ARTICLES
https://doi.org/10.1038/s41593-019-0428-x
nature neuroscience

Coexisting representations of sensory and mnemonic information in human visual cortex

Rosanne L. Rademaker[1,2*], Chaipat Chunharas[1,3] and John T. Serences[1,4,5*]

Traversing sensory environments requires keeping relevant information in mind while simultaneously processing new inputs. Visual information is kept in working memory via feature-selective responses in early visual cortex, but recent work has suggested that new sensory inputs obligatorily wipe out this information. Here we show region-wide multiplexing abilities in classic sensory areas, with population-level response patterns in early visual cortex representing the contents of working memory alongside new sensory inputs. In a second experiment, we show that when people get distracted, this leads to both disruptions of mnemonic information in early visual cortex and decrements in behavioral recall. Representations in the intraparietal sulcus reflect actively remembered information encoded in a transformed format, but not task-irrelevant sensory inputs. Together, these results suggest that early visual areas play a key role in supporting high-resolution working memory representations that can serve as a template for comparison with incoming sensory information.

When trying to attain behavioral goals, the ability to flexibly juggle thoughts is key. Visual working memory (VWM) provides the mental workspace to keep visual information online, allowing this information to guide visual search or to be recalled at a future moment in time. Neuroimaging studies have firmly established that the contents of VWM can be decoded from occipital cortex including primary visual area V1 (refs. [1–4]) and that the quality of this information predicts behavioral performance[5,6], suggesting that early sensory areas are involved in the representation of visual memories.

That said, previous studies have typically relied on a traditional delayed-match-to-sample (DMTS) task in which a sample memory stimulus is encoded and remembered across a blank delay interval before a test stimulus appears for comparison. However, in everyday perception, VWM maintenance needs to be robust to the continuous influx of new visual inputs that come with each exploratory saccade or change in the environment. Thus, a delay period devoid of other visual inputs is divorced from typical visual experience. Based on this mismatch between experimental and real-world scenarios, some have argued that recruiting early sensory areas to store relevant VWM information would be counterproductive in everyday life, as new sensory inputs would destructively interfere with concurrent mnemonic representations[6–8].

Based on this logic, one recent study[6] employed a DMTS task with task-irrelevant pictures of faces and gazebos sometimes presented during the delay period. The authors used functional magnetic resonance imaging (fMRI) and found that activation patterns in early visual cortex represented the contents of VWM during blank delays, but not when delays were filled with predictable distractors (that is, the task-irrelevant pictures). The authors concluded that representations initially encoded in early visual cortex were recoded in a more durable format in parietal cortex to insulate mnemonic representations from interference induced by new sensory input. Furthermore, the authors argued that the disengagement of primary sensory regions was strategic, as it occurred only when participants expected the task-irrelevant pictures. This study challenged the importance of early visual cortex during VWM, attributing previous findings of sensory recruitment to overly artificial tasks (see ref. [9] for potential caveats of this work).

The framework proposed by this previous study[6] implies a fundamental limitation of cortical information processing: a sensory area, such as the primary visual cortex, cannot represent both top-down biasing signals associated with internal cognitive operations such as VWM and bottom-up sensory inputs evoked by newly encountered stimuli in the environment. However, this strong stance is questionable for at least two reasons. First, from a functional point of view, success on a DMTS task relies on comparing an internally stored representation to a new sensory input. For this, a 'local comparison circuit', able to jointly represent remembered and perceived items in the same local circuit, could be ideal. Second, separable bottom-up and top-down inputs could theoretically support the coexistence of multiple simultaneous representations, a concept we term 'region-wide multiplexing' (after the more common usage of 'multiplexing' to refer to flexible coding in single neurons). Bottom-up input from the lateral geniculate nucleus primarily projects to layer 4 of primary visual cortex, whereas top-down input arrives primarily in superficial layers and layer 5 (refs. [10,11]). When information from layer 4 is conveyed to the superficial layers, different populations of neurons might be recruited to keep bottom-up and top-down inputs anatomically segregated[12]. In addition, the format of the codes might differ, with bottom-up signals driving changes in spike rate and top-down signals modulating membrane potentials[7]. Such a system could promote match detection via response gain when memory and sensory information are aligned[13].

Results

In Experiment 1, we evaluated the ability of early visual areas to act as a multiplexing comparison circuit during VWM. Participants performed a working memory task where they remembered randomly oriented visual gratings while looking at either a blank screen

[1]Department of Psychology, University of California, San Diego, La Jolla, CA, USA. [2]Donders Institute for Brain, Cognition and Behavior, Radboud University, Nijmegen, the Netherlands. [3]Department of Medicine, King Chulalongkorn Memorial Hospital, Chulalongkorn University, Bangkok, Thailand. [4]Neurosciences Graduate Program, University of California, San Diego, La Jolla, CA, USA. [5]Kavli Institute for Brain and Mind, University of California, San Diego, La Jolla, CA, USA. *e-mail: rosanne.rademaker@gmail.com; jserences@ucsd.edu

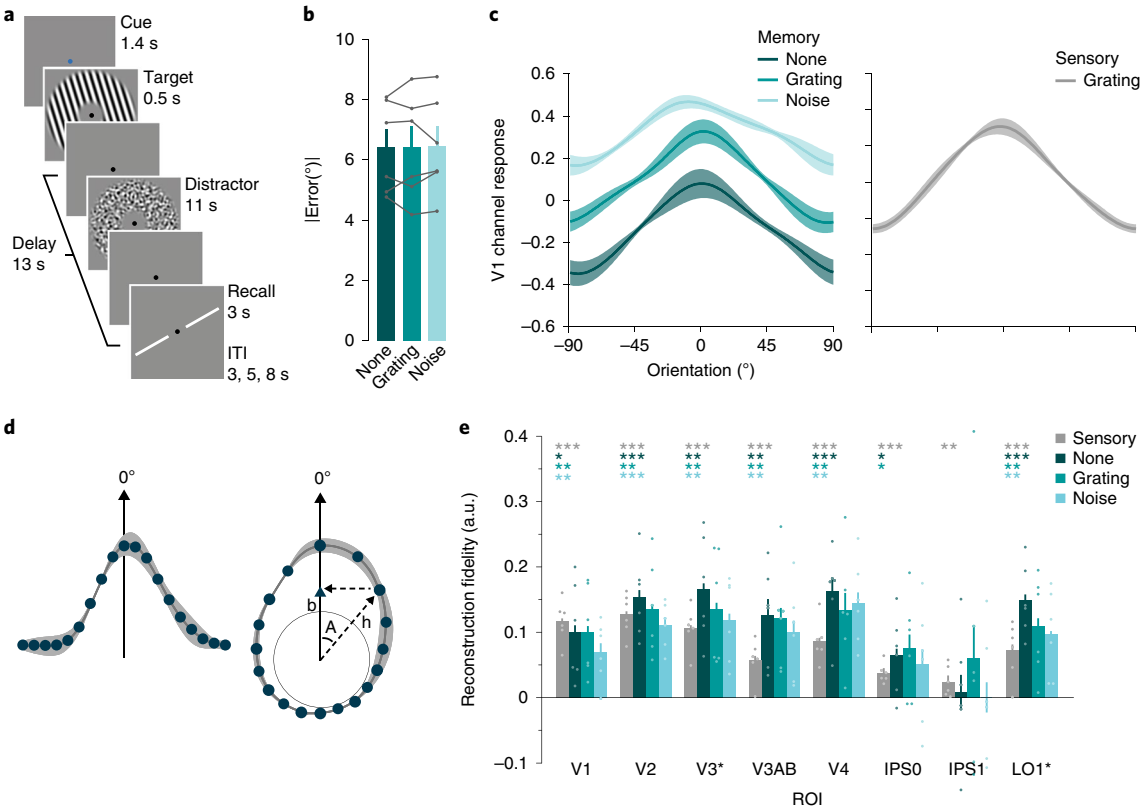

Fig. 1 | Experiment 1 paradigm and results. a, After a valid cue about the distractor condition (here, the blue fixation dot cued a noise distractor), a 0.5-s target orientation was remembered for 13 s. During this delay, participants viewed a gray screen or an 11-s contrast-reversing distractor. Distractors could be a Fourier-filtered noise stimulus (depicted) or an oriented grating (its orientation was pseudo-randomly selected on every trial). After the delay, participants had 3 s to rotate a recall probe to match the remembered orientation. Participants then continued to the next trial after an inter-trial interval (ITI) of 3, 5 or 8 s. **b**, There were no differences in behavioral error between the three distractor conditions, as indicated by a non-parametric one-way repeated-measures within-subject ANOVA ($F_{(2,10)} = 0.044$; $P = 0.943$). Gray lines indicate individual subjects. **c**, Model-based reconstructions of the remembered orientation during the three different distractor conditions (left) and of the physically present orientation on trials with a grating distractor (right). Reconstructions were based on the average activation patterns 5.6–13.6 s after target onset. **d**, The degree to which memory and sensory stimuli were represented during the delay was quantified by projecting the channel response at each degree onto a vector centered on the true orientation (that is, zero), and taking the mean of all these projected vectors. Left: a cartoon reconstruction is defined by 18 points (note, in reality there were 180°). Right: this cartoon reconstruction is wrapped onto a circle. We show for one point how the channel response (h) is projected onto the true orientation (remembered or sensed), resulting in vector b. Knowing the angle (A) between the true orientation and the orientation at this particular point, we solve for b using trigonometric ratios for right triangles (that is, $\cos A = b/h$). The mean of all projected vectors (all b) indexes the amount of information at the true orientation and is our metric for reconstruction fidelity (in arbitrary units, a.u.). **e**, Reconstruction fidelity for remembered (shades of teal) and sensed distractor (gray) orientations is significantly above chance in almost all ROIs (based on one-sided randomization tests comparing fidelity in each condition and ROI to zero; see Methods). Black asterisks next to ROI names (under the x axis) indicate significant differences in memory fidelity between the three distractor conditions in that ROI, as determined by non-parametric one-way repeated-measures within-subjects ANOVA analyses performed separately for each ROI (see Methods; for exact P values and post hoc tests, see Supplementary Tables 1 and 2). $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ (uncorrected for multiple comparisons). Dots indicate individual subject fidelities in each condition and ROI. For **b**, **c** and **e**, error bars/areas represent ±1 within-subject s.e.m. around the average across $n = 6$ independent subjects.

or a sequence of contrast-reversing visual distractors (Fig. 1a). Each trial started with a 100% valid cue (1.4 s) indicating the distractor condition during the subsequent delay. Next, a target orientation was shown for 0.5 s and remembered throughout a 13-s delay. To ensure a relatively uniform sampling of orientation space, the target orientation was pseudo-randomly drawn from one of six orientation bins (each bin contained 30 orientations, in integer increments), with an equal number of draws from each bin. During the middle portion of the delay, participants viewed either a gray screen or an 11-s contrast-reversing distractor. Distractors were either a Fourier-filtered white noise stimulus (an example is shown in Fig. 1a—a novel noise structure was generated on every trial) or an oriented grating with pseudo-random angular offset relative to the memory target orientation (its orientation was similarly drawn from one of six bins, counterbalanced with respect to the target orientation bin; see Supplementary Fig. 1a). After the delay, participants had 3 s to rotate a recall probe to match the remembered orientation as precisely as possible before continuing to the next trial. Although the presence or absence of distractors was fully predictable, distractors were irrelevant to the task and had no observable impact on behavior during fMRI scanning (Fig. 1b and Supplementary Fig. 2a). As expected, distractors effectively drove a sustained and highly robust increase in the overall univariate response amplitude in V1 and other visual areas (Supplementary Fig. 3a). Note that the distractors, when presented, function as visual masks. This was a deliberate choice, optimizing the design to be most favorable to the no-distractor condition (that is, a blank screen without any visual interference during the delay).

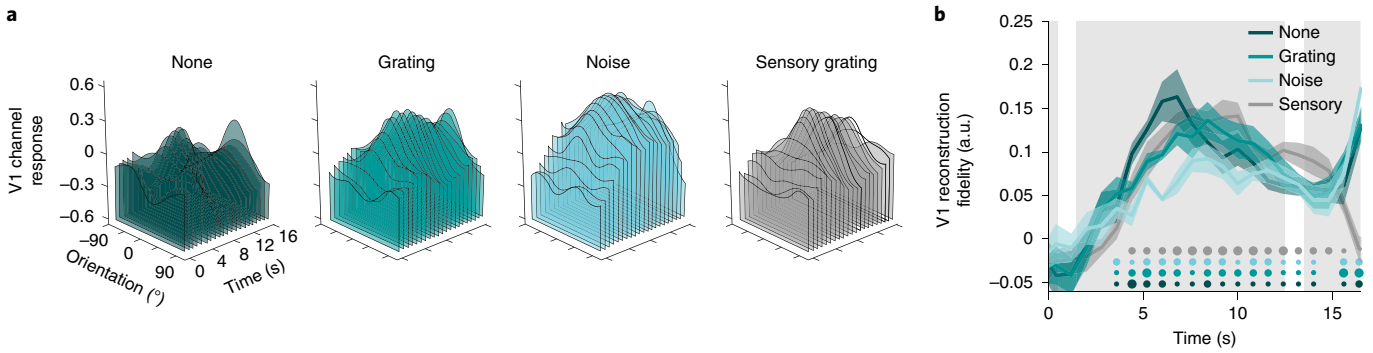

Fig. 2 | Model-based reconstructions of remembered orientations and sensed distractor orientations over time in V1. a, The time axis starts at 0, which is trial onset, and each slice shows the mean reconstruction across participants at each 800 ms TR (for a total of 21 TRs). Reconstructions for the remembered orientation are shown in the three left-most panels (shades of teal), and sensed distractor orientation reconstructions are shown in the right-most panel (gray). **b**, The fidelity of timepoint-by-timepoint reconstructions in V1 (quantification of **a**), with time 0 representing the target onset. The three gray background panels represent the target, distractor, and recall epochs of the working memory trial. Small, medium, and large dots at the bottom indicate significance at each time point at $P < 0.05$, $P < 0.01$ and $P < 0.001$, respectively (based on one-sided randomization tests comparing fidelity in each condition and at each timepoint to zero, uncorrected for multiple comparisons; see Methods). Shaded error areas represent ± 1 within-subject s.e.m. around the average across $n = 6$ independent subjects.

