹³. These time points were selected because they accounted for hemodynamic response lag, but were uncontaminated by test stimulus presentations. This delay time period is one TR longer than that used previously¹³, but is still early enough to be uncontaminated by the presentation of the test stimulus (which occurred at 13 s). We saw no reason to exclude the additional data point, as the inclusion of this data point would increase power in our analysis. fMRI responses across all the voxels in each ROI were then normalized using z-score transformation to remove any effects related to overall amplitude differences between ROIs (decoding performance was similar for normalized and non-normalized data). The resulting fMRI response pattern from each ROI was then used in the decoding analysis.

Using the leave-one-run-out cross validation procedure, we trained a linear support vector machine (SVM) to either discriminate the orientation of the remembered grating or the type of distractors shown during the delay period. Analysis was performed in MATLAB using the CLOP toolbox (Challenge Learning Object Package, <http://clonet.com/CLOP/>). Decoding accuracy was expressed as the proportion of test patterns that were correctly classified, with chance level performance being 50%. Significance was assessed within an ROI using paired *t* tests, one tailed for comparisons to chance, two-tailed for comparisons between conditions.

Behavioral and neural representation similarity analysis. To construct neural representation similarity measures between orientations in experiment 4, we compared the decoding accuracy between all pairs of orientations. As in experiments 1 and 3, the decoding analysis was performed on the average delay period fMRI response (including time points 6–10 s or TRs 4–6) from each voxel in a given ROI. These time points were selected because they accounted for hemodynamic response lag but were far enough from encoding and test to be uncontaminated by stimulus presentations. Removing the earliest time point in this analysis produced similar results, suggesting that our results are driven by memory representations, not any lingering perceptual representations present during the encoding period which were unlikely given the masking procedure we used. fMRI responses across all the voxels within each ROI were then normalized using z-score transformation to remove any effects related to overall amplitude differences between ROIs. The resulting fMRI response pattern from each ROI was then used in the decoding analysis.

Using the leave-one-run-out cross validation procedure, we trained a linear support vector machine (SVM) to discriminate between each pair of orientations (for example, 10° and 40°, 10° and 70°, etc.). As before, the analysis was performed in MATLAB using the CLOP toolbox. Decoding accuracy was expressed as the proportion of test patterns that were correctly classified, with chance level performance being 50%. This produced a neural representation similarity matrix, showing how dissociable each orientation was from the others. The decoding accuracies for each orientation pair was calculated separately for each participant and then averaged across all participants to create the group-level neural representation similarity matrix²⁹.

The behavioral representation similarity matrix was created by comparing the reaction time to detect a change for each orientation paired with each of the other five orientations. We then averaged all trials for each orientation pair, regardless of which orientation was the target and which was the test stimulus. Reaction time was calculated separately for each participant and then averaged

across participants to form the group-level behavioral representation similarity matrix for each orientation pair.

We then directly correlated the behavioral and neural representation similarity matrices in each ROI. If the neural representation in a region reflects the storage of the item in VSTM, then we should see a strong correlation between the behavioral and neural measures of representation similarity²⁹. The significance of each correlation was evaluated using a permutation test in which the values within the behavioral and neural measures of representation similarity were randomly shuffled and then correlated. We ran the permutation test over 10,000 iterations to derive the mean and s.d. of the baseline correlation value distribution.

Univariate response amplitude analysis. In addition to the MVPA analysis, we also examined the univariate fMRI response amplitudes to both trials with and without distractors in each ROI in experiments 1–3. fMRI response amplitudes for each stimulus condition were measured in percent signal change, calculated by taking the difference in average signal intensity between each trial type and the fixation trials, then dividing this difference by that of the fixation trials and multiplying it by 100. Differences in encoding period and delay period activity were analyzed by using, in each individual participant, the maximum signal change during the encoding period and either the maximum (when activity increased during the delay) or minimum (when activity decreased during the delay) signal change during the delay period.

Statistics. We used a within-subject design in all the experiments included here, as such, all participants received all the test conditions. Consequently, randomization and blinding were not required. Simple *t*-tests and analysis of variance (ANOVA) were used to assess the difference between conditions at the group level. Following previously published studies that made similar measurements^{13,17}, data distribution was assumed to be normal, but this was not formally tested. Although no statistical methods were used to pre-determine the sample sizes for each experiment, our sample sizes are similar to those reported in previous publications^{13,17}.

A Supplementary Methods Checklist is available.

42. DeYoe, E.A. *et al.* Mapping striate and extrastriate visual areas in human cerebral cortex. *Proc. Natl. Acad. Sci. USA* **93**, 2382–2386 (1996).
43. Engel, S.A. *et al.* fMRI of human visual cortex. *Nature* **369**, 6481 (1994).
44. Sereno, M.I. *et al.* Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* **268**, 889–893 (1995).
45. Swisher, J.D., Halko, M.A., Merabet, L.B., McMains, S.A. & Somers, D.C. Visual topography of human intraparietal sulcus. *J. Neurosci.* **27**, 5326–5337 (2007).
46. Bressler, D.W. & Silver, M.A. Spatial attention improves reliability of fMRI retinotopic mapping signals in occipital and parietal cortex. *Neuroimage* **53**, 526–533 (2010).
47. Xu, Y. & Chun, M.M. Selecting and perceiving multiple visual objects. *Trends Cogn. Sci.* **13**, 167–174 (2009).
48. Straw, A.D. Vision egg: an open-source library for realtime visual stimulus generation. *Front. Neuroinform.* **2**, 4 (2008).
49. Brainard, D.H. The psychophysics toolbox. *Spat. Vis.* **10**, 433–436 (1997).
50. Dale, A.M., Fischl, B. & Sereno, M.I. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage* **9**, 179–194 (1999).
51. Fischl, B., Liu, A. & Dale, A.M. Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex. *IEEE Trans. Med. Imaging* **20**, 70–80 (2001).
52. Fischl, B., Sereno, M.I. & Dale, A.M. Cortical surface-based analysis. II. Inflation, flattening, and a surface-based coordinate system. *Neuroimage* **9**, 195–207 (1999).
53. Reuter, M., Rosas, H.D. & Fischl, B. Highly accurate inverse consistent registration: a robust approach. *Neuroimage* **53**, 1181–1196 (2010).
54. Fischl, B. *et al.* Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* **33**, 341–355 (2002).
55. Fischl, B. *et al.* Sequence-independent segmentation of magnetic resonance images. *Neuroimage* **23** (suppl. 1), S69–S84 (2004).
56. Cowan, N. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* **24**, 87–114, discussion 114–185 (2001).
57. Destrieux, C., Fischl, B., Dale, A. & Halgren, E. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage* **53**, 1–15 (2010).
58. Haynes, J.-D. & Rees, G. Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nat. Neurosci.* **8**, 686–691 (2005).
59. Kamitani, Y. & Tong, F. Decoding the visual and subjective contents of the human brain. *Nat. Neurosci.* **8**, 679–685 (2005).