Report

Current Biology

Parietal-driven visual working memory representation in occipito-temporal cortex

Highlights

- Visual working memory (VWM) content is present in multiple human cortical areas
- They include occipito-temporal cortex (OTC) and posterior parietal cortex (PPC)
- Their representations for real-world objects differ more in perception than in VWM
- This supports the role of feedback from PPC in shaping the content of VWM in OTC

Authors

Yaoda Xu

Correspondence

yaoda.xu@yale.edu

In brief

Using fMRI pattern decoding and representational similar analysis, Xu provides evidence supporting the role of feedback from the human posterior parietal cortex in shaping the content of visual working memory in occipito-temporal cortex for real-world visual objects.

Xu, 2023, Current Biology *33*, 4516–4523
October 23, 2023
https://doi.org/10.1016/j.cub.2023.08.080

CellPress

CellPress

Current Biology

Report

Parietal-driven visual working memory representation in occipito-temporal cortex

Yaoda Xu[1,2,*]
[1]Department of Psychology, Yale University, 100 College Street, New Haven, CT 06510, USA
[2]Lead contact
*Correspondence: yaoda.xu@yale.edu
https://doi.org/10.1016/j.cub.2023.08.080

SUMMARY

Human fMRI studies have documented extensively that the content of visual working memory (VWM) can be reliably decoded from fMRI voxel response patterns during the delay period in both the occipito-temporal cortex (OTC), including early visual areas (EVC), and the posterior parietal cortex (PPC).[1–4] Further work has revealed that VWM signal in OTC is largely sustained by feedback from associative areas such as prefrontal cortex (PFC) and PPC.[4–9] It is unclear, however, if feedback during VWM simply restores sensory representations initially formed in OTC or if it can reshape the representational content of OTC during VWM delay. Taking advantage of a recent finding showing that object representational geometry differs between OTC and PPC in perception,[10] here we find that, during VWM delay, the object representational geometry in OTC becomes more aligned with that of PPC during perception than with itself during perception. This finding supports the role of feedback in shaping the content of VWM in OTC, with the VWM content of OTC more determined by information retained in PPC than by the sensory information initially encoded in OTC.

RESULTS

To understand the impact of feedback on VWM representation, we need to directly compare the representational content across different brain regions and different VWM processing stages. This is not trivial, as different neuronal populations are involved in different brain regions and, even within the same region, the exact representational format may differ across VWM processing stages. For example, neural responses from the same population have been shown to rotate in the representational space or transform into a simpler code in VWM.[11–13] Such changes are adaptive to meet the demand of attentional control, resist distractor interference, and provide efficient coding, but do not necessarily change the informational content of VWM. To account for this, instead of directly comparing neural responses, here we evaluated the content of neural representation by applying representational similar analysis (RSA).[14] RSA measures how similar objects are represented with respect to each other through the computation of representational dissimilarity matrix (RDM). It is unit-free and can thus support direct comparisons of VWM contents independent of the exact neural code carrying them.

To examine how feedback may impact VWM representations in OTC, we also need to find a situation in which perceptual representation differs between OTC and associative areas such as PPC and test if such difference is maintained during VWM. In a recent human fMRI study with eight real-world basic-level object categories, we find robust category decoding during visual perception across areas in OTC and PPC[10,15,16] (Figures 1A and 1B). Critically, object RDMs during perception are distinct between OTC and PPC and there exists an information-driven two-pathway separation of brain regions in how they represent visual objects, with OTC regions arranging hierarchically along one pathway and PPC regions arranging along a second pathway[10,17] (Figure 1A). This is likely due to how visual features are coded in these different regions (see Vaziri-Pashkam and Xu[10] for a detailed discussion). This finding is replicated in ten independent datasets and across both natural and artificial basic-level object categories. This result provides us with a unique opportunity to understand the role of feedback in VWM: if feedback during VWM sustains the object representations formed during perception in OTC, then object RDMs should remain distinct between OTC and PPC in VWM as they do in perception; however, if feedback from associative areas such as PPC actively reshapes VWM representation in OTC, then OTC VWM representations may become more aligned with those of PPC during perception than with those of OTC during perception.

To test these possibilities, in this fMRI study, we asked 12 human participants to retain two target objects from the same basic-level category in each VWM trial (Figure 1C). After an extended delay period, participants judged whether the probe object shown at the end of the trial matched one of the target objects or was a new exemplar from the same category. As in our previous studies,[10,15,17] we used exemplars from eight basic-level object categories, including bodies, cars, cats, chairs, elephants, faces, houses and scissors, and with exemplars within a category varying in identity, pose and viewing angles (Figure 1A; see Figure S1A for the full set of images used). These basic-level categories were chosen to cover the range of real-world basic-level object categories in our visual environment and included some of the typical categories used in previous studies of object representations in OTC.[18,19] Behavioral performance accuracies