Next, using a multivariate inverted encoding model[14,15] (IEM, see Supplementary Fig. 4) trained on independent data, we generated model-based reconstructions of the remembered orientation from delay-period activity patterns in primary visual cortex (Fig. 1c, left panel), and all other retinotopic areas that we mapped along the visual hierarchy (Supplementary Fig. 5a), irrespective of whether a distractor was present during the delay. The baseline offset observed between distractor conditions (vertical shift up or down the *y* axis in Fig. 1c) largely reflects the univariate effect of distractor presence during the delay interval, with higher baselines during trials with distractors (see Supplementary Fig. 3a). As a measure of reconstruction fidelity, we projected the channel response at each degree in orientation space onto the remembered orientation and then took the mean of these projected vectors (Fig. 1d). Fidelity was significantly above zero, indicating that there was information about the remembered orientation during all distractor conditions (Fig. 1e, teal bars) and in every retinotopic region of interest (ROI), except areas in the intraparietal sulcus (IPS). Note that the independent data used to train the IEM were collected while participants directly viewed orientation stimuli. Thus, generalization from sensory evoked response patterns to memory-related response patterns during the delay epoch implies that mnemonic information was represented in a sensory-like format.

We were also able to reconstruct the orientation of the distractor that was physically present on the screen during grating-distractor trials (Fig. 1c, right panel and Fig. 1e, gray bars), using the exact same delay-period data. This demonstrates that, contrary to simple feedforward models that posit V1 as a passive filter, early visual cortex can represent incoming sensory information alongside mnemonic information that is no longer tethered to external inputs. While the fidelity of remembered and sensed orientations was roughly equivalent in V1–V2 (Fig. 1e, compare mid-teal bars and gray bars), the fidelity of the sensed distractor grating dropped against the fidelity of the remembered grating when ascending the visual hierarchy to V3–V4 (interaction: $F_{(4,20)} = 5.67$, $P = 0.002$; note that this analysis does not include IPS and LO1, as their relative hierarchical relationships are less clear). This finding captures the top-down nature of mnemonic signals, as top-down signals are thought to have more traction than bottom-up sensory inputs in higher-level regions.

Notably, timepoint-by-timepoint reconstructions revealed that remembered and perceived representations evolved together over time in primary visual cortex (Fig. 2), indicating that these representations coexisted throughout most of the delay period. This was also true for other early retinotopic areas along the visual hierarchy, but not for later retinotopic IPS areas (Supplementary Fig. 6a). Note that claims about the coexistence of information are limited by the temporal resolution of our measurement (that is, a repetition time (TR) of 800 ms). The notion of a comparison circuit at the level of early sensory cortex was further supported by a boost in representational quality when target and distractor orientations were similar, compared to dissimilar (Supplementary Figs. 7 and 8b). This finding is mirrored by behavioral demonstrations showing higher fidelity memory recall for similar targets and distractors[16,17] (Supplementary Fig. 9b).

In a second experiment we evaluated the impact of more naturalistic distractors, namely, face and gazebo pictures (after ref. [6]). Moreover, instead of contrast-reversing our visual distractors (as in Experiment 1), we flickered distractors on (250 ms) and off (250 ms) for 11 s during the delay, and randomized their spatial phase on each 'on' cycle, to maximize the unpredictability of contrast changes at every pixel. The task structure (Fig. 3a) was similar to that of Experiment 1, with participants again remembering a pseudo-randomly oriented grating while ignoring other inputs during the delay. A 100% predictable cue (1.4 s) was presented before the to-be-remembered memory target (0.5 s), with the cue indicating one of three possible events during the 12 s delay: no distractor (that is, blank screen), an 11 s grating distractor (with pseudo-random angular offset relative to the target orientation, see Supplementary Fig. 1b) or 11 s of picture distractors (example shown in Fig. 3a). Participants had 4 s to report the remembered orientation as precisely as possible by rotating a dial. Distractors drove robust univariate responses in all of our ROIs (Supplementary Fig. 3b).

Notably, the presence of distractors in Experiment 2 negatively affected behavioral performance during fMRI scanning (Fig. 3b and Supplementary Fig. 2b). This drop in behavioral performance was accompanied by qualitatively poorer memory reconstructions when distractors were presented during the delay (Fig. 3c, left panel and Supplementary Fig. 5b). Memory fidelity in V1–V4 and LO1 was reduced when grating and picture distractors were shown during the delay, compared to fidelity without distractors (Fig. 3d). Alongside these (reduced) memory representations, the directly sensed distractor orientation was represented in a robust manner (Fig. 3c, right panel and Fig. 3d, gray bars), as were the directly sensed picture distractors (Supplementary Fig. 10). As in Experiment 1, the IEM was trained on independent data, collected while participants directly viewed oriented gratings. Generalization to data from the

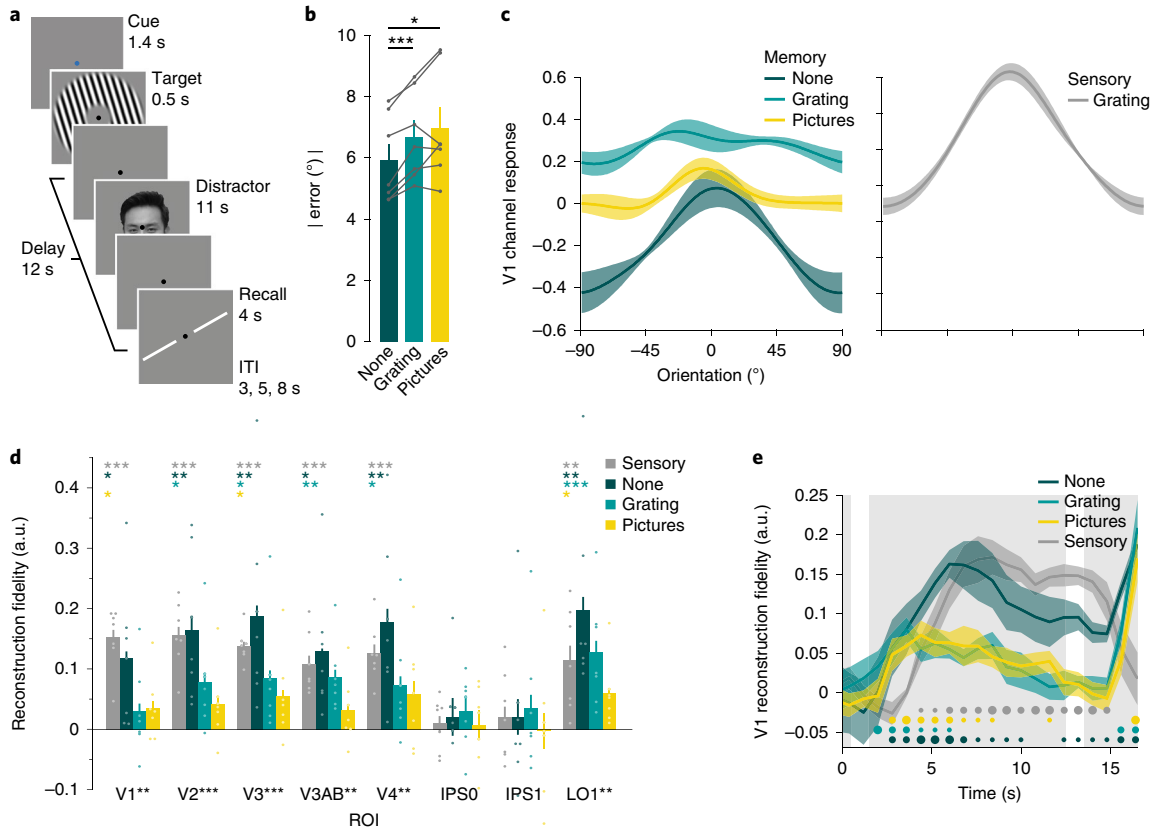

Fig. 3 | Experiment 2 paradigm and results. a, Irrelevant distractors were cued with 100% validity by a change in fixation color (here blue indicated that picture distractors would be shown during the delay) before a 500 ms target presentation. Participants remembered the target orientation for 12 s, while they viewed either a gray screen or an 11-s on–off flickering distractor (a pseudo-randomly oriented grating, or pictures of faces or gazebos). After the memory delay, participants rotated a dial to match the remembered orientation. Photo used with permission (S. Itthipuripat). **b**, Distractor presence negatively affected behavioral performance, as indicated by a non-parametric one-way repeated-measures within-subjects ANOVA ($F_{(2,12)}=10.154$; $P<0.001$). Errors were smaller when no distractor was shown during the delay, compared to when distractor gratings ($t_{(6)}=6.272$; $P<0.001$) or pictures ($t_{(6)}=3.375$; $P=0.018$) were shown. Performance did not differ between grating and picture distractors ($t_{(6)}=1.184$; $P=0.184$). Post hoc tests were non-parametric uncorrected paired-sample *t*-tests. Gray lines indicate individual subjects. **c**, Model-based reconstructions of the remembered orientation during the three different distractor conditions (left) and of the sensed distractor orientation on trials with a grating distractor (right). Reconstructions were generated with an IEM trained on independent localizer data and based on the average activation patterns 5.6–13.6 s after target onset. **d**, Reconstruction fidelity for remembered orientations without distraction (dark teal) and for sensed distractor orientations (gray) is significantly above zero in all ROIs except IPS0 and IPS1 (based on one-sided randomization tests in each condition and ROI; see Methods). However, reconstruction fidelity is less robust when a distractor was presented throughout the delay (mid-teal and yellow for grating and picture distractors, respectively). Black asterisks next to ROI names indicate significant differences in memory fidelity during the three distractor conditions in that ROI, as determined by non-parametric one-way repeated-measures within-subjects ANOVA analyses performed separately for each ROI (see Methods; for exact *P* values and post hoc tests, see Supplementary Tables 3 and 4). Dots indicate individual subject fidelities in each condition and ROI. **e**, The fidelity of timepoint-by-timepoint reconstructions in V1. Time 0 represents target onset, and the three gray panels represent the target, distractor, and recall epochs of the working memory trial. Small, medium or large dots at the bottom of **e** indicate significance levels of $^{*}P<0.05$, $^{**}P<0.01$ or $^{***}P<0.001$, respectively (uncorrected). For **b**–**e**, error bars and shaded areas represent ± 1 within-subject s.e.m. around the average across $n=7$ independent subjects.

memory delay implies that a sensory-like code is used to represent mnemonic information, and that this representation is less robust when people are distracted by visual inputs during the delay. Note that in IPS0 and IPS1 there was no evidence of mnemonic information, nor did these regions represent the sensed distractor grating, indicating that these areas do not represent information in a manner that is generalizable from directly viewed sensory inputs.

A direct comparison between the remembered and sensed orientations on trials with a grating distractor (compare mid-teal bars and gray bars in Fig. 3d) again revealed a relative increase in mnemonic compared to sensed information when ascending the visual hierarchy (interaction: $F_{(4,24)}=7.418$, $P=0.001$)—a hallmark of top-down processing[18–21]. Timepoint-by-timepoint analyses (Fig. 3e) showed sustained memory representations throughout the delay when there was no distractor present, while memory representations were less sustained in the presence of visual distraction. Note that this did not hold true for IPS, where there was no evidence for representations of either the remembered orientation or the sensed distractor orientation, sustained or otherwise (Supplementary Fig. 6b).

In the analyses presented thus far, we trained the IEM solely on data from independent sensory localizers, demonstrating that visual areas up to IPS encode remembered features in a sensory-like format. Here, 'sensory-like' refers to a format akin to that of a stimulus-driven sensory response. An independent sensory localizer makes no demands on memory, and gives rise to information in a stimulus-driven format. However, not all mnemonic signals are necessarily stored in this format, and might also be stored in a format that is somehow transformed. For example, pixel-by-pixel representations

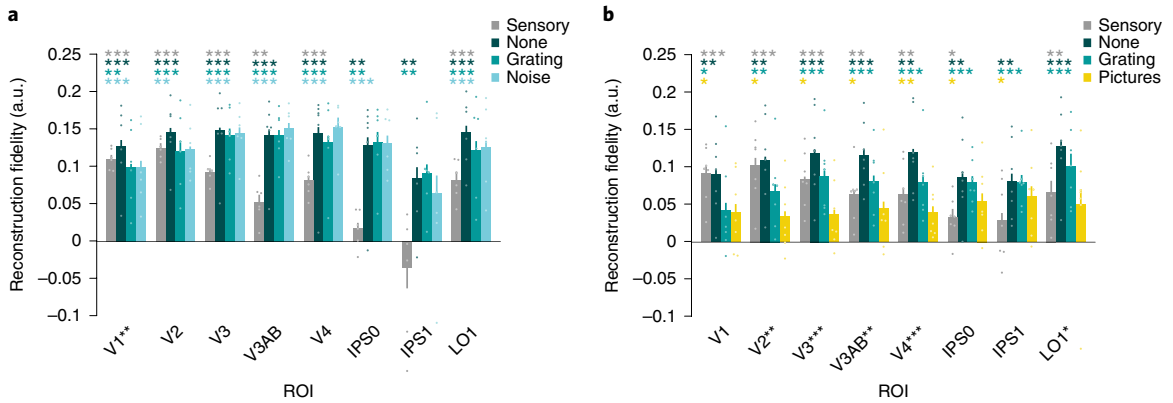

Fig. 4 | Reconstruction fidelity when training and testing an IEM on data from the memory delay. a,**b**, Experiment 1 (**a**) and Experiment 2 (**b**). There are robust memory representations throughout the visual hierarchy, including retinotopic IPS. This implies that the representational format in IPS is not in a stimulus-driven format. The proposed transformed nature of the IPS code is also supported by the lack of information about the directly sensed grating distractor (gray bars). As before, differences in memory fidelity between the three distractor conditions (black asterisks next to ROI names) were virtually absent in Experiment 1 (**a**; for exact *P* values and post hoc tests see Supplementary Tables 5 and 6), while in Experiment 2 the presence of distractors was accompanied by a drop in memory fidelity in many ROIs (**b**; for exact *P* values and post hoc tests see Supplementary Tables 7 and 8). Note, however, that mnemonic representations in IPS were unaffected by visual distraction (see also Supplementary Fig. 13). $^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$. Dots indicate individual subject fidelities in each condition and ROI. Error bars represent ±1 within-subject s.e.m. (for $n=6$ and $n=7$ independent subjects in **a** and **b**, respectively). Statistical testing was identical to that in Figs. 1e and 3d. When ascending the visual hierarchy from V1 to V4, a weakening sensory representation paired with a strengthening mnemonic representation illustrates the top-down nature of VWM (compare gray and mid-teal bars). This signature interaction was present in both Experiment 1 ($F_{(4,20)}=13.6$, $P<0.001$) and Experiment 2 ($F_{(4,24)}=7.769$, $P<0.001$).

in early visual cortex might undergo some dimensionality reduction in upstream cortical sites[22,23]. To look for a mnemonic code that is not necessarily sensory-like, we trained the IEM on data from the memory delay via a leave-one-out procedure (see Methods). In both Experiments 1 and 2 we saw robust VWM representations, despite visual distraction, in all retinotopic ROIs, including IPS0 and IPS1 (Fig. 4 and Supplementary Figs. 11 and 12), and including more anterior and non-visually responsive IPS ROIs (Supplementary Fig. 13). Despite the ubiquity of mnemonic information in IPS revealed by this analysis, there was still no apparent information about the directly sensed distractor grating (gray bars) in IPS (Fig. 4 and Supplementary Fig. 13). Taken together, this implies that IPS does represent mnemonic information, and that it uses a code that is transformed away from the stimulus-driven driven response. Furthermore, representations in IPS were impervious to visual distraction, with equivalent memory fidelity during all distractor conditions even in Experiment 2 (Supplementary Fig. 13), despite differences at the behavioral level (Fig. 3b). Thus, representations in V1–V4 and LO1 (lower fidelity during visual distraction), but not IPS (stable fidelity), mirrored how well people did on a VWM task.

Note that this analysis does not necessarily speak to the representational format in early visual areas V1–V4 or LO1. While generalization from independent sensory data demonstrated a sensory-like mnemonic format, generalization within the memory delay does not by definition indicate a non-stimulus-driven format. After all, training and testing on a sensory-like mnemonic code would yield robust reconstructions as long as there was information present. Thus, we can only ascertain the presence of non-stimulus-driven transformed codes in IPS, as uncovering mnemonic information in IPS is only possible after training the IEM on memory (and not sensory) data.

Finally, none of the findings reported here depend on our choice of analysis approach (IEM model), and conventional decoding analyses yield similar patterns of results (Fig. 5).

Discussion

Visual information held in mind to attain behavioral goals should withstand interference from ongoing sensory inputs. In Experiment 1, we demonstrated that recall of an orientation was unimpeded by irrelevant visual distractors, and that the fidelity of mnemonic representations in visual cortex was similar with and without distractors. By contrast, participants in Experiment 2 did get distracted, showing impairments at both the behavioral and neural level. Participants viewed noise distractors in Experiment 1, and picture distractors (faces and gazebos) in Experiment 2. Different distractor types might result in different degrees of distractibility, which could explain the discrepancy between the two experimental outcomes. However, grating distractors were shown during both experiments, and even for this shared condition there was a drop in the fidelity of behavioral and brain responses in Experiment 2. This was also true for the three participants who completed both experiments and had not been affected by distractors in Experiment 1. Instead of distractor type, a likely variable causing the differences in distractibility is the 'intensity' of the distractors, namely, whether they were contrast-reversing (Experiment 1) or flickering on and off (Experiment 2). Integrated over a contrast-reversal cycle, the contrast at every pixel is always mean gray. In comparison, flickering between a grating with a random phase ('on') and a mean gray screen ('off') results in contrast fluctuations from cycle to cycle, and thus a stronger temporal gradient of change at every pixel.

A previous paper reported one experiment implying that task-irrelevant visual distraction wiped out mnemonic representations in V1 at no behavioral cost[6,9]. Our results from Experiment 2 are reminiscent of this finding, where distractor presence caused a marked drop in mnemonic information in early retinotopic areas. However, information did not generally dissipate altogether (Figs. 3d, 4b and 5 and Supplementary Figs. 5b, 6b, 11b and 12b). This reduction of information was mirrored by a clear decline in behavioral performance (Fig. 3b). Previous failures to uncover behavioral effects when only a single feature is remembered can be readily explained by the need for adequate statistical power in paradigms where memory fidelity tends to be high[24]. Instead of a traditional DMTS task with only two answer alternatives, the recall procedure we used here allowed a much more sensitive detection of numerically small behavioral effect. Thus, differences in distractor intensity, as well as a behavioral paradigm that lacks sensitivity, can

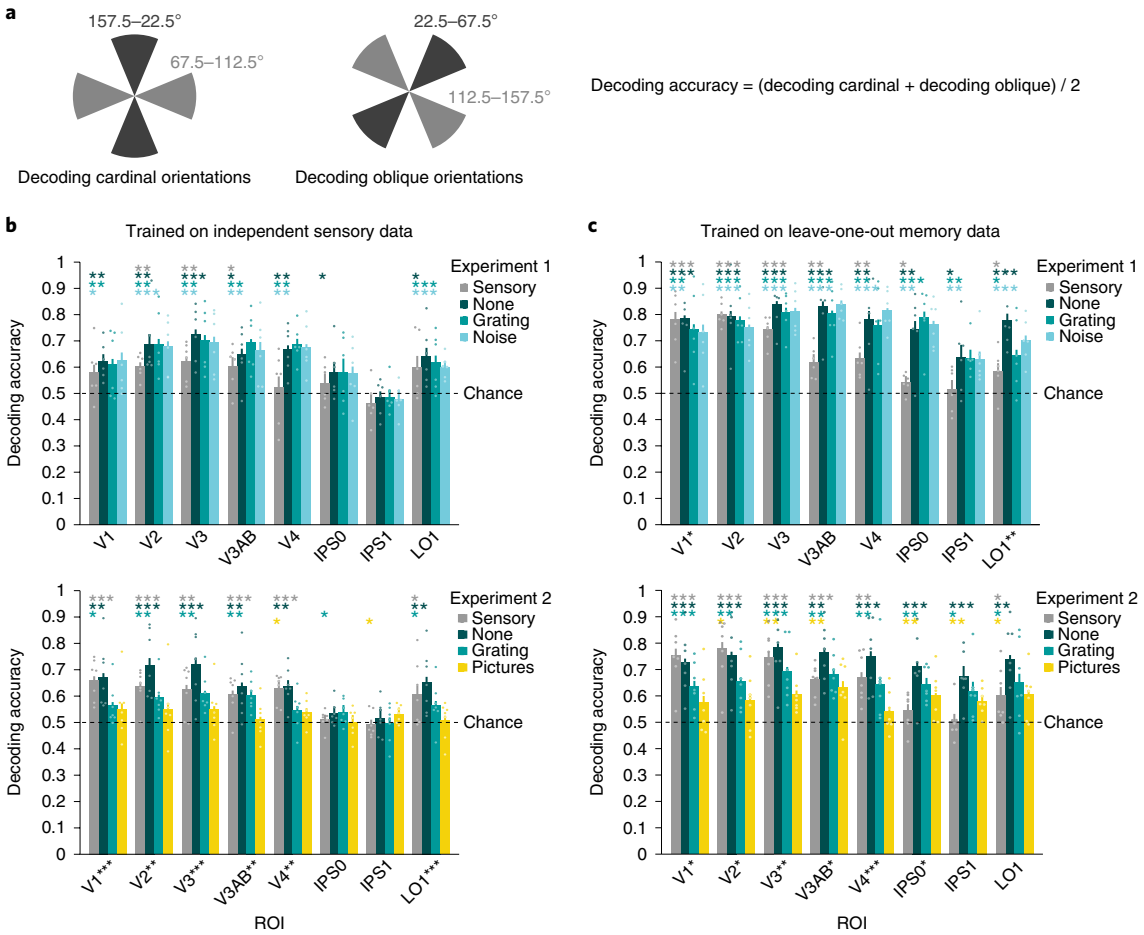

Fig. 5 | Decoding analyses yield highly comparable results to the IEM analyses. a, In Experiments 1 and 2 we used random orientations (1–180°), while relevant previous work has used orthogonal orientations[1,6]. To closely mimic the two-way classification performed in previous work, we divided our random orientations into four bins and performed two two-way classifications. The first classification determined whether orientations were around vertical (between 157.5 and 22.5°) or horizontal (between 67.5 and 112.5°), shown schematically in the left diagram. The second classification determined whether orientations were around one or the other obliques (that is, between 22.5 and 67.5° or between 112.5 and 157.5°), shown schematically in the right diagram. Decoding performance was averaged across these two-way classifications to yield an overall classification accuracy for each ROI. For all decoding analyses we ensured balanced training sets. **b**, We trained the SVM on independent data from the visual mapping tasks. Results mirrored those from the IEM analyses. In Experiment 1 (top) we found above chance decoding in V1–V4 and LO1, but not IPS0 and IPS1. There were no differences between the three distractor conditions in any of the ROIs (all $F_{(5,10)} < 1.024$, all $P > 0.429$). Also, in Experiment 2 (bottom) there was little above-chance decoding in IPS regions. In V1–V4 and LO1, memory decoding in Experiment 2 differed between the three distractor conditions (all $F_{(5,10)} > 10.419$, all $P < 0.004$) and was generally better when no visual distraction was presented during the delay, compared to delays with a grating or a picture distractor. In both Experiments 1 and 2, the grating-distractor condition revealed an interaction between remembered and sensed representations (compare mid-teal and gray bars), considered a signature of top-down processing ($F_{(4,20)} = 2.469$, $P = 0.046$ and $F_{(4,24)} = 3.198$, $P = 0.024$, respectively). **c**, We also trained the SVM on data from the memory delay via a leave-one-out cross-validation procedure. This led to robust decoding of mnemonic information in IPS0 and IPS1 for both Experiments 1 (top) and 2 (bottom), indicating a non-stimulus-driven mnemonic code in these areas. Lack of information about the ignored sensory distractor orientation (gray bars) further corroborates that IPS uses non-stimulus-driven codes to represent task-relevant information. In Experiment 1 (top) the three distractor conditions differed in V1 and LO1 ($F_{(2,10)} = 3.517$, $P = 0.045$ and $F_{(2,10)} = 12.723$, $P = 0.003$, respectively) but not in any other ROIs (all $F_{(2,10)} < 1.062$, all $P > 0.386$). In Experiment 2 (bottom) the three distractor conditions differed in almost all ROIs (V1–IPS0, all $F_{(2,12)} > 5.399$, all $P < 0.022$;). Again, both Experiments 1 and 2 revealed an interaction between remembered and sensed representations (compare mid-teal and gray bars) in the grating-distractor condition ($F_{(4,20)} = 11.499$, $P < 0.001$ and $F_{(4,24)} = 3.331$, $P = 0.029$, respectively). For both **b** and **c**, statistical testing was identical to that in Figs. 1e, 3d and 4 except that randomization tests were against chance (0.5) instead of zero (see also Methods). $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. Dots indicate individual subject decoding in each condition and ROI. Error bars represent ±1 within-subject s.e.m. (for $n = 6$ and $n = 7$ independent subjects in Experiments 1 and 2, respectively).

account for the biggest discrepancies between previous work[6] and our current findings.