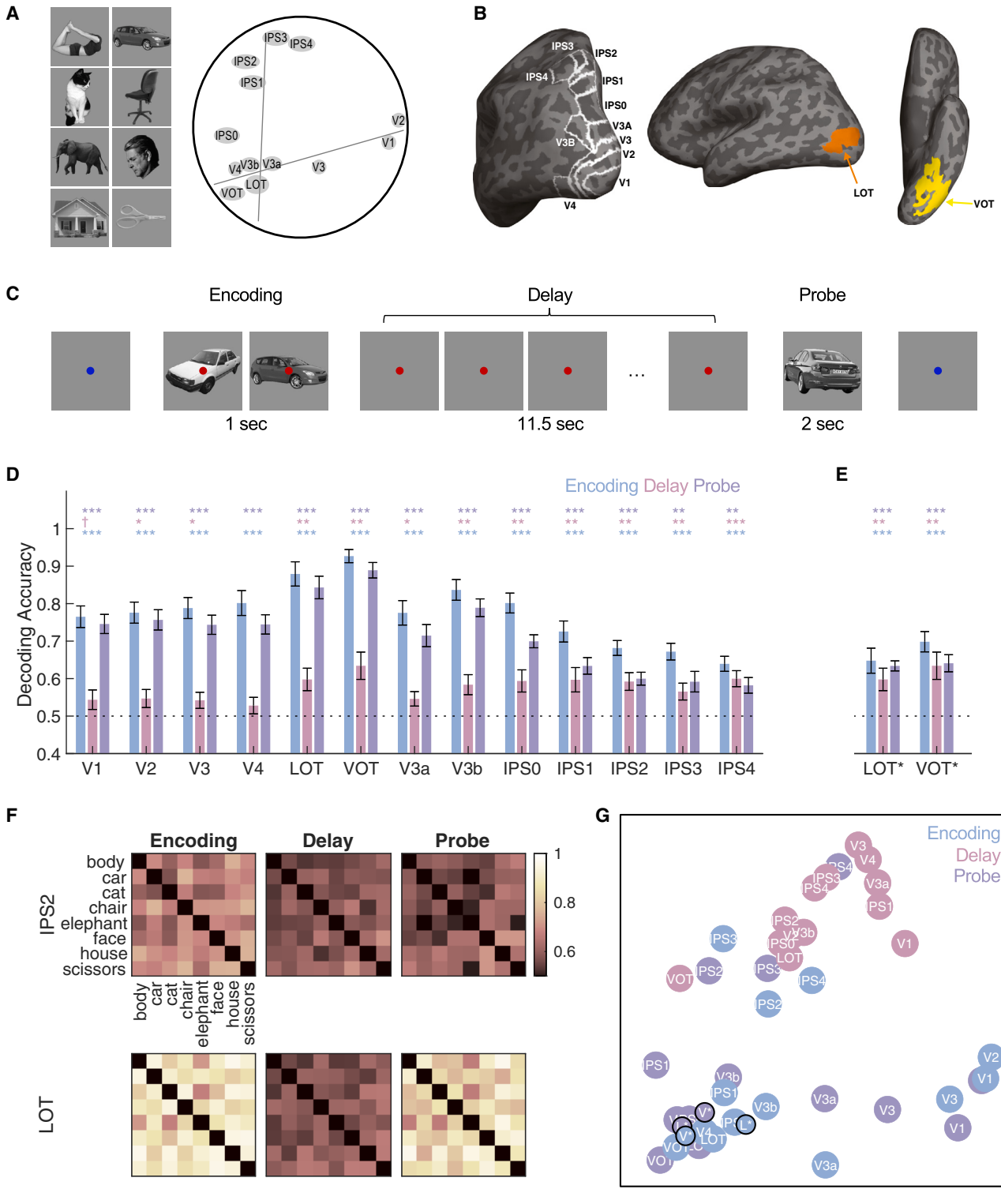

Figure 1. Experimental stimuli and VWM responses across OTC and PPC

(A) The eight real-world object categories used, and the information-driven two-pathway separation of brain regions as reported by Vaziri-Pashkam and Xu.[10]

(B) Example ROIs shown on the inflated cortical surfaces.

(C) An example trial showing the trial sequence with encoding, delay, and probe stages.

(D) Category decoding accuracy for all the ROIs during VWM encoding, delay, and probe stages. Asterisks above the bars mark decoding significance compared with chance (.5)—one-tailed and corrected.

(legend continued on next page)

were overall high, with above 90% accuracies for all the categories (Figure S1B). Responses from trials containing exemplars from the same category were averaged together to generate responses at the category level. These category responses were then examined in regions of interest (ROIs) in early visual areas V1 to V4, object shape processing regions in lateral occipito-temporal (LOT) and ventral occipito-temporal (VOT) cortex, and topographic areas along the intra-parietal sulcus (IPS) including V3a, V3b, and IPS0 to IPS4 (Figure 1B). fMRI responses were analyzed directly on the inflated cortical surface (vertices) rather than on the cortical volume (voxels) of each participant, as surface-based analysis has been shown to exhibit more sensitivity and better spatial selectivity[20,21] (see Figure S1C for the average number of vertices in each ROI).

From the averaged fMRI response time courses from each ROI (Figure S1D), we defined encoding, delay, and probe periods and averaged responses within each period for each ROI and each category (see STAR Methods). As reported in Figure 1D, all ROIs showed highly significant decoding during encoding and probe for these basic-level categories ($ts > 3.34$, $ps < .01$; one-tailed and corrected for multiple comparisons using the Benjamini-Hochberg method[22] for the three comparisons performed in each ROI; this applies to all subsequently reported stats of the same type; decoding for each time point is shown in Figure S2A). During VWM delay, category decoding was marginally above chance in V1 ($t(11) = 1.68$, $p = .061$), not above chance in V4 ($t(11) = 1.27$, $p = .12$), but was significantly above chance for all the other ROIs ($ts > 1.94$, $ps < .039$).

To understand how object representation may change across the different stages of the task, we performed cross-decoding across the different time points. Specifically, we trained a classifier at each time point and asked it to decode the responses at all time points. The results from all ROIs are shown in Figure S2B, with some notable cross-decoding drops between encoding/probe and delay periods across all the ROIs (a streamlined analysis of these results is reported later, see Figure 2C). A cross-decoding drop could be caused by a rotation in the representational space while preserving the representational structure, a change in the actual representational structure without a rotation, or both. To isolate and measure the change in the representation structure, we next performed RSA.

To evaluate the object representational geometry, from all the pairwise object category decoding accuracies we constructed an object-wise RDM separately for encoding, delay, and probe for each ROI of each participant (Figure 1F). We vectorized the off-diagonal elements of these RDMs and averaged them across participants. We next performed pairwise correlations among these averaged vectors from all the ROIs and all three stages of VWM processing to form a region-processing-wise RDM. Using multidimensional scaling (MDS), we projected this RDM onto a two-dimensional (2D) space, with the distances between brain regions at particular processing stages denoting their similarities in how they represent the eight object categories in the representational space (Figure 1G). Replicating our previous findings,[10,17] we saw the two-pathway separation among the ROIs during encoding and probe, with OTC and PPC regions arranging hierarchically along these two pathways, respectively. We also saw a high degree of overlap of the same ROI during encoding and probe. This is not surprising as both involved perceptual processing. We thus replicated our previous finding showing that, during perception, OTC and PPC differed in how they represent visual objects. Strikingly, however, this two-pathway separation disappeared during VWM delay, with all VWM ROIs clustered around higher perceptual PPC ROIs. Object representations thus became more similar between OTC and PPC during VWM delay than during perception, with OTC VWM representation becoming more aligned with those of PPC during perception than with those of OTC during perception.

Given the higher object category decoding accuracy obtained in OTC during VWM encoding and probe than during delay, it could be argued that fMRI decoding performance might have been saturated and reached ceiling performance during encoding and probe. For example, in the extreme case of "1" or "near 1" for each pair of category decoding, the intricate structure of the representational space would be distorted or lost. This would lead to distorted object-wise RDMs and distorted region-wise MDS projections during perception. In other words, the two-pathway structure we observed here could have been simply driven by the overall decoding performance of an ROI, with high-performing ROIs clustering around the ventral pathway and low-performing ones around the dorsal pathway, rather than the actual RDM differences between these ROIs. This is unlikely, however, as when decoding accuracy was much lower in OTC during perception in several cases in our previous study, we still obtained the same two-pathway separation.[10] To directly verify this, we need to lower the decoding performance during perception in OTC and see if a difference in RDM correlation between OTC and PPC still exists. To do so, instead of using peak encoding and probe responses for pattern decoding in LOT and VOT, we used instead the rising part of encoding and the falling part of probe responses as lower decoding performance was observed at these time points (see Figures S1D and S2A). Category decoding accuracies during these off-peak encoding and probe stages in LOT and VOT were both significant ($ts > 4.42$, $ps < .001$; one-tailed and corrected) and similar to those during delay in these regions and similar to those in the PPC regions across all three VWM stages (Figure 1E). Critically, in the region-processing-wise MDS projection plot, even with the lowered category decoding during encoding and probe, these off-peak LOT and VOT RDMs still landed right next to the corresponding peak LOT and VOT RDMs during encoding and probe, and away from these regions during delay (Figure 1G). The

(E) Category decoding accuracy for LOT and VOT during early encoding, delay, and late probe stages, with comparable decoding across the three stages. Asterisks mark decoding significance as in (D).

(F) Example object-wise RDMs for two ROIs for each of the three VWM stages.

(G) MDS projection of the region-processing-wise RDM showing the similarity of the ROIs in how they represent the eight categories of objects at the three VWM stages. V* and L* denote VOT and LOT at early encoding and late probe stages, respectively.

Error bars indicate SE. †$p < .10$, *$p < .05$, ** $.01 < p < .001$, ***$p < .001$.

See also Figures S1 and S2.

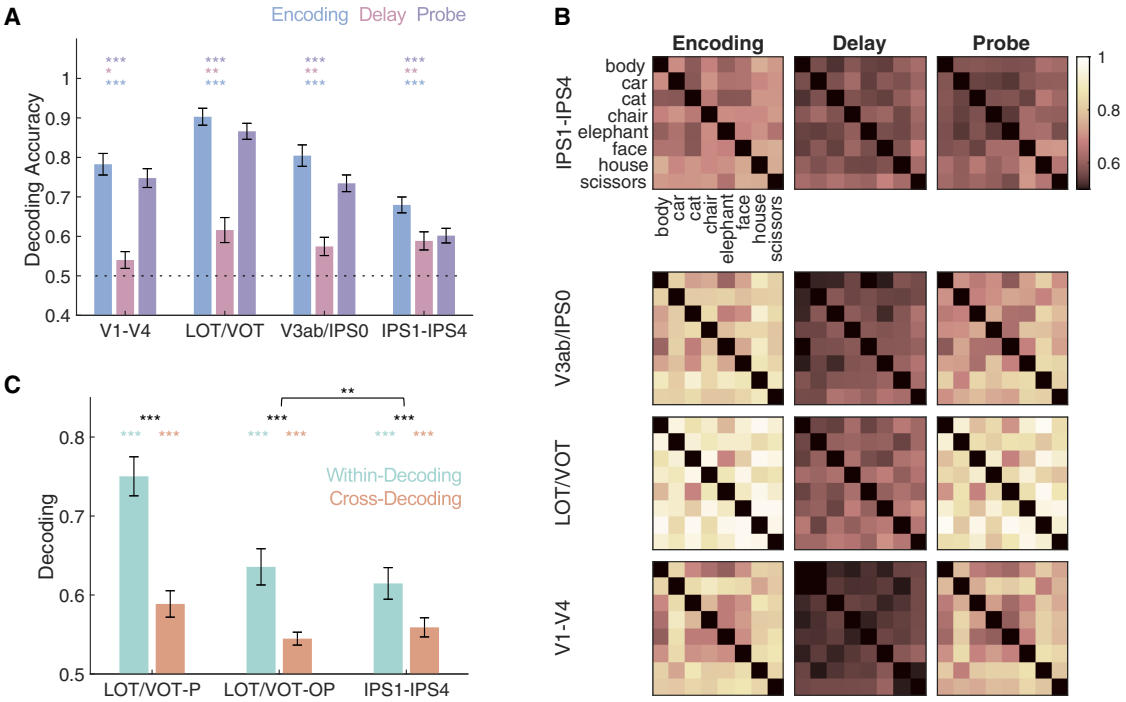

Figure 2. Results of the four ROI sectors (averaged across the ROIs in each sector)

(A) Category decoding accuracy for the four sectors during the three VWM stages. Asterisks above the bars mark decoding significance compared with chance (.5)—one-tailed and corrected.

(B) Object-wise RDMs for the four sectors for each of the three VWM stages.

(C) Within- and cross-decoding between perception (averaged between encoding and probe) and VWM (delay) for LOT/VOT and IPS1–IPS4. Colored asterisks above the bars mark decoding significance compared with chance (.5)—one-tailed and corrected. Black asterisks above the bars mark the significance of the cross-decoding drop—one-tailed and corrected. Black asterisks bridging between the sectors mark the significance of the difference in cross-decoding drop—two-tailed.

LOT/VOT-P, peak perception responses from LOT/VOT; LOT/VOT-OP, off-peak perception responses from LOT/VOT.

Error bars indicate SE. †p < .10, *p < .05, ** .01 < p < .001, ***p < .001.

See also Figure S2.

observed change in OTC RDM between perception and VWM delay therefore could not be attributed to saturation in the fMRI decoding performance.

To quantify the above observation and to streamline the analysis, based on the overall decoding performance and projections in the representational space (Figures 1D and 1G), we averaged responses from different ROIs to create four sectors: V1–V4, LOT/VOT, V3a/V3b/IPS0, and IPS1–IPS4. Decoding performance across all three VWM stages and all four sectors was significantly above chance (*ts* > 1.88, *ps* < .044; one-tailed and corrected; see Figures 2A and 2B for the decoding results and the RDMs). Given that responses from encoding and probe showed similar decoding performance and RDMs, we further averaged results involving encoding and probe and labeled it as perception. We focused our analyses on LOT/VOT and IPS1–IPS4, as V1–V4 showed overall weak VWM decoding and RDMs from all regions were clustered around higher IPS ROIs during VWM delay.

Cross-decoding between perception and VWM for LOT/VOT peak and off-peak responses and for IPS1–IPS4 revealed significant above-chance cross-decoding between perception and VWM in both regions (*ts* > 4.87, *ps* < .001, one-tailed and corrected; Figure 2C), but also a significant drop between within- and cross-decoding (*ts* > 5.08, *ps* < .001, one-tailed and corrected). In the LOT/VOT off-peak case, although the within-decoding performance was similar to that of IPS1–IPS4 ($t(11)$ = 1.88, p = .088, two-tailed), the drop in cross-decoding was significantly greater in LOT/VOT than in IPSI–PS4 ($t(11)$ = 3.20, p < .01, two-tailed). There were thus significant transformations in object representation between perception and VWM in both the ventral and dorsal regions, with the magnitude of the transformation being greater in the ventral than dorsal regions.

To understand the nature of the representational transformation between perception and VWM, we next quantified the RDM correlation results. Consistent with the region-processing-wise MDS plot (Figure 1G), we found that RDM correlation between LOT/VOT and IPS1–IPS4 was greater during VWM delay than during perception for both peak and off-peak LOT/VOT encoding and probe responses ($t(11)$ = 2.22, p = .032, and $t(11)$ = 2.16, p = .036 for peak and off-peak responses, respectively; one-tailed and corrected for a total of four comparisons made here, separately for peak and off-peak responses; this applies to all subsequently reported stats of the same type; Figure 3A). In fact, RDM correlation between perception and VWM within a region was greater in IPS1–IPS4 than LOT/VOT ($t(11)$ = 3.15, p = .0092, and $t(11)$ = 2.70, p = .021,

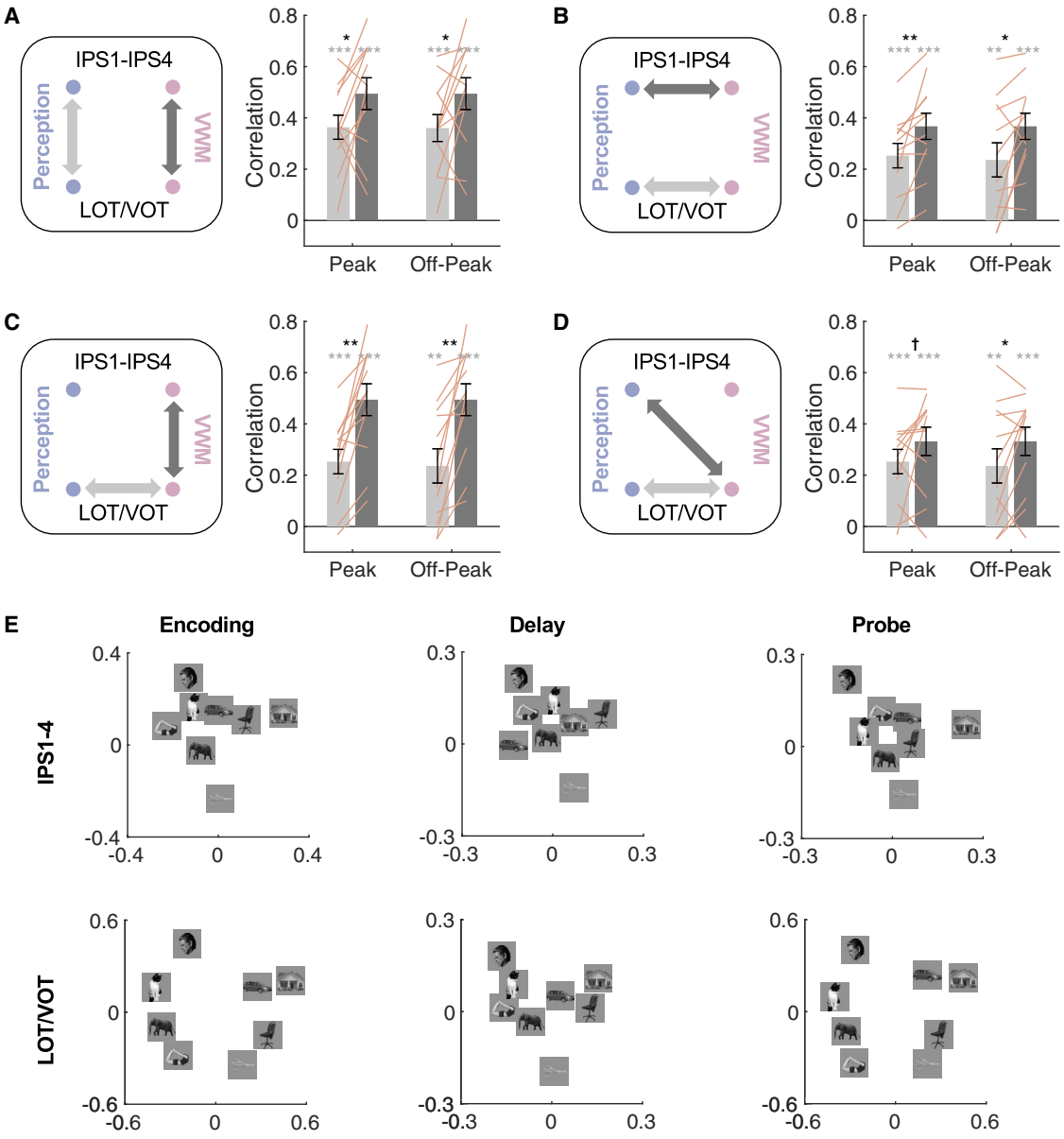

Figure 3. RDM correlations

(A) RDM correlation between LOT/VOT and IPS1–IPS4 during perception vs. during VWM.

(B) RDM correlation between perception and VWM within LOT/VOT vs. within IPS1–IPS4.

(C) RDM correlation between perception and VWM within LOT/VOT vs. RDM correlation between LOT/VOT and IPS1–IPS4 during VWM.

(D) RDM correlation between perception and VWM within LOT/VOT vs. RDM correlation between LOT/VOT during VWM and IPS1–IPS4 during perception.

Across (A) to (D), Peak and Off-Peak are results from RDMs extracted from peak and off-peak encoding and probe responses from LOT/VOT, respectively. Gray asterisks mark correlation significance compared with 0 and black asterisks mark pairwise correlation comparisons—one-tailed and corrected. Orange lines indicate responses from individual participants.

(E) MDS projections of the object-wise RDM showing the representational similarity of the eight object categories for each of the three VWM stages in LOT/VOT and IPS1–IPS4. These projections are manually rotated with faces always shown at the top left corner. They show a U-shaped object arrangement in LOT/VOT during encoding and probe, but a T-shaped object arrangement in LOT/VOT during delay and in IPS1–IPS4 during all three VWM stages.

Error bars indicate SE. †p < .10, *p < .05, ** .01 < p < .001, ***p < .001.

See also Figure S3.

respectively, for peak and off-peak responses; Figure 3B). Object representations thus appear to undergo a greater amount of change in OTC than in PPC between perception and VWM delay. Further comparisons revealed that LOT/VOT RDM during VWM delay was more correlated with that of IPS1–IPS4 during both VWM and perception than with itself during perception (comparison with IPS1–IPS4 during VWM, $t(11) = 4.57$, $p = .016$, and $t(11) = 4.15$, $p = .0032$, respectively, for peak and

off-peak responses; comparison with IPS1–IPS4 during perception, $t(11) = 1.58$, $p = .074$, and $t(11) = 1.89$, $p = .043$, respectively, for peak and off-peak responses; Figures 3C and 3D). Similar, albeit slightly weaker, results were obtained when an earlier delay period was examined (Figures S3A–S3D; see STAR Methods for details). This was largely due to an increased RDM correlation between perception and VWM in LOT/VOT in the peak responses (likely resulting from the presence of lingering encoding responses due to the hemodynamic response lag), with all four correlation comparisons remaining significant for LOT/VOT off-peak responses. The same results were not obtained, however, when the cells in the RDMs were randomly shuffled (Figures S3A–S3D), indicating that the correlation results obtained reflected structure in the data and not noise.

Overall, these results showed that not only did object representations become more similar between OTC and PPC during VWM delay than perception but also that those of OTC during VWM delay became more aligned with the representations of PPC during perception. This can be directly seen when we projected the object-wise RDMs of LOT/VOT and IPS1–IPS4 for each VWM stage onto separate 2D spaces using MDS (Figure 3E). These projections show a U-shaped object arrangement in LOT/VOT during encoding and probe, but a T-shaped object arrangement in LOT/VOT during delay and in IPS1–IPS4 for all three VWM stages.

DISCUSSION

Taking advantage of a recent finding showing that object representational structure (i.e., RDM) during perception differs between OTC and PPC, here we report that this difference decreases during VWM, with OTC object representations during VWM becoming more aligned with those of PPC during perception than with OTC itself during perception. Given the previously documented role of feedback in driving VWM representation in sensory regions,[5–7] our results support the view that feedback does not only sustain VWM signals in OTC but can also actively shape the content of VWM there, more so than the initial sensory input in OTC.

The present finding echoes a recent direct neurophysiological study demonstrating that attentional feedback signals can influence sensory coding in macaque visual cortex.[23] Using fMRI and the RSA approach, here we provide evidence supporting a role of feedback in actively shaping the content of VWM in OTC in the human brain. Previous research has reported that the neural code for VWM may rotate to provide better distractor resistance or meet attentional selection priority[11,12] and may be recoded into an abstract and more efficient code.[13] VWM and perception may also engage different neuronal populations.[7,9] Indeed, using the cross-decoding approach (in which a decoder trained to differentiate the neural responses of a pair of objects at one VWM processing stage is asked to do so at another stage), we also found a transformation in object representation between perception and VWM in both the ventral and dorsal regions, with the magnitude of the transformation being greater in the ventral than dorsal regions. Such a cross-decoding drop between perception and VWM has also been noted by others in both the ventral and dorsal regions.[24,25] However, cross-decoding drop alone cannot inform us of the nature of the transformation: it could be caused by a rotation in the representational space while preserving the representational structure, a change in the actual representational structure without a rotation, or both. By applying the RSA approach, here we are able to isolate and measure the change in the representational structure (i.e., representational content) independent of the rotation in the representational space and provide evidence supporting the role of feedback in shaping the representational content of OTC during VWM delay. This makes RSA a useful tool in unveiling the content of VWM within and across brain regions, even when the exact format of VWM representation may change.

Further research is needed to understand if PPC uniquely shapes the VWM content of OTC, or if both regions may be influenced by a third region, such as PFC. However, existing neurophysiological evidence as reviewed previously[4] indicates that VWM delay period signals in PPC are not triggered by feedback originating from PFC. For example, widespread, task-dependent, and content-specific frontoparietal synchronization during VWM delay period has been shown to exhibit an asymmetry, with information transmission appearing to be governed more by influences arising in PPC than those from PFC.[26] Similarly, stronger, more reliable, and shorter latency VWM category signals during the delay period are found in PPC than those in PFC,[27] and that neuronal activity correlating with behavioral VWM performance appeared much earlier in PPC than in PFC.[28] More neurophysiology and lesion studies are needed to fully establish the casual connection between the different brain regions in determining the content of VWM representation.

Previous research has shown that object representations in PPC are more task-driven than those in OTC.[15,29–31] This suggests that while OTC automatically encodes a large set of object information during perception, given that only a subset of this information is relevant for a VWM task, information encoded into PPC and retained in VWM would be more selective and task-relevant. Consequently, although the RDMs of the OTC during perception will be determined by a greater set of features automatically extracted during perception, those of PPC during perception and VWM and those of OTC during VWM will more closely track the task-relevant feature encoded and retained. This makes an interesting prediction that item interference in VWM for these natural objects will be determined more by their representational similarity in PPC than in OTC. It will be important for future research to test this prediction and establish how feedback from associative regions in the brain may directly impact behavioral performance, and how this may be further modulated by whether real-world objects or simple visual features are retained.

In the present study, basic-level object category can also be decoded in EVC during perception and VWM. This may not be surprising, as the seminal work by Rosch et al.[32] showed that exemplars of the same basic-level category share a common set of visual features. Indeed, many basic-level categories can be differentiated based on low-level features even with an exemplar identity and viewpoint change, such as curvature and the presence of unique features like limbs. Rice et al.[33] quantified this and found that even for exemplars varying in identity and viewpoint as in the present study, the low-level image similarity between exemplars was still higher within the same categories

than between different categories. Moreover, fMRI response pattern correlation across the entire ventral visual cortex, including EVC, tracked the low-level visual feature similarity across the categories. Weaker but significant category effects were also found in EVC in another study and were believed to be driven by the presence of low-level visual similarities among the exemplars of the same category.[19] The shared low-level visual features among the exemplars within the same category thus likely drove category decoding in EVC in the present study and in our previous studies.[10,15,16]

In an effort to precisely control feature changes, previous studies have predominantly investigated VWM for low-level visual features (e.g., orientation, color, motion, and spatial position).[24,34–37] In everyday life, however, we rarely hold on to such stimuli in VWM; but rather, it is the complex real-world objects from our natural environment that we need to retain to guide our thoughts and actions. Thus, it is important to study these objects directly if we want to ultimately understand the neural underpinning of VWM supporting real-world visual cognition. The present study not only contributes to such a research effort, but also shows that the richness of real-world objects can afford us with unique opportunities to ask and answer new questions. Given the simplicity of low-level visual stimuli, similar representations are likely formed across brain regions, preventing us from examining the role of feedback in determining the content of VWM. With complex real-world objects and with PPC and OTC showing differential representational geometry for these objects, we are able to examine here how feedback may impact the content of VWM in OTC. Studying VWM of real-world objects thus not only unveils the specific neural mechanisms involved in processing such objects but can also enrich our understanding of the general mechanisms supporting VWM in the human brain.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Main experiment
  - Localizer experiments
  - MRI method
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - ROI definitions
  - VWM decoding analysis
  - RDM analysis
  - Statistical analyses

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2023.08.080.

ACKNOWLEDGMENTS

This research was supported by National Institutes of Health grants (1R01EY030854) to Y.X. I thank Hillary Nguyen for her assistance in fMRI data collection and Marvin Chun for his general support.

AUTHOR CONTRIBUTIONS

Y.X. conceived and designed the study, collected and analyzed all data, and wrote the manuscript.

DECLARATION OF INTERESTS

The author declares no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: April 6, 2023
Revised: July 24, 2023
Accepted: August 25, 2023
Published: September 22, 2023

REFERENCES

1. D'Esposito, M., and Postle, B.R. (2015). The cognitive neuroscience of working memory. Annu. Rev. Psychol. *66*, 115–142.
2. Serences, J.T. (2016). Neural mechanisms of information storage in visual short-term memory. Vis. Res. *128*, 53–67.
3. Christophel, T.B., Klink, P.C., Spitzer, B., Roelfsema, P.R., and Haynes, J.D. (2017). The distributed nature of working memory. Trends Cognit. Sci. *21*, 111–124.
4. Xu, Y. (2017). Reevaluating the sensory account of visual working memory storage. Trends Cognit. Sci. *21*, 794–815.
5. Liebe, S., Hoerzer, G.M., Logothetis, N.K., and Rainer, G. (2012). Theta coupling between V4 and prefrontal cortex predicts visual short-term memory performance. Nat. Neurosci. *15*. 456-62.
6. Mendoza-Halliday, D., Torres, S., and Martinez-Trujillo, J.C. (2014). Sharp emergence of feature-selective sustained activity along the dorsal visual pathway. Nat. Neurosci. *17*, 1255–1262.
7. van Kerkoerle, T., Self, M.W., and Roelfsema, P.R. (2017). Layer-specificity in the effects of attention and working memory on activity in primary visual cortex. Nat. Commun. *8*, 13804.
8. Leavitt, M.L., Mendoza-Halliday, D., and Martinez-Trujillo, J.C. (2017). Sustained activity encoding working memories: not fully distributed. Trends Neurosci. *40*, 328–346.
9. Lawrence, S.J.D., van Mourik, T., Kok, P., Koopmans, P.J., Norris, D.G., and de Lange, F.P. (2018). Laminar organization of working memory signals in human visual cortex. Curr. Biol. *28*, 3435–3440.e4.
10. Vaziri-Pashkam, M., and Xu, Y. (2019). An information-driven 2-pathway characterization of occipitotemporal and posterior parietal visual object representations. Cerebr. Cortex *29*, 2034–2050.
11. Parthasarathy, A., Herikstad, R., Bong, J.H., Medina, F.S., Libedinsky, C., and Yen, S.C. (2017). Mixed selectivity morphs population codes in prefrontal cortex. Nat. Neurosci. *20*, 1770–1779.
12. Yu, Q., Teng, C., and Postle, B.R. (2020). Different states of priority recruit different neural codes in visual working memory. PLoS Biol. *18*, e3000769.
13. Kwak, Y., and Curtis, C.E. (2022). Unveiling the abstract format of mnemonic representations. Neuron *110*, 1822–1828.e5.
14. Kriegeskorte, N., and Kievit, R.A. (2013). Representational geometry: integrating cognition, computation, and the brain. Trends Cognit. Sci. *17*, 401–412.

15. Vaziri-Pashkam, M., and Xu, Y. (2017). Goal-directed visual processing differentially impacts human ventral and dorsal visual representations. J. Neurosci. *37*, 8767–8782.
16. Vaziri-Pashkam, M., Taylor, J., and Xu, Y. (2019). Spatial frequency tolerant visual object representations in the human ventral and dorsal visual processing pathways. J. Cognit. Neurosci. *31*, 49–63.
17. Xu, Y., and Vaziri-Pashkam, M. (2019). Task modulation of the 2-pathway characterization of occipitotemporal and posterior parietal visual object representations. Neuropsychologia *132*, 107140.
18. Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., and Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science *293*, 2425–2430.
19. Kriegeskorte, N., Mur, M., Ruff, D.A., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K., and Bandettini, P.A. (2008). Matching categorical object representations in inferior temporal cortex of man and monkey. Neuron *60*, 1126–1141.
20. Oosterhof, N.N., Wiestler, T., Downing, P.E., and Diedrichsen, J. (2011). A comparison of volume-based and surface-based multi-voxel pattern analysis. Neuroimage *56*, 593–600.
21. Brodoehl, S., Gaser, C., Dahnke, R., Witte, O.W., and Klingner, C.M. (2020). Surface-based analysis increases the specificity of cortical activation patterns and connectivity results. Sci. Rep. *10*, 5737.
22. Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Stat. Soc. B *57*, 289–300.
23. Debes, S.R., and Dragoi, V. (2023). Suppressing feedback signals to visual cortex abolishes attentional modulation. Science *379*, 468–473.
24. Harrison, S.A., and Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. Nature *458*, 632–635.
25. Rademaker, R.L., Chunharas, C., and Serences, J.T. (2019). Coexisting representations of sensory and mnemonic information in human visual cortex. Nat. Neurosci. *22*, 1336–1344.
26. Salazar, R.F., Dotson, N.M., Bressler, S.L., and Gray, C.M. (2012). Content-specific fronto-parietal synchronization during visual working memory. Science *338*, 1097–1100.
27. Swaminathan, S.K., and Freedman, D.J. (2012). Preferential encoding of visual categories in parietal cortex compared with prefrontal cortex. Nat. Neurosci. *15*, 315–320.
28. Jacob, S.N., and Nieder, A. (2014). Complementary roles for primate frontal and parietal cortex in guarding working memory from distractor stimuli. Neuron *83*, 226–237.
29. Bracci, S., Daniels, N., and Op de Beeck, H. (2017). Task context overrules object- and category-related representational content in the human parietal cortex. Cerebr. Cortex *27*, 310–321.
30. Xu, Y. (2018). A tale of two visual systems: Invariant and adaptive visual information representations in the primate brain. Annu. Rev. Vis. Sci. *4*, 311–336.
31. Xu, Y. (2018). The posterior parietal cortex in adaptive visual processing. Trends Neurosci. *41*, 806–822.
32. Rosch, E., Mervis, C.B., Gray, W.D., Johnson, D.M., and Boyes-Braem, P. (1976). Basic objects in natural categories. Cognit. Psychol. *8*, 382–439.
33. Rice, G.E., Watson, D.M., Hartley, T., and Andrews, T.J. (2014). Low-level image properties of visual objects predict patterns of neural response across category-selective regions of the ventral visual pathway. J. Neurosci. *34*, 8837–8844.
34. Serences, J.T., Ester, E.F., Vogel, E.K., and Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. Psychol. Sci. *20*, 207–214.
35. Bettencourt, K.C., and Xu, Y. (2016). Decoding under distraction reveals distinct occipital and parietal contributions to visual short-term memory representation. Nat. Neurosci. *19*, 150–157.
36. Sprague, T.C., Ester, E.F., and Serences, J.T. (2016). Restoring latent visual working memory representations in human cortex. Neuron *91*, 694–707.
37. Yu, Q., and Shim, W.M. (2017). Occipital, parietal, and frontal cortices selectively maintain task-relevant features of multi-feature objects in visual working memory. Neuroimage *157*, 97–107.
38. Chang, C.C., and Lin, C.J. (2011). LIBSVM: a library for support vector machines. Acm. T. Intel. Syst. Tec. *2*, 1–27.
39. Sereno, M.I., Dale, A.M., Reppas, J.B., Kwong, K.K., Belliveau, J.W., Brady, T.J., Rosen, B.R., and Tootell, R.B. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science *268*, 889–893.
40. Swisher, J.D., Halko, M.A., Merabet, L.B., McMains, S.A., and Somers, D.C. (2007). Visual topography of human intraparietal sulcus. J. Neurosci. *27*, 5326–5337.
41. Kourtzi, Z., and Kanwisher, N. (2000). Cortical regions involved in perceiving object shape. J. Neurosci. *20*, 3310–3318.
42. Jeong, S.K., and Xu, Y. (2017). Task-context dependent linear representation of multiple visual objects in human parietal cortex. J. Cognit. Neurosci. *29*, 1778–1789.
43. Taylor, J., and Xu, Y. (2022). Representation of color, form, and their conjunction across the human ventral visual pathway. Neuroimage *251*, 118941.
44. Brainard, D.H. (1997). The psychophysics toolbox. Spatial Vis. *10*, 433–436.
45. Dale, A.M., Fischl, B., and Sereno, M.I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage *9*, 179–194.
46. Bettencourt, K.C., and Xu, Y. (2013). The role of transverse occipital sulcus in scene perception and its relationship to object individuation in inferior intraparietal sulcus. J. Cognit. Neurosci. *25*, 1711–1722.
47. Bettencourt, K.C., and Xu, Y. (2016). Understanding location- and feature-based processing along the human intraparietal sulcus. J. Neurophysiol. *116*, 1488–1497.
48. Malach, R., Reppas, J.B., Benson, R.R., Kwong, K.K., Jiang, H., Kennedy, W.A., Ledden, P.J., Brady, T.J., Rosen, B.R., and Tootell, R.B. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc. Natl. Acad. Sci. USA. *92*, 8135–8139.
49. Grill-Spector, K., Kushnir, T., Edelman, S., Itzchak, Y., and Malach, R. (1998). Cue-invariant activation in object-related areas of the human occipital lobe. Neuron *21*, 191–202.
50. Dale, A.M., and Buckner, R.L. (1997). Selective averaging of rapidly presented individual trials using fMRI. Hum. Brain Mapp. *5*, 329–340.
51. Shepard, R.N. (1980). Multidimensional scaling, tree-fitting, and clustering. Science *210*, 390–398.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Processed fMRI data	This paper	https://osf.io/p7m6v/
Software and algorithms		
MATLAB 2019B	MathWorks, Inc.	https://www.mathworks.com/
Analysis code	This paper	https://osf.io/p7m6v/
Experiment code	This paper	https://osf.io/p7m6v/
PsychToolBox	PsychToolBox	http://psychtoolbox.org/
FreeSurfer	FreeSurfer	http://surfer.nmr.mgh.harvard.edu
LibSVM	Chang and Lin[38]	https://www.csie.ntu.edu.tw/~cjlin/libsvm/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Yaoda Xu (yaoda.xu@yale.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- De-identified human data (behavior, processed fMRI results) have been deposited at OSF. They are publicly available as of the date of publication. DOIs are listed in the key resources table.
- All original code has been deposited at OSF and is publicly available as of the date of publication. DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this work paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Twelve (seven females) healthy human participants with normal or corrected to normal visual acuity, all right-handed, and aged between 18 and 35 took part in the experiment. All participants gave their informed consent prior to the experiment and received payment for their participation. The experiment was approved by the Committee on the Use of Human Subjects at Yale University.

METHOD DETAILS

Main experiment

This experiment used the same eight real-world basic-level object category exemplar images as in a prior study.[10] These categories included bodies, cars, cats, chairs, elephants, faces, houses, and scissors, and were some of the typical basic-level categories used in previous investigations of object representation in human OTC[18,19] (Figure 1A). For each object category, there were ten exemplar images that varied in identity, pose and viewing angle to minimize the low-level similarities among them (see Figure S1A for the complete set of images used). All images were placed on a gray square (subtended 9.73 ° × 9.73 °) and shown on a larger background of the same gray color. Each VWM trial contained a sequential central presentation of two target object images from the same category, a prolonged blank delay, and a probe object (Figure 1C). The probe object was either an exact match of one of the target object images or a different exemplar from the same category. Each trial was 15 s long, with the timing of the different events as follows: fixation (.5 s) – in the form of a looming red dot to alert the participants the imminent presentation of the target object images, first target object image (.5 s), second target object image (.5 s), blank delay with a red fixation dot (11.5 s), and probe object image (2 s). There were a total of 16 trials in each run, two for each category. Each run started and ended with an 8-s blank period with a blue fixation dot. Successive VWM trials were sandwiched by a blank period with a blue fixation dot. Of the 15 such inter-trial blank periods, 3 were 8-s

long and 12 were 2-s long, and they were randomly distributed. Fourteen runs of data were collected from each participant, with each run lasting 5 min 4 s.

Behavioral performance accuracies for the VWM trials are reported in Figure S1B. Accuracies were overall high, with above 90% accuracies for all the categories. Accuracies were higher for some than other categories ($F(7,88)$ = 3.09, p = .0059), likely driven by greater dissimilarity among the exemplars in some of the categories and/or due to category specific object coding for categories such as faces. Note that accuracy was a result of decision making at the probe stage; the same information could be retained in VWM and produced either a correct or incorrect response depending on the similarity between the target and probe items. Since we averaged responses from trials of the same category and examined responses at the category level rather than at the individual exemplar level, even if representations for the specific exemplars were poor, as long as some share category information was retained in VWM, we should still be able to differentiate the different categories based on the averaged fMRI response patterns. We have additionally z-normalized fMRI response patterns before further analysis to remove fMRI response amplitude differences across the different categories and VWM processing stages (see later).

Localizer experiments

Topographic visual regions

These regions were mapped with flashing checkerboards using standard techniques[39,40] with parameters optimized following Swisher et al.[40] to reveal maps in parietal cortex. Specifically, a polar angle wedge with an arc of 72° swept across the entire screen (19.07 × 13.54° of visual angle). The wedge had a sweep period of 32 s, flashed at 4 Hz, and swept for 8 cycles in each run (for more details, see Swisher et al.[40]). Participants completed 4 to 6 runs, each lasting 4 min 36 s. All participants were asked to detect a dimming that could occur anywhere within the polar angle wedge, thereby ensuring attention to the whole wedge.

Lateral and ventral occipito-temporal regions (VOT and LOT)

To identify LOT and VOT ROIs, following Kourtzi and Kanwisher[41] and as we have done previously,[10,15,16,42,43] participants viewed blocks of object and scrambled object (all subtended approximately 9.73 ° × 9.73 °). The images were photographs of gray-scaled common objects (e.g., cars, tools, and chairs) and phase-scrambled versions of these objects. Participants monitored a slight spatial jitter which occurred randomly once in every 10 images. Each run contained 4 blocks of each of objects, phase-scrambled objects, and two other conditions that were used to define another brain region. Each block lasted 16s and contained 20 unique images, with each appearing for 750 ms and followed by a 50 ms blank display. Besides the stimulus blocks, 8-s fixation blocks were included at the beginning, middle, and end of each run. Each participant was tested with 2 runs, each lasting 4 min 40 s.

MRI method

Each participant completed an experimental session (1.5 h) and a localizer session (1.5 h) containing topographic mapping and functional localizers. MRI data were collected using a Siemens Prisma 3T scanner, with a 32-channel receiver array head coil. Participants lay on their backs inside the scanner and viewed the back-projected display through an angled mirror mounted inside the headcoil. The display was projected using an LCD projector at a refresh rate of 60 Hz and a spatial resolution of 1280 × 1024. An Apple Macbook Pro laptop was used to create the stimuli and collect motor responses. Stimuli were created using MATLAB and Psychtoolbox.[44]

A high-resolution T1-weighted structural image (0.8 × 0.8 × 0.8 mm) was obtained from each participant for surface reconstruction. All blood-oxygen-level-dependent (BOLD) data were collected via a $T2^*$- weighted echo-planar imaging pulse sequence that employed multiband RF pulses and simultaneous multi-slice (SMS) acquisition. For both the main experiment and the localizers, 72 axial slices (2 mm isotropic), 0 skip, covering the entire brain were collected (TR = 800 ms, TE = 37 ms, flip angle = 52°, SMS factor = 8).

QUANTIFICATION AND STATISTICAL ANALYSIS

FMRI data were analyzed using FreeSurfer (http://surfer.nmr.mgh.harvard.edu), FsFast,[45] and in-house MATLAB codes. LibSVM software[38] was used for the MVPA support vector machine (SVM) analysis. FMRI data preprocessing included 3D motion correction and linear and quadratic trend removal. After reconstructing the inflated 3D cortical surface of each participant using the high-resolution anatomical data, we projected the fMRI data of that participant onto their native cortical surface. All fMRI responses were analyzed directly on the inflated cortical surface (vertices) rather than on the cortical volume (voxels) of each participant, including ROI definition and the main VWM analysis, as surface-based analysis has been shown to exhibit more sensitivity and better spatial selectivity.[20,21]

ROI definitions

Following the detailed procedures described in Swisher et al.[40] and as was done in our prior publications,[10,15,16,35,46,47] by examining phase reversals in the polar angle maps, we identified areas V1 to V4, V3a, V3b, and IPS0 to IPS4 in each participant (Figure 1B). Following Kourtzi and Kanwisher[41] and as was done in our prior studies,[10,15,16,42,43] LOT and VOT were defined as a cluster of continuous voxels in the lateral and ventral occipital cortex, respectively, that responded more to the intact than to the scrambled object images (Figure 1B). LOT and VOT loosely corresponded to the location of LO and pFs[41,48,49] but extending further into the temporal cortex in an effort to capture the continuous activations often seen extending into the ventral temporal cortex. The average number of vertices for each ROI is reported in Figure S1C.

VWM decoding analysis

With the length of our VWM trials being 15 s, for each surface vertex, we estimated the fMRI response amplitude at each TR from the onset of the trial up to 24 s, totaling 30 TRs (with each TR being 800 ms). This was done separately for the trials in each object category. To obtain these estimates, we first constructed 30 finite impulse response (FIR) functions corresponding to each TR of each trial. As each category only appeared twice in a run and with adjacent trials overlapping with each other, to obtain accurate response estimates