The coexistence of bottom-up information about current sensory inputs, combined with feature-selective top-down inputs that carry information about remembered items, could provide a powerful local mechanism for comparing memory contents to the sensory environment[12,25,26]. To support this functionality, sensory and memory information could be multiplexed by different populations of neurons[12]. For example, if neurons tuned to the features of a memory target were selectively activated during the delay, the output of a local comparison circuit would be relatively high when a matching test stimulus was encountered that selectively excited similarly tuned neurons. On the other hand, a mismatch test stimulus would drive a set of differently tuned neurons, which may lead

to a lower overall response due to inhibitory competition with the already active neurons tuned to the sample feature[13,27]. The same logic also applies if the top-down modulations supporting VWM do not lead to sustained patterns of spiking in sensory cortices, and instead only influence sub-threshold potentials[7,28]. Differences in the local comparison circuit output would still be expected due to interactions between the top-down feature-selective bias and the sensory response evoked by the test stimulus. Moreover, content-specific patterns of sub-threshold membrane potentials could persevere through bouts of local spiking driven by sensory inputs during a delay, and thus protect mnemonic contents from distraction[13]. Because fMRI measures an aggregate of signals (that is, spikes, local field potentials and so on), here we can only speculate about the exact nature of this comparison circuit and cannot draw inferences about single neurons, the likely scale of anatomical separation, or precise temporal integration of the comparison circuit.

Here, we find distractor-resistant mnemonic representations throughout the delay in early visual cortex (Experiment 1), whereas classic single-neuron physiology has generally found that mnemonic representations in later stages of visual cortex are disrupted by visual transients[12,29,30]. For example, when monkeys viewed a target image and subsequently looked for matches in a series of test images, neuronal responses in inferior temporal cortex signaled an active memory trace via enhanced firing for matches[31]. However, this memory signal did not bridge the intervening delays between test images. By contrast, delay activity in prefrontal cortex survived intervening test stimuli and was maintained during each delay[32,33]. Note how this task, unlike ours, required the animal to perform a matching operation on each intervening stimulus, and thus each 'distractor' was actually a behaviorally relevant image that required attentive processing. In a set of studies more directly comparable to our experiments, monkeys had to mentally trace a curved line that was no longer in view, which led to sustained delay-period spiking in V1. Spiking was briefly interrupted by an irrelevant mask, but reinstated soon thereafter[11]. Thus, the status of a distractor as relevant or irrelevant might play an important role in how memories are maintained. Also, different memory contents (that is, highly familiar categorical objects versus fine-grained line orientations) might require different levels of representational precision.

Prefrontal and parietal cortices play a central role in maintaining distractor-resistant memory representations[32–34], and feedback from these regions likely supports the persistent mnemonic representations in early visual cortex found here. Early retinotopic representations were sensory-like in nature, as evidenced by the generalization from independent sensory data (used to train our multivariate models) to delay epoch data. A sensory-like format would indeed be well suited for a local comparison circuit, readily able to contrast mnemonic information and ongoing sensory inputs. By contrast, sensory information did not generalize to the memory delay in IPS. Instead, only training and testing on data obtained during the memory delay itself revealed information in IPS. This indicates a code that is transformed away from a purely stimulus-driven format. The notion of such a non-stimulus-driven code in IPS is further supported by the absence of information about the directly sensed grating distractor. Maintaining multiple replicas of a remembered sensory stimulus at all cortical levels would be computationally expensive and inefficient. Instead, high-resolution pixel-by-pixel representations might be condensed into stable and low dimensional representations in higher cortical regions[22]. Accordingly, these stable and potentially compressed representations might not support high-fidelity mnemonic information. Indeed, behavioral performance in Experiment 2 was impaired in the same conditions where representations in early visual areas were disrupted, while mnemonic representations in IPS remained intact. Of course, our failure to find sensory-like representations in IPS does not mean they do not exist there. For one, IPS does not have orientation columns in the same way that early retinotopic regions do, which could impede our ability to pick up on macroscopic information at the voxel level. Moreover, participants did not attend the orientations of the sensory stimuli used to train our model (instead, they performed an orthogonal task). A confluence of both perception and attention might be required to get reliable sensory responses from IPS[35,36]. Even though we cannot exclude the possibility of a mnemonic code in IPS that reflects stimulus-driven responses, our data do demonstrate that a transformed non-stimulus-driven code exists in IPS.

Previous studies have shown that there are interactions between remembered and seen stimuli, such as interference by[16,17,37–41] and attraction toward[16,17,42–47] irrelevant distractors (see also Supplementary Fig. 9). One recent fMRI study[46] looked at visual cortex representations of a remembered orientation in the delay period before and after a brief (0.5 s) irrelevant grating. The irrelevant grating always differed 40–50° from the target orientation. In the delay before the irrelevant grating, the remembered orientation could be recovered from early visual areas V1–V3 combined. In the delay after the irrelevant grating, the recovered orientation was shifted in the direction of the distractor, dovetailing with known behavioral attraction biases toward irrelevant orientations[16,17,47] (see also Supplementary Fig. 9a). However, this previous study only looked at memory representations before and after distraction, so nothing can be said about the joint representation of information. Furthermore, the target and distractor orientations were yoked together, so the representations associated with the target and the distractor could not be independently assessed. In the present work, we were able to detect biases toward irrelevant gratings during distraction (Supplementary Figs. 7 and 8a) while using randomized target–distractor differences (Supplementary Fig. 1). However, when people remember an orientation while a grating distractor is presented, both the target and distractor orientations contribute to the measured brain response, which could explain such biases. At the single-trial level, the relative contributions of mnemonic and perceived signals to an IEM reconstruction can be hard to untangle. Nevertheless, the uncorrelated nature of target and distractor orientations allowed us to assess memory representations in the presence of an orientation distractor across many trials. The finding that the average fidelity of mnemonic representations in Experiment 1 were unaltered by concurrent sensory inputs can therefore not be an artifact of the distractor orientation. This is further supported by comparably durable memory representations in the presence of grating and noise distractors alike—the latter having no discernible orientation information.

What if, instead of coexisting mnemonic and sensory representations, people were exclusively representing either the target or the distractor orientation on some fraction of trials or time points within a trial? This alternative account is unlikely for several reasons. First, switching between representations would impose a drop in the representation of the memory target. No such drop from the no-distractor condition to the grating-distractor condition was observed in Experiment 1. Second, the 11 s continuous presentation of distractors necessarily activates V1. Thus, while V1 is representing ongoing sensory inputs, mnemonic information can still be recovered at every TR throughout the delay.

Neuroimaging studies on working memory routinely use a retro-cue paradigm where two stimuli are presented in quick succession, followed by a numerical cue indicating which of the two to remember[1]. Using this paradigm, information can be decoded equally well when the first stimulus was cued instead of the second[1], demonstrating a robustness to potential interference from the second stimulus. In Experiment 1, we extend this finding by showing that mnemonic representations persisted in the presence of visual masks shown for 11 s (the distractors). Mnemonic representations were just as robust during distractor and no-distractor conditions—the latter entirely without visual interference by deliberate omission of the retro-cue

paradigm. Furthermore, mnemonic information could be recovered at every time point during the 13 s delay (Fig. 2) despite the poorer signal-to-noise of single TR data (compared to data averaged over multiple TRs). Note that this timeframe far surpasses the duration of the stimulus-evoked blood-oxygen-level-dependent (BOLD) response. A comprehensive body of work has shown that stimulus-evoked BOLD alone is generally insufficient for stimulus information to persist into the working memory delay. For example, when people make their response immediately after a retro-cue stimulus sequence, instead of after a long delay, the cued target cannot be decoded[1]. In addition, when presented with a stimulus that has two independent features, only the attended and remembered feature can be decoded during the delay period[2]. This means that, despite identical sensory inputs and task demands at encoding, stimulus-evoked BOLD responses do not carry information about a cued target in the absence of a continued memory requirement. Indeed, once active maintenance of a stimulus feature is no longer needed, information about that feature rapidly drops to chance[1–3,48–50]. Thus, active mnemonic maintenance, and not stimulus-evoked BOLD, can drive the information contained in multivariate fMRI signals during the working memory delay.

In sum, new sensory inputs do not automatically purge working memory information from early retinotopic cortex. Salient and distracting information can, not surprisingly, negatively affect neural representations and behavioral performance. Together, these data suggest that early visual areas actively participate in both sensory and mnemonic processing, possibly serving as a local comparison circuit, and that high-fidelity memories rely on sustained representations in early visual cortex.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41593-019-0428-x.

Received: 4 January 2018; Accepted: 16 May 2019;
Published online: 1 July 2019

References

1. Harrison, S. A. & Tong, F. Decoding reveals the contents of visual working memory in early visual areas. *Nature* **458**, 632–635 (2009).
2. Serences, J. T., Ester, E. F., Vogel, E. K. & Awh, E. Stimulus-specific delay activity in human primary visual cortex. *Psych. Sci.* **20**, 207–214 (2009).
3. Riggall, A. C. & Postle, B. R. The relationship between working memory storage and elevated activity as measured with functional magnetic resonance imagine. *J. Neurosci.* **32**, 12990–12998 (2012).
4. 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).
5. Ester, E. F., Anderson, D. E., Serences, J. T. & Awh, E. A neural measure of precision in visual working memory. *J. Cog. Neurosci.* **25**, 754–761 (2013).
6. Bettencourt, K. C. & Xu, Y. Decoding the content of visual short-term memory under distraction in occipital and parietal areas. *Nat. Neurosci.* **19**, 150–157 (2016).
7. 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).
8. Stokes, M. G. 'Activity-silent' working memory in prefrontal cortex: a dynamic coding framework. *Trends Cog. Sci.* **19**, 394–405 (2015).
9. Ester, E. F., Rademaker, R. L. & Sprague, T. S. How do visual and parietal cortex contribute to visual short-term memory? *eNeuro* **3**, e0041–16 (2016). 2016 1–3.
10. Nassi, J. J. & Callaway, E. M. Parallel processing strategies of the primate visual system. *Nat. Rev. Neurosci.* **10**, 360–372 (2009).
11. Van Kerkoerle, T., Self, M. W. & Roelfsema, P. R. Layer-specificity in the effects of attention and working memory on activity in primary visual cortex. *Nat. Comm.* **8**, 13804 (2017).
12. Miller, E. K., Li, L. & Desimone, R. Activity of neurons in anterior inferior temporal cortex during a short-term memory task. *J. Neurosci.* **13**, 1460–1478 (1993).
13. Serences, J. T. Neural mechanisms of information storage in visual short-term memory. *Vis. Res.* **128**, 53–67 (2016).
14. Brouwer, G. J. & Heeger, D. J. Decoding and reconstructing color from responses in human visual cortex. *J. Neurosci.* **29**, 13992–14003 (2009).
15. Sprague, T. C., Saproo, S. & Serences, J. T. Visual attention mitigates information loss in small- and large-scale neural codes. *Trends Cogn. Sci.* **19**, 215–226 (2015).
16. Rademaker, R. L., Bloem, I. M., De Weerd, P. & Sack, A. S. The impact of interference on short-term memory for visual orientation. *J. Exp. Psychol. Hum. Percept. Perform.* **41**, 1650–1665 (2015).
17. Wildegger, T., Meyers, N. E., Humphreys, G. & Nobre, A. C. Supraliminal but not subliminal distracters bias working memory recall. *J. Exp. Psychol. Hum. Percept. Perform.* **41**, 826–839 (2015).
18. Silver, M. A., Ress, D. & Heeger, D. J. Topographic maps of visual spatial attention in human parietal cortex. *J. Neurophysiol.* **94**, 1358–1371 (2005).
19. Serences, J. T. & Yantis, S. Selective visual attention and perceptual coherence. *Trends Cogn. Sci.* **10**, 38–45 (2006).
20. Poltoratski, S., Ling, S., McCormack, D. & Tong, F. Characterizing the effects of feature salience and top-down attention in the early visual system. *J. Neurophysiol.* **118**, 564–573 (2017).
21. Sprague, T. C., Itthipuripat, S., Vo, V. A. & Serences, J. T. Dissociable signatures of visual salience and behavioral relevance across attentional priority maps in human cortex. *J. Neurophysiol.* **119**, 2153–2165 (2018).
22. Murray, J. D. et al. Stable population coding for working memory coexists with heterogeneous neural dynamics in prefrontal cortex. *Proc. Natl Acad. Sci. USA* **114**, 394–399 (2017).
23. DiCarlo, J. J., Zoccolan, D. & Rust, N. C. How does the brain solve visual object recognition? *Neuron* **73**, 415–434 (2012).
24. Rademaker, R. L., Park, Y. E., Sack, A. T. & Tong, F. Evidence of gradual loss of precision for simple features and complex objects in visual working memory. *J. Exp. Psychol. Hum. Percept. Perform.* **44**, 925–940 (2018).
25. Bisley, J. W., Zaksas, D., Droll, J. A. & Pasternak, T. Activity of neurons in cortical area MT during a memory for motion task. *J. Neurophysiol.* **91**, 286–300 (2004).
26. Zaksas, D. & Paternak, T. Direction signals in the prefrontal cortex and in area MT during a working memory for visual motion task. *J. Neurosci.* **26**, 11726–11742 (2006).
27. Gayet, S. et al. Visual working memory enhances the neural response to matching visual input. *J. Neurosci.* **37**, 6638–6647 (2017).
28. Merrikhi, Y. et al. Spatial working memory alters the efficacy of input to visual cortex. *Nat. Comms.* **8**, 15041 (2017).
29. Miller, E. K., Li, L. & Desimone, R. A neural mechanism for working and recognition memory in inferior temporal cortex. *Science* **254**, 1377–1379 (1991).
30. Maunsell, J. H. R., Sclar, G., Nealey, T. A. & DePriest, D. D. Extraretinal representations in area V4 in the macaque monkey. *Vis. Neurosci.* **7**, 561–573 (1991).
31. Miller, E. K. & Desimone, R. Parallel neuronal mechanisms for short-term memory. *Science* **263**, 520–522 (1994).
32. Miller, E. K., Erickson, C. A. & Desimone, R. Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *J. Neurosci.* **16**, 5154–5167 (1996).
33. 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).
34. 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).
35. Silver, M. A. & Kastner, S. Topographic maps in human frontal and parietal cortex. *Trends Cogn. Sci.* **13**, 488–495 (2009).
36. 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).
37. Deutsch, D. Tones and numbers: specificity of interference in immediate memory. *Science* **168**, 1604–1605 (1970).
38. Deutsch, D. Interference in memory between tones adjacent in the musical scale. *J. Exp. Psychol.* **100**, 228–231 (1973).
39. Magnussen, S., Greenlee, M. W., Asplund, R. & Dyrnes, S. Stimulus-specific mechanisms of visual short-term memory. *Vis. Res.* **31**, 1213–1219 (1991).
40. Magnussen, S. & Greenlee, M. W. Retention and disruption of motion information in visual short-term memory. *J. Exp. Psychol. Learn. Mem. Cogn.* **18**, 151–156 (1992).
41. Pasternak, T. & Zaksas, D. Stimulus specificity and temporal dynamics of working memory for visual motion. *J. Neurophysiol.* **90**, 2757–2762 (2003).
42. Van der Stigchel, S., Merten, H., Meeter, M. & Theeuwes, J. The effects of a task-irrelevant visual event on spatial working memory. *Psychon. Bull. Rev.* **14**, 1066–1071 (2007).
43. Huang, J. & Sekuler, R. Distortions in recall from visual memory: two classes of attractors at work. *J. Vis.* **10**, 1–27 (2010).

44. Nemes, V. A., Parry, N. R., Whitaker, D. & McKeefry, D. J. The retention and disruption of color information in human short-term visual memory. *J. Vis.* **12**, 1–14 (2012).
45. Bae, G. Y. & Luck, S. J. Interactions between visual working memory representations. *Atten. Percep. Psychophys.* **79**, 2376–2395 (2017).
46. Lorenc, E. S., Sreenivasan, K. K., Nee, D. E., Vandenbroucke, A. R. E. & D'Esposito, M. Flexible coding of visual working memory representations during distraction. *J. Neurosci.* **38**, 5267–5276 (2018).
47. Chunharas, C., Rademaker, R. L., Brady, T. F. & Serences, J. T. Adaptive memory distortion in visual working memory. Preprint at *PsyArXiv* https://psyarxiv.com/e3m5a/ (2019).
48. Sprague, T. C., Ester, E. F. & Serences, J. T. Restoring latent visual working memory representations in human cortex. *Neuron* **91**, 694–707 (2016).
49. Christophel, T. G., Iamshchinina, P., Yan, C., Allefeld, C. & Haynes, J. D. Cortical specialization for attended versus unattended working memory. *Nat. Neurosci.* **21**, 494–496 (2018).
50. Rose, N. S. et al. Reactivation of latent working memories with transcranial magnetic stimulation. *Science* **354**, 1136–1139 (2016).

Acknowledgements

This work was supported by grant no. NEI R01-EY025872 to J.T.S., and by the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie Grant Agreement No 743941 to R.L.R. We thank A. Jacobson at the UCSD CFMRI for assistance with multiband imaging protocols. We also thank R. van Bergen for assistance setting up an FSL/FreeSurfer retinotopy pipeline, A. Chakraborty for collecting the behavioral data shown in Supplementary Fig. 9 and V. Vo for discussions on statistical analyses.

Author contributions

This study was designed by R.L.R., C.C. and J.T.S. Data were collected by R.L.R., and C.C. and R.L.R. preprocessed the data. R.L.R. and J.T.S. did the main analyses and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41593-019-0428-x.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to R.L.R. or J.T.S.

Peer review information: *Nature Neuroscience* thanks Thomas Christophel and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2019

Methods

Participants. Six volunteers (five female) between the ages of 21 and 32 years (s.d. = 3.67) participated in Experiment 1, and seven volunteers (five female) between the ages of 24 and 35 years (s.d. = 3.994) participated in Experiment 2. Three volunteers (S03, S04 and S05) participated in both experiments. No statistical methods were used to pre-determine sample sizes, but our sample sizes are similar to those reported in previous publications[1,2,6]. Participants had varying amounts of experience with fMRI experiments, ranging from scanner-naïve (S02, S07 and S09) to highly experienced (that is, >10 h in the scanner; S01, S04, S05 and S10). For a separate behavioral experiment (Supplementary Fig. 9) we recruited 21 participants (14 female; mean age = 20.12, s.d. = 2.007), of whom 17 were included in the analysis (three dropped out and one was excluded due to chance-level performance). The study was conducted at the University of California, San Diego, and approved by the local Institutional Review Board. All participants provided written informed consent, had normal or corrected-to-normal vision and received monetary reimbursement for their time ($10 per hour for behavior, $20 per hour for fMRI, except for S10, who is one of the authors).

Stimuli and procedure Experiment 1. All stimuli in Experiment 1 were projected on a 120 × 90 cm screen placed at the foot-end of the scanner and viewed through a tilted mirror from ~370 cm in an otherwise darkened room. Stimuli were generated on a Macbook Air running OSX using MATLAB 8.1 and the Psychophysics toolbox[51,52]. The luminance output from the projector was linearized in the stimulus presentation code. All stimuli were presented against a 62.82 cd m^{-2} uniform gray background. Stimuli presented during the memory task (targets and distractors, Fig. 1a) were configured in a donut-shaped circular aperture with a 1.5° and 7° inner and outer radius, respectively, and smoothed edges (1° Gaussian kernel; s.d. = 0.5°). Memory targets were full contrast sinusoidal gratings with a spatial frequency of two cycles per degree. Distractors were either gratings or Fourier-filtered noise stimuli, both with a Michelson contrast of 50%. Noise distractors were created by filtering white noise to include only spatial frequencies between 1 and 4 cycles per degree (all stimulus code is available on OSF).

To ensure that the distribution of remembered orientations was approximately uniform across all trials in the experiment, the orientation of the memory target on each trial was chosen from one of six orientation bins. Each bin contained 30 orientations in integer increments, and orientations were drawn randomly from each of the six bins with equal probability. The orientation of the distractor, on trials that contained an oriented grating distractor, was chosen using the same procedure. Moreover, we counterbalanced the orientation bins from which target and distractor orientations were drawn. This ensured that distractor orientations were also distributed uniformly across all trials in the experiment and that the target and distractor orientations were uncorrelated across trials (see also Supplementary Fig. 1a).

On every trial, we randomly chose the spatial phase of the memory target grating. Depending on the distractor condition, we also selected either a random spatial phase for the distractor grating or a random seed to generate the noise distractor. Each initial stimulus was then toggled back and forth between its original and inverted contrast at 4 Hz, without blank gaps in between, for as long as the stimulus was on the screen. Thus, the memory target (500 ms total duration) cycled through one contrast reversal (that is, 250 ms per contrast). This single counter-phase contrast reversal was specifically designed to minimize afterimages induced by the memory target[53]. Similarly, distractors (11 s total duration) contrast reversed for 22 cycles. The recall probe consisted of two white line segments that were 5.5° long and 0.035° wide, with each segment presented at the same distance from fixation as the donut-shaped target and distractor stimuli. A 0.4° central black dot was presented continuously on each block of trials to facilitate fixation.

Each trial of the memory task (Fig. 1a) started with a 1.4 s change in the color of the central fixation dot, indicating with 100% validity the distractor condition during the delay (no distractor, grating distractor, noise distractor). Cues could be blue, green or red. The pairing of cue-colors with distractor conditions was randomized across participants. Following the cue, a memory target was shown for 500 ms and participants remembered its orientation over a 13 s delay. A contrast-reversing noise (one-third of trials) or grating (one-third of trials) distractor was presented for 11 s during the middle portion of the delay, or the screen remained gray throughout the 13 s delay (one-third of trials). After the delay, participants used four buttons to rotate the recall probe around fixation, matching the remembered orientation as precisely as possible. The left two buttons rotated the line counter clockwise, while the right two buttons rotated it clockwise. Using the outer- or inner-most buttons would result in faster or slower rotation of the recall probe, respectively. Participants had 3 s to respond before being presented with the next memory target 3, 5, or 8 s later. Each run consisted of 12 memory trials, and lasted 4 min and 40.8 s. Distractor type (none, grating or noise) and the orientation bin (one of six) from which the target or distractor grating orientations were drawn, were fully counterbalanced across nine consecutive runs of the memory task. Data for 27 total runs were acquired across three separate scanning sessions. Before starting the fMRI experiment, participants practiced the memory task outside the scanner until they were comfortable using the response buttons to recall the target orientation within the temporally restricted response window, and mean absolute response error was <10° (this took between 6 and 12 trials for all participants).

In addition to the memory task, Experiment 1 also included an independent mapping task. During this task, participants viewed 9-s blocks of donut-shaped gratings (same dimensions as in the memory task) or circle-shaped gratings (1.5° radius) that were contrast-reversing at 5 Hz. The orientation of each grating was chosen at random from one of ten orientation bins, and from each bin equally often during a run, to approximate an even sampling of orientation space. Per run, 20 blocks of donut-shaped gratings were alternated with 20 blocks of circle-shaped gratings, with four fixation blocks interspersed. Each run took 7 min. Participants performed a detection task to ensure attention at the physical location of the stimuli: Grating contrast was probabilistically dimmed twice every 9 s, from 100 to 80% for 200 ms. Because the contrast change was probabilistic, there was no change on some stimulus blocks, while on others there were >2 changes. Participants maintained fixation on a 0.4° mean gray dot with a 0.2° magenta dot on top. Note that the donut-shaped stimuli in the mapping task occupied the same physical location as the donut-shaped target and distractor stimuli in the main memory task. This allowed us to independently identify voxels in early visual areas that selectively responded to the spatial position of the memory target. During each scanning session, participants completed 4–6 runs of the mapping task (15–17 total runs across days). Three participants (S02, S03 and S04) practiced one block of the mapping task before the experiment.

Stimuli and procedure Experiment 2. In Experiment 2, all stimuli were projected on a 21.3 cm × 16 cm screen placed inside the scanner bore, viewed from ~40 cm through a tilted mirror. Stimuli were generated using Ubuntu 14.04, MATLAB 9.3 and the Psychophysics toolbox[51,52]. During the memory task, memory targets were full contrast circular sinusoidal gratings (radius = 14.58°) with smoothed edges (1.33° kernel; s.d. = 0.67°) and a spatial frequency of 1.5 cycles per degree. Distractor stimuli were either gratings shown at 50% Michelson, or pictures of faces[54] and gazebos[6] (maximal extent = 27.83°, adapted after ref. [6]). All pictures had the same mean luminance, which was equal to the gray background. The memory target contrast-reversed once, just as in Experiment 1. However, unlike Experiment 1 (where distractors were also contrast reversing), distractors in Experiment 2 were toggled on and off at 4 Hz (that is, one cycle consisted of a 250 ms distractor image and a 250 ms blank screen). On a picture distractor trial, we either showed the full set of 22 unique face images or the full set of 22 unique gazebo images in a randomly shuffled order. On a grating-distractor trial, we showed 22 gratings, each with the same orientation but a randomly chosen phase (0–2π). Target and distractor grating orientations were pseudo-randomly chosen from one of six orientation bins to ensure a roughly uniform sampling of orientation space, identical to the procedure used in Experiment 1 (see also Supplementary Fig. 1b). The recall probe consisted of a 0.056° wide and 29.17° long black line. This line was interrupted by a 0.53° black central fixation dot presented on top of a 0.81° mean gray circle. This fixation dot was presented throughout to aid fixation.

The procedure during the memory task (Fig. 3a) was identical to that of Experiment 1, with the following exceptions: the noise distractor condition was replaced with a picture distractor condition. On half of these trials, pictures of faces were shown and, on the other half, pictures of gazebos were shown. In both the grating and picture distractor conditions, the distractors started flickering on and off 1 s into the 12-s delay, and the recall probe appeared immediately after the last off period. Participants had 4 s to rotate the dial. Participants were scanned on three separate days, completing a total of 27 total runs (9 runs per day, 12 trials per run) of the memory task. Before scanning, participants practiced for 12–24 trials until their absolute performance was <10°. Only S09 did not reach this criterion during practice, with a performance of 11.5° after 36 practice trials.

Experiment 2 used two different mapping tasks: During the first mapping task, participants viewed a series of gratings (50% of trials), face pictures (25% of trials) or gazebo pictures (25% of trials) that were flickered on (250 ms) and off (250 ms) at 4 Hz for a total of 5.5 s (that is, 11 stimuli per trial). Each trial was followed by a 3, 5, or 8 s inter-trial interval. Grating and picture stimuli were identical to the ones described above for the Experiment 2 memory task. On grating trials, the orientation was chosen at random from one of 12 orientation bins (to ensure approximately uniform sampling of orientation space as in Experiment 1). Each of the 22 unique face images was shown three times during a run. Face images were randomly shuffled across all trials in a run, with the restriction that the same image was never shown twice in a row. The same was true for gazebo images. Participants completed 24 trials per run (4 min and 31.2 s per run). Across the three scanning days, participants completed between 20 and 29 total runs of this first mapping task. Three participants (S04, S05 and S09) practiced one run of the task before going into the scanner.

The second mapping task of Experiment 2 comprised trials showing either a circle-shaped (1.06° radius) or donut-shaped (1.06° inner and 14.74° outer radius) grating stimulus (spatial frequency 1.43 cycles per degree; edges smoothed with 0.69° kernel and s.d. = 0.36°). On every trial, a 6 s grating was contrast-reversing (as in Experiment 1) at 4 Hz (that is, 250 ms per contrast), followed by a 3, 5, or 8 s inter-trial interval. Grating orientation was randomly chosen from one of nine orientation bins on each trial, and equally often from each bin within a run. Participants completed 36 trials per run (18 circle-shaped grating trials and 18

donut-shaped grating trials, randomly interleaved), and each run took 7 min and 5.6 s. A central black dot (0.56°) aided fixation throughout. Data for this second mapping task were collected separately from the other Experiment 2 data (that is, different scanning sessions). Participants completed between 10 and 20 total runs of the second mapping task.

During both mapping tasks in Experiment 2, we occasionally (0–3 times per trial) superimposed small smoothed circles (of a uniform light gray color) on the mapping stimuli for 250 ms. These brief 'blobs' could be centered at any distance from fixation occupied by a stimulus (although no closer than 0.056° and no further than 13.78°), and at any angle relative to fixation (1–360°). Blobs were scaled for cortical magnification[55], such that all blobs (that is, at every distance from fixation) stimulated roughly 1 mm of cortex. In terms of visual angle, this means blobs had radii spanning from 0.18° to 0.75°. No blobs were presented during the first or last 500 ms of a trial, or within 500 ms of each other. Participants pressed a button every time they detected a blob superimposed on a stimulus image, such that they stayed alert and attending the location of the mapping stimuli.

Magnetic resonance imaging. All scans were performed on a General Electric Discovery MR750 3.0T scanner located at the University of California, San Diego (UCSD), Keck Center for Functional Magnetic Resonance Imaging (CFMRI). High-resolution (1 mm^3 isotropic) anatomical images were acquired during a retinotopic mapping session, using an Invivo eight-channel head coil. Functional echo-planar imaging (EPI) data for the current experiment were acquired using a Nova Medical 32-channel head coil (NMSC075-32-3GE-MR750) and the Stanford Simultaneous Multi-Slice EPI sequence (MUX EPI), using nine axial slices per band and a multiband factor of eight (total slices = 72; 2 mm^3 isotropic; 0 mm gap; matrix = 104 × 104; field of view = 20.8 cm; repetition time/echo time (TR/TE) = 800/35 ms, flip angle = 52°; inplane acceleration = 1). At sequence onset, the initial 16 TRs served as reference images critical to the transformation from *k*-space to image space. Un-aliasing and image reconstruction procedures were performed on local servers using CNI-based reconstruction code. Forward and reverse phase-encoding directions were used during the acquisition of two short (17 s) 'topup' datasets. From these images, susceptibility-induced off-resonance fields were estimated[56] and used to correct signal distortion inherent in EPI sequences, using FSL topup[57,58].

Preprocessing. All imaging data were preprocessed using software tools developed and distributed by FreeSurfer and FSL (free to download at https://surfer.nmr.mgh.harvard.edu and http://www.fmrib.ox.ac.uk/fsl). Cortical surface gray-white matter volumetric segmentation of the high-resolution anatomical image was performed using the 'recon-all' utility in the FreeSurfer analysis suite[59]. Segmented T1 data were used to define ROIs for use in subsequent analyses. The first volume of every first functional run of a scanning session was then coregistered to this common anatomical image. Transformation matrices were generated using FreeSurfer's manual and boundary-based registration tools[60]. These matrices were then used to transform each four-dimensional functional volume using FSL FLIRT[61,62], such that all cross-session data from a single participant was in the same space. Next, motion correction was performed using the FSL tool MCFLIRT[62] without spatial smoothing, a final sinc interpolation stage, and 12 degrees of freedom. Slow drifts in the data were removed last, using a high pass filer (1/40 Hz cutoff). No additional spatial smoothing was applied to the data apart from the smoothing inherent to resampling and motion correction.

Signal amplitude time-series were normalized via *Z*-scoring on a voxel-by-voxel and run-by-run basis. *Z*-scored data were used for all further analyses. Trial events were jittered with respect to TR onsets, and trial events were rounded to the nearest TR. To recover the univariate BOLD time courses for all three memory distractor conditions in Experiments 1 and 2, we estimated the hemodynamic response function for each voxel at each time point of interest (0–19.5 s from memory target onset). This was done using a finite impulse response function model[63] consisting of a column marking the onset of each event (memory target onset) with a '1', and then a series of temporally shifted version of that initial regressor in subsequent columns to model the BOLD response at each subsequent time point. Estimated hemodynamic response functions were then averaged across all voxels in each ROI (see also Supplementary Fig. 3). The analyses performed after preprocessing was completed were all done in MATLAB 9.1 using custom functions.

Identifying ROIs. To identify voxels that were visually responsive to the donut-stimuli, a general linear model was performed on data from the mapping task (for Experiment 2 we used data from the second mapping task) using FSL FEAT (FMRI Expert Analysis Tool, v.6.00). Individual mapping runs were analyzed using the brain extraction tool[64] and data prewhitening using FILM[65]. Predicted BOLD responses were generated for blocks of 'donut and circle' stimuli by convolving the stimulus sequence with a canonical gamma hemodynamic response function (phase = 0 s, s.d. = 3 s, lag = 6 s). The temporal derivative was included as an additional regressor to accommodate slight temporal shifts in the waveform to yield better model fits and to increase explained variance. Individual runs were combined using a standard weighted fixed effects model. Voxels that were significantly more activated by the donut compared to the circle ($P = 0.05$; false discovery rate corrected) were defined as visually responsive and used in subsequent analyses.

Standard retinotopic mapping procedures[66,67] were employed to define nine a priori ROIs in early visual (V1–V3, V3AB, hV4) and parietal (IPS0–IPS3) cortex. Retinotopic mapping data were acquired during an independent scanning session that used both meridian mapping techniques (with checkerboard 'bowtie' stimuli shown alternating between the horizontal and vertical meridian) and polar angle techniques (with a slowly rotating checkerboard wedge) to identify the visual field preferences of voxels (stimuli described in more detail in ref. [68]). Anatomical and functional retinotopy analyses were performed using a set of custom wrappers that encapsulated existing FreeSurfer and FSL functionality. ROIs were combined across left and right hemispheres and across dorsal and ventral areas (for V1–V3) by concatenating voxels.

For all analyses (except the one presented in Supplementray Fig. 13), only visually responsive voxels, selected using the localizer procedure described above, were included in the ROI of each retinotopic area. We only included data for retinotopic areas in which the number of visually responsive voxels exceeded 20 for every single participant. Exact voxel counts for each participant in each ROI can be found in Supplementary Tables 9 and 10 for Experiments 1 and 2, respectively.

fMRI analyses: IEM. To generate model-based reconstructions of remembered and perceived orientations from voxel responses, an IEM was implemented[14,15] with orientation as the feature dimension. The first step in this analysis is to estimate an encoding model using voxel responses in a cortical region of interest. Data used during this step are considered the 'training set' (Supplementary Fig. 4a, left), and are combined with nine idealized tuning functions, or 'channels' (Supplementary Fig. 4a, right), to parameterize an orientation sensitivity profile for each voxel. The second step in the analysis combines the estimated sensitivity profiles in each voxel with a novel pattern of all voxel responses in a ROI on a single trial (data from the 'test set', Supplementary Fig. 4b, left) to reconstruct a model-based representation of the orientation that was remembered or viewed on that trial (Supplementary Fig. 4b, right). The encoding model for a single voxel has the general form

$$R_j = \sum_{i}^{9} w_i \, c_i \tag{1}$$

where R_j is the response R of voxel j, and c_i is the channel magnitude c at the ith of nine channels. A voxel's sensitivity profile over orientation space is captured by nine weights, w. Channels were modeled as

$$c(\theta) = \cos\left((\theta - \mu)\frac{\pi}{180}\right)^8 \tag{2}$$

where θ is degrees in orientation space (defined over 180°) and the channel center is μ. Channel centers were spaced 20° apart.

For the first step of the IEM, equation (1) can be expressed as

$$B_1 = WC_1 \tag{3}$$

Here, a matrix of observed BOLD responses B_1 (m voxels × n trials) is related to a matrix of modeled channel responses C_1 (k channels × n trials) by a weight matrix W (m voxels × k channels). For each trial, C_1 contains the pointwise product of a stimulus mask (that is, 1 at the true stimulus orientation, 0 at all other orientations) with the idealized tuning functions. W quantifies the sensitivity of each voxel at each idealized orientation channel, and can be computed with least-squares linear regression:

$$\hat{W} = B_1 C_1^{\mathrm{T}} (C_1 C_1^{\mathrm{T}})^{-1} \tag{4}$$

Estimating the sensitivity profiles concludes the first encoding step of the IEM. The second step of the IEM inverts the model, using the estimated sensitivity profiles of all voxels $\hat{W}$ (m voxels × k channels) in combination with a test set of novel BOLD response data B_2 (m voxels × n trials) to estimate the response of each channel on each trial $\hat{C}_2$ (k channels × n trials):

$$\hat{C}_2 = (\hat{W}^{\mathrm{T}}\hat{W})^{-1}\hat{W}^{\mathrm{T}}B_2 \tag{5}$$

This step uses the Moore–Penrose pseudoinverse of $\hat{W}$ and it is multivariate in nature, since it uses the sensitivity profiles across all voxels to jointly estimate channel responses $\hat{C}_2$ for each trial of the test set. This effectively forms a model-based 'reconstruction' of the remembered or seen stimulus feature on a trial-by-trial basis.

Because grating orientations could take any integer value between 1° and 180°, both the encoding (equation (4)) and inversion (equation (5)) steps of the IEM were repeated 20 times. On each repeat, the centers of nine idealized tuning functions were shifted by 1° (equation (2)), and we estimated the channel responses $\hat{C}_2$ at those nine centers, until the entire 180° orientation space was estimated in 1°

steps. This procedure thus yielded estimated channel responses $\hat{C}_2$ for each degree in orientation space. After generating reconstructions for each trial, all single-trial reconstructions were re-centered on the remembered orientation (when looking at information for mnemonic orientations) or on the orientation of the directly viewed distractor grating (when looking at information for viewed orientations).

For the IEM analyses investigating mnemonic codes based on sensory-driven responses, we used independent data from the mapping tasks as the training set and data from the memory task as the test set. For Experiment 2, we combined data across the two mapping tasks to comprise the training set. For the IEM analyses investigating mnemonic codes that are not necessarily based on sensory-driven responses, we used a within-condition leave-one-out procedure. Here, all but one trial is used as the training set, and the left-out trial constitutes the test set. This procedure is repeated until all trials in a given condition have been left out once. Of note, we also did these analyses leaving one session out, yielding qualitatively similar results. To obtain single-trial activity estimates, memory data were averaged over a time window of 5.6–13.6 s (7–17 TRs) after target onset. Mapping data in Experiment 1 were averaged over 4.8–9.6 s (6–12 TRs) after donut onset. Mapping data from both Experiment 2 mapping tasks were averaged over 2.4–7.2 s (3–9 TRs) after donut onset.

fMRI analyses: reconstruction fidelity. Model-based reconstructions of orientation were quantified using a fidelity metric derived from trigonometry (Fig. 1d)[69,48]. Unless specified otherwise, this fidelity metric was applied to the average of 108 single-trial reconstructions (that is, all trials from a given distractor condition), separately for each condition, participant, and ROI. For each reconstruction, the fidelity metric was calculated by taking the channel response at each degree in orientation space (wrapped onto a 2π circle), and projecting this vector onto the center of the stimulus space (that is, onto 0°) via $\cos A_{\mathrm{abs}(0°-d)} = \frac{b}{h}$, where A_{abs} is the absolute angle between the reconstruction center (at 0°) and the degree in orientation space being evaluated (d), and h is the channel response at d (that is, the hypotenuse of a right triangle). In other words, we projected the length of vector h onto 0° by solving for b (that is, the adjacent side of a right triangle). This procedure was repeated for all 180° in orientation space, after which we calculated the mean of all 180 projected vectors. Thus, the mean projected vector—our fidelity metric—reflects the amount of 'energy' at a remembered or sensed orientation. Note that this metric, by design, gets rid of additive offsets and captures only the amount of information at the center of the reconstruction.

fMRI analyses: decoding. In addition to using an inverted encoding model to analyze our data (that is, the IEM analysis described above), we also used a more conventional multivariate pattern analysis decoding approach. This allowed us to evaluate the extent to which our results generalized across different analysis techniques. It also allowed us to more directly compare our results to previous work by Bettencourt and Xu[6], who used a decoder to perform a two-way classification between orthogonal orientations. To analyze our data in an analogous manner, despite our use of continuous orientations (1°–180°), we performed two two-way classifications. For the first classification, we binned all orientations within a 45° window around vertical, and we binned all orientations within a 45° window around horizontal. We then performed a two-way classification to decode between the two cardinal axes (Fig. 5a). For the second classification, we binned all orientations within a 45° window around the two oblique axes (that is, around 45° and 135°) to classify between the two oblique orientations (Fig. 5a). Finally, decoding performance was averaged across the two two-way classifications (that is, the cardinal classification and the oblique classification) before performing statistics and plotting (Fig. 5b). We used the MATLAB built-in multi-class support vector machine (SVM) ('fitcecoc' and 'predict' functions). As with our IEM analyses, we performed the decoding analysis in two different ways: (1) training the SVM on independent localizer data and decoding the orientation from the working memory delay epoch, and (2) training and testing on data from the memory delay, using a leave-one-trial-out procedure. Decoding results based on the independent and leave-one-out training schemes are plotted on the left and right side of Fig. 5b, respectively. For Experiment 2, we also evaluated whether the picture distractors shown during the delay (faces and gazebos) could be decoded. Using the same classifier described above, we trained an SVM on independent data from the first Experiment 2 mapping task (that is, the one with trials showing pictures of faces and gazebos) to do a two-way classification. We then decoded the presence of either a face of gazebo distractor during the working memory delay-period epoch (Supplementary Fig. 10).

Statistical procedures. All statistical statements reported here were based on permutation (or randomization) testing over 1,000 iterations with scrambled data labels. Note that this constrains the resolution of our P values to a lower limit of $P \leq 0.001$. To test whether fidelity metrics were significantly greater than zero, we generated permuted null distributions of fidelities for each participant, ROI, and condition (and for each timepoint, in the analyses shown in Figs. 2b and 3e and Supplementary Figs. 6 and 12). On each permutation, we first reshuffled target orientation labels for all trials before performing the inversion step of the IEM (effectively randomizing single-trial reconstructions relative to the true orientation). Second, we calculated the fidelity for the trial-averaged shuffled orientation reconstructions in a manner identical to calculating fidelity for the intact reconstructions. This resulted in one 'null' fidelity estimate per permutation. Combining the null fidelities across all participants (so six and seven fidelities for Experiments 1 and 2, respectively) resulted in one t-statistic per permutation $t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$, where $\bar{x}$ and s are the mean and standard deviation of fidelities across participants, μ_0 is the null hypothesis mean (that is, 0) and n is the sample size. To test across-participant fidelities against zero (Figs. 1e, 2b, 3d,e and 4 and Supplementary Figs. 6, 12 and 13) we compared the t-statistic calculated from the intact data against the permuted null distribution of t-statistics for that condition, ROI, and timepoint. Reported tests against zero were one-sided and uncorrected. Note that the same procedure was used for the decoding analyses (Fig. 5 and Supplementary Fig. 10), with the exception that the null hypothesis for the t-statistic was 0.5 (that is, chance level). Significant fidelity (and decoding) is indicated in our figures by colored and gray asterisks.

To test whether there were differences in fidelity between distractor conditions (within each ROI; Figs. 1e, 3d, 4 and 5 and Supplementary Figs. 8 and 13), we used within-subjects repeated-measures one-way analysis of variance (ANOVA). First, we calculated the F-statistic from the intact data for the effect of condition. Next, we generated permuted null distributions of F by shuffling the condition labels per subject, and calculating the null F-statistic on each permutation. Each data-derived F was then compared to its null distribution of F to get the P value. Significant effects were followed up with post-hoc paired-sample t-tests: We performed pairwise comparisons between each of the conditions, comparing the data-derived $t = \frac{\overline{X}_D - \mu_0}{s_D / \sqrt{n}}$ (with $\overline{X}_D$ and s_D denoting the mean and s.d. of the pairwise differences), against a permuted t distribution generated by reshuffling condition labels on each iteration. Significant comparisons are indicated in our figures by black asterisks.

To look for a signature of top-down related processing, we used the grating-distractor condition, as it allowed us to directly compare fidelities from remembered and sensed orientations (which were derived from the same exact data). We used a within-subjects repeated-measures two-way ANOVA to track differences in fidelity across ROIs and memory/sensory condition (Figs. 1e, 3d, 4 and 5). We only included V1, V2, V3, V3AB, and V4 as our ROIs, as the hierarchical relationship between these areas is still fairly clear. Permutations were done as described above, but now based on F-statistics for the two main effects of ROI and memory/sensory condition, as well as their interaction. Specifically, the aim of this analysis was to look for significant interactions between ROI and memory/sensory condition, and to test whether memory and sensory representations became more and less pronounced, respectively, as one ascends the visual hierarchy. There are various ways in which ROIs might systematically differ from one another, such as their size (that is, number of voxels), sensitivity to neural activity, or signal-to-noise ratio. Because these factors might affect the attainable accuracy of multivariate analysis tools, a direct comparison across ROIs is generally not recommended[70]. However, because here we are looking for an interaction specifically (and not an absolute difference between ROIs), and because we apply our multivariate techniques to the exact same data in each ROI (that is, information about either the remembered or sensed orientation) these caveats are not of concern in this particular case. One remaining concern is that also the scaling might not be comparable between ROIs (that is, a difference of X in one ROI may not mean the same as a similar difference of X in another ROI), although this concern is not reflected by the data presented here.

When circular statistics were used, these were calculated using the circular statistics toolbox[71].

Glitches. S01 completed four sessions of scanning, but on the first scanning day the projector settings had been changed such that we were presenting stimuli as ovals rather than circles. Data from this session were excluded from analysis. S02 was also scanned four times, but data transfer after one of the sessions failed, and the data were removed from the scanner center's servers before being backed up—and thus lost. For S04, we only collected four mapping runs on the first day of scanning because the scanner computer hard drive was full by the time we approached the end of the scan, causing the computer to freeze. On the first day of scanning S05 the scanner computer started spontaneously deleting data files half way through the session. Consequently, data from the second mapping run were deleted and the fifth memory run was aborted (with imaging data collection incomplete, while behavioral data collection was complete). To ensure full counterbalancing, an exact replica of this fifth memory run was repeated as the first run on the second day of scanning. During the last scanning session of S09, the subject reported repeated but brief instances of falling asleep. Probed further, S09 indicated having also slept occasionally during the previous three sessions. Because we could no longer determine the runs during which S09 was asleep, we decided to keep all S09 data for analyses and to scan one additional subject for Experiment 2.

Miscellaneous. During data collection, participants were not blinded to the experimental conditions (that is, they could clearly perceive the distractor condition on every trial), while experimenters were blinded (that is, they were not in the room and conditions were interleaved). Analyses were not performed blind to the conditions of the experiments.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

We have uploaded all preprocessed fMRI and behavioral data, from each subject and ROI, to the Open Science Framework (OSF) at https://osf.io/dkx6y. An accompanying wiki is available here as well, providing an overview of all the data and code.

Code availability

The experiment code used during data collection and the analysis code used to generate the figures in the main manuscript and Supplementary Materials is available from the Open Science Framework (OSF) at https://osf.io/dkx6y.

References

51. Brainard, D. H. The Psychophysics Toolbox. *Spat. Vis.* **10**, 433–436 (1997).
52. Kleiner, M. et al. What's new in psychtoolbox-3. *Perception* **36**, 1–16 (2007).
53. Tyler, C. W. & Nakayama, K. Grating induction: a new type of aftereffect. *Vis. Res.* **20**, 437–441 (1980).
54. Goeleven, E., De Raedt, R., Leyman, L. & Verschuere, B. The Karolinska directed emotional faces: a validation study. *Cogn. Emot.* **22**, 1094–1118 (2008).
55. Rovamo, J. & Virsu, V. An estimation and application of the human cortical magnification factor. *Exp. Brain Res.* **37**, 495–510 (1979).
56. Andersson, J. L. R., Skare, S. & Ashburner, J. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *Neuroimage* **20**, 870–888 (2003).
57. Smith, S. M. et al. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* **23**, 208–219 (2004).
58. Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W. & Smith, S. M. FSL. *Neuroimage* **62**, 782–790 (2012).
59. Dale, A. M., Fischl, B. & Sereno, M. I. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage* **9**, 179–194 (1999).
60. Greve, D. & Fischl, B. Accurate and robust brain image alignment using boundary-based registration. *Neuroimage* **48**, 63–72 (2009).
61. Jenkinson, M. & Smith, S. M. A global optimisation method for robust affine registration of brain images. *Med. Image Anal.* **5**, 143–156 (2001).
62. Jenkinson, M., Bannister, P., Brady, J. M. & Smith, S. M. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* **17**, 825–841 (2002).
63. Dale, A. M. Optimal experimental design for event-related fMRI. *Hum. Brain Mapp.* **8**, 109–114 (1999).
64. Smith, S. M. Fast robust automated brain extraction. *Hum. Brain Mapp.* **17**, 143–155 (2002).
65. Woolrich, M. W., Ripley, B. D., Brady, M. & Smith, S. M. Temporal autocorrelation in univariate linear modeling of FMRI data. *Neuroimage* **14**, 1370–1386 (2001).
66. Engel, S. A. et al. fMRI of human visual cortex. *Nature* **369**, 525 (1994).
67. 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).
68. Sprague, T. C. & Serences, J. T. Attention modulates spatial priority maps in the human occipital, parietal and frontal cortices. *Nat. Neurosci.* **16**, 1879–1887 (2013).
69. Wolff, M. J., Jochim, J., Akyürek, E. G. & Stokes, M. G. Dynamic hidden states underlying working-memory-guided behavior. *Nat. Neurosci.* **20**, 864–871 (2017).
70. Haynes, J. D. A primer on pattern-based approaches to fMRI: principles, pitfalls, and perspectives. *Neuron* **87**, 257–270 (2015).
71. Berens, P. CircStat: a MATLAB toolbox for circular statistics. *J. Stat. Softw.* **31**, 1–21 (2009).

nature research

Corresponding author(s): Rosanne L. Rademaker

Last updated by author(s): Apr 25, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Data were collected using task code written under Matlab (2013a & 2017b for Experiments 1 & 2, respectively) and the Psychophysics Toolbox (Brainard, 1997).
Data analysis	functional Magnetic Resonance Imaging (fMRI) data were preprocessed using the Oxford FMRIB Software Library (abbreviated FSL) and the FreeSurfer Software Suite. After preprocessing, custom analysis code was written in Matlab (using versions R2016b and R2017a), some of this code uses the circular statistics toolbox (Berens, 2009). Behavioral data were analyzed in Matlab 2017a, also occasionally using the circular toolbox.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All fMRI data, behavioral data, experiment code, and analysis code required to generate all the figures in this paper are publicly available through the open Science Framework at https://osf.io/dkx6y.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	We collected quantitative behavioral and neuroimaging data while human participants performed a simple computer task (remembering orientations over a brief delay).
Research sample	Six human research participants (5 female) between the ages of 21 and 32 years took part Experiment 1 (fMRI). Seven human research participants (5 female) between the ages of 24 and 35 years took part in Experiment 2 (fMRI). Twenty-one human research participants (14 female) between the ages of 18 and 24 took part in the behavioral study. All our participants had normal or corrected-to-normal vision, were neurologically intact, and reported no history of psychiatric disorders. All of our participants received monetary reimbursement for their time ($20 an hour for fMRI, $10 an hour for behavior).
Sampling strategy	The fMRI sample sizes were based on previous fMRI studies decoding memory for orientation using a single scanning session's worth of data with sample sizes between 6 (Harrison & Tong, Nature, 2009) and 10 (Serences, Ester, Vogel, & Awh, Psych. Sci., 2009; Bettencourt & Xu, Nat. Neurosci., 2016) observers. Another study collected 2–3 session's worth of data with a sample size of 6 (Ester, Sprague, & Serences, Neuron, 2015). Thus, to be overly cautious, we chose to collect 3 sessions worth of experimental fMRI data for 6 participants (in addition to other sessions collecting the data required to define retinotopic maps). The sample size in the behavioral experiment was based on previous work with a highly comparable paradigm (Rademaker et al., JEP:HPP, 2015) using between 8 and 10 participants across a series of 4 experiments. To account for increases in target-orientation offsets in the current behavioral experiment, we increased both the number of participants and the amount of testing (3–4 days of testing instead of just 1 day of testing) relative to this previous series of experiments. Participant recruitment was based on a sample of convenience. Because for the fMRI experiments independent retinotopy scans were required, and the experiments themselves involved multiple sessions, we recruited people around the Psychology Department willing to commit their time on multiple occasions and who would be around for a year or longer. These people therefore included lab members, graduate students from other labs, and undergraduate research assistants. For the behavioral experiments the participants were recruited via flyer's, and were primarily UCSD undergraduate students.
Data collection	For the fMRI experiments we used laptop computers for stimulus presentation (a Macbook Air running OS X, and a HP running Ubuntu 14.04, for Experiments 1 and 2, respectively). fMRI scanning was performed on a General Electric (GE) Discovery MR750 3.0T scanner. During fMRI data collection both the first and second author were present outside of the scanner room while the participant was in the scanner. Thus, the experimenters were blind to the experimental conditions, as experimenters were not in the room, and all experimental conditions were randomly interleaved. Participants were blind to the purpose of our experiment (with exception of S10 in Experiment 2). Participants were instructed to recall the remembered orientation as precisely possible, and to use the cues (indicating distractor condition) to their advantage. For the behavioral experiment, stimuli were presented on a Dell desktop computer running Ubuntu 14.04. Data collection for the behavioral study was performed by an undergraduate research assistant who was blind to the conditions and hypotheses of the experiment.
Timing	The first scanning session for Experiment 1 was performed on December 15th 2016, and the last session was completed on April 24th 2017. For Experiment 2, the first scanning session was on March 22nd 2018 and the last session was completed on July 23rd 2018. For the behavioral study, data collection started on the 29th of September 2016, and was completed on the 14th of March 2018.
Data exclusions	In the fMRI experiment we excluded one scan session from S01, because during that session the projector settings had been changed without our knowledge, and it turned out that we had been presenting our stimuli as ovals rather than circles. This scan session was repeated at another date. All data from one of S02's sessions were accidentally deleted off the scanner center's servers during data transfer without having been backed up, and were irretrievably lost. This scan session was repeated at another date. Finally, on the first day of scanning S05 the scanner computer was full, and started deleting data files half way through the session. Consequently, data from the 2nd mapping run was deleted and the 5th memory run was aborted halfway through (with imaging data collection incomplete, and thus not usable). To ensure exact counterbalancing, an exact replica of this 5th memory run was repeated as the first run on the next day of scanning. The number of mapping runs varied between sessions, and the lost mapping run was therefore not explicitly repeated (we just collected new mapping runs in subsequent sessions). During the last scanning session of S09, the subject reported repeated but brief instances of falling asleep. Probed further, S09 indicated having also slept occasionally during the 3 sessions prior. Because we could no longer determine the runs during which S09 was asleep, we decided to keep all S09 data for analyses, and to scan one additional subject for Experiment 2. These exclusions, and some other glitches, are also reported in the Online Methods. For the behavioral experiment, we had three participants drop out (completing only 5, 6, or 12 runs, instead of the required 18). Data from these participants was thus excluded from analysis. We also excluded one participant whose behavioral performance was at chance level. This left 17 participants for data analysis.

Non-participation	No participants dropped out from the fMRI experiments. For the behavioral experiment, we had three participants drop out with no reason reported (other than boredom).
Randomization	We did not compare experimental groups of participants, and therefore did not have to randomly assign participants to groups. This was a within-subject design. We did control for the possible impact of the cue color at the start of each experimental trial (the cue indicating distractor condition) by randomly assigning the three colors to three conditions for each participant individually (thus: different colors indicated different conditions to different participants).

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☐	☒ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	We used a population of healthy human subjects, consisting of UCSD postdocs, graduate students, and undergraduate students.
Recruitment	fMRI participants were recruited based on a sample of convenience. We would ask people around the department if they would be interested in participating, if they did, we asked if they were willing to participate in multiple scanning sessions (4 or more) and if they would be around for at least another year for potential follow-up experiments. This selection procedure biases our sample to highly educated individuals who likely had some experience being in Psychology experiments. Participants in the behavioral experiment were primarily by UCSD undergraduate students. By recruitment through flyering around campus, our sample included mostly young adults with at minimum a high school education and potential experience being in Psychology experiments.
Ethics oversight	The study was conducted at the University of California, San Diego, and was approved by the local Institutional Review Board. All participants provided written informed consent.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Magnetic resonance imaging

Experimental design

Design type	The main runs collected during scanning (both in Experiments 1 and 2) consisted of a memory task of which the trials were presented in an event-related manner. A second type of run were the mapping runs. The mapping runs in Experiment 1 were presented in a blocked design, while the mapping runs in Experiment 2 were presented in an event-related manner.
Design specifications	For the main memory task (both in Experiments 1 and 2), each participant completed a total of 27 runs across 3 separate scanning sessions. Within a run, each trial started with a 1.4s color cue, followed by a memory target (0.5s), a delay (13s), and a response period (3s). The inter-trial interval was jittered, and could be 3s, 5s, or 8s. Note that the cue was presented during the tail-end of each preceding inter-trial interval. Thus, trial lengths were 19.5s, 21.5s, or 24.5s. Each run had 12 trials, with each trial length occurring 4 times. Together with fixation periods at the beginning (3s) and end (3s) of each run, the total run length was 268s. In Experiment 1, participants completed 4–6 runs of the mapping task during each scanning session (15–17 total runs across days). Per run, 9s blocks of donut-shaped or circle-shaped grating stimuli were each alternated 20 times per run, with 4 fixation blocks (also 9s long) interspersed. Together with fixation periods at the beginning (3s) and end (8.2s) of each run, the total run length was 407.2s. Per scanning-session we collected 4–6 runs of the mapping task, totaling 15–17 runs across all 3 days of scanning. In Experiment 2 there were two mapping tasks: For the first mapping task participants completed between 20 and 29 total runs across the three scanning days. Each run had 24 trials and lasted 4 minutes and 31.2s. Each trial lasted 5.5s, and the inter-trial interval could be 3, 5, or 8s. For the second mapping task participants completed between 10–20

total runs. Each run had 36 trials and took 7 minutes and 5.6s. Each trial lasted 6s, and the inter-trial interval could be 3, 5, or 8s.

Note: at the scanner, each run took 12.8 extra seconds, because our multiband protocol demanded 16 'calibration TR's' at the start of each run. These calibration TR's are not part of the data.

Behavioral performance measures

There is a button box with 4 buttons at the scanner. All button presses and their latencies were recorded. During the memory task, the button presses rotated a response dial on the screen. The left two buttons rotated the probe counter clockwise, while the right two buttons rotated it clockwise. The outer- or inner-most buttons rotated the probe faster or slower, respectively. To ensure participants were performing the task as expected, we calculated the angle of the dial at the end of the 3s response window, and compared this response to the target orientation shown at the beginning of the trial. Behavioral error was the difference in degrees between recalled and target orientation. Responses centered on an error of 0º indicated that the participants were performing the task as expected. Chance level performance in such a task is 45º. The average errors from our participants were mostly < 10º.

Acquisition

Imaging type(s)

Functional (structural data came from an independent retinotopy session)

Field strength

3 tesla

Sequence & imaging parameters

High resolution (1 mm3 isotropic) anatomical images were acquired during a separate retinotopic mapping session, using an Invivo 8-channel head coil. Functional echo-planar imaging (EPI) data for the current experiments were acquired using a Nova Medical 32-channel head coil (NMSC075-32-3GE-MR750) and the Stanford Simultaneous Multi-Slice (SMS) EPI sequence (MUX EPI), utilizing 9 axial slices per band and a multiband factor of 8 (total slices = 72; 2 mm3 isotropic; 0 mm gap; matrix = 104 x 104; FOV = 20.8 cm; TR/TE = 800/35 ms, flip angle = 52º; inplane acceleration = 1). At sequence onset, the initial 16 TR's served as reference images critical to the transformation from k-space to image space. Unaliasing and image reconstruction procedures were performed on local servers using CNI based reconstruction code. Forward and reverse phase-encoding directions were utilized during the acquisition of two short (17 s) "topup" datasets. From these images, susceptibility-induced off-resonance fields were estimated and used to correct signal distortion inherent in EPI sequences using FSL topup.

Area of acquisition

Whole brain scans were acquired.

Diffusion MRI ☐ Used ☒ Not used

Preprocessing

Preprocessing software

All imaging data were preprocessed using software tools developed and distributed by FreeSurfer and FSL (free to download at https://surfer.nmr.mgh.harvard.edu and http://www.fmrib.ox.ac.uk/fsl). Cortical surface gray-white matter volumetric segmentation of the high resolution anatomical image was performed using the "recon-all" utility in the FreeSurfer analysis suite. Segmented T1 data were used to define Regions of Interest (ROIs) for use in subsequent analyses. The first volume of every functional run was then coregistered to this common anatomical image. Transformation matrices were generated using FreeSurfer's manual and boundary based registration tools. These matrices were then used to transform each 4D functional volume using FSL FLIRT such that all cross-session data from a single participant was in the same space. Next, motion correction was performed using the FSL tool MCFLIRT without spatial smoothing, a final sinc interpolation stage, and 12 degrees of freedom. Slow drifts in the data were removed last, using a high pass filer (1/40 Hz cutoff). No additional spatial smoothing was applied to the data apart from the smoothing inherent to resampling and motion correction.

Normalization

Functional data were kept in each participant's native space. The data were not normalized to a common across-subject template. We did align the functional data to the within-subject anatomy (described above). Signal amplitude timeseries were normalized via Z-scoring on a voxel-by-voxel and run-by-run basis (i.e. also within-subject).

Normalization template

Functional data were kept in each participant's native space.

Noise and artifact removal

We did not use tissue or physiological signals to remove artifact and structured noise. For other preprocessing procedures (motion correction, high pass filtering, etc.) see above.

Volume censoring

we did not apply volume censoring.

Statistical modeling & inference

Model type and settings

In the mapping task, we identified visually responsive voxels by performing a General Linear Model (GLM) using FSL FEAT (FMRI Expert Analysis Tool, version 6.00). Individual mapping runs were analyzed using BET brain extraction and data prewhitening using FILM. Predicted BOLD responses were generated for blocks of "donut" and "circle" stimuli by convolving the stimulus sequence with a canonical gamma hemodynamic response function (phase = 0 s, sd = 3 s, lag = 6 s). The temporal derivative was also included as an additional regressor to accommodate slight temporal shifts in the waveform to yield better model fits and to increase explained variance. Individual runs were combined using a standard weighted fixed effects model. Voxels that were significantly more activated by the donut compared to the circle (q = 0.05; FDR corrected) were defined as visually responsive and used in all subsequent analyses.

In the main memory task we recovered the univariate BOLD time courses for all three distractor conditions by estimating the Hemodynamic Response Function (HRF) for each voxel at each time point of interest (0–19.5 seconds

from memory target onset). Estimated HRF's were averaged across all voxels in each ROI. Next, we reconstructed remembered and perceived orientations using an Inverted Encoding Model.

Effect(s) tested

For the mapping task in Experiment 1 (and one of the mapping tasks in Experiment 2) we had two within-subject conditions, comparing visual activation from a donut-shaped stimulus, to visual activation from a circle-shaped stimulus. We aimed to localize voxels for which donut > circle activity. For the second Experiment 2 mapping task we compared gratings versus fixation.

In the main memory task we had three within-subject conditions: While participants remembered a grating orientation, (1) there could be a Fourier filtered noise stimulus on the screen (or, in Experiment 2 this was a picture of a face or gazebo), (2) there could be another grating on the screen whose orientation was pseudo-random relative to the remembered orientation, or (3) the screen stayed blank. Our statistical analyses were within-subject repeated-measures ANOVA's of which the test statistics were based on permutation testing over 1000 permutations (see also Online Methods). We compared the three conditions.

Specify type of analysis: ☐ Whole brain ☒ ROI-based ☐ Both

Anatomical location(s)

Regions of interest (ROI's) were determined from the whole brain data based on standard retinotopic mapping procedures. We defined 9 a priori ROIs in early visual (V1–V3, V3AB, hV4) and parietal (IPS0–IPS3) cortex. These ROIs were combined across left and right hemispheres and across dorsal and ventral areas (for V2–V3) by concatenating voxels.

Furthermore, To identify voxels that were visually responsive, a General Linear Model (GLM) was performed on data from the mapping task using FSL FEAT (FMRI Expert Analysis Tool, version 6.00). Voxels defined as visually responsive were used in all subsequent analyses (with exception of Supplementary Fig. 12). Only visually responsive voxels were included in the ROI of each retinotopic area. We only included data for retinotopic areas in which the number of visually responsive voxels exceeded 20 for every single participant.

Statistic type for inference
(See Eklund et al. 2016)

Voxels in each ROI were included based on the significance of their response in the independent mapping task as determined by a t-test. We then used a multivariate method to assess the aggregate information across the entire set of voxels in each ROI during the main memory task. The results of this multivariate analysis were then subjected to repeated measures ANOVAs and the p-values were assessed using a non-parametric randomization procedure (see also the Online Methods).

Correction

Voxels in each ROI were included for further analysis if they passed an FDR corrected threshold of p = 0.05 based on data in the independent mapping task.

Models & analysis

n/a	Involved in the study
☒	☐ Functional and/or effective connectivity
☒	☐ Graph analysis
☐	☒ Multivariate modeling or predictive analysis

Multivariate modeling and predictive analysis

We used a multivariate Inverted Encoding Model (IEM) analysis. The first step in the IEM is estimating the encoding model. Each voxel differs with respect to the orientation it preferentially responds to, and this response profile is quantified by showing many orientations over many trials. Response profiles for each voxel are the weighted sum of 9 hypothetical orientation channels. The second step in the IEM is inverting the encoding model, to generate orientation reconstructions from voxel activity patterns. Channel weights represent each voxel's orientation selectivity, and when a new response is evoked, the combined selectivity of all voxels is used to reconstruct this new orientation from the voxel pattern. Reconstructions, in turn, were quantified by calculating a "fidelity metric". For each reconstruction, the fidelity metric was calculated by taking the channel response at each degree in orientation space (wrapped onto a 2πcircle), and projecting this vector onto the center of the stimulus space (i.e. onto zero degrees). The fidelity was taken as a proxy for the amount of information present. We quantified if this fidelity was >0 by running a resampling analysis, starting at the second step of the IEM, but using shuffled labels. This generated a resampled null-distribution of fidelities against which the empirical fidelities were compared.

Note: We also ran an additional analysis using a support vector machine - all of the main findings were confirmed using this more model-free approach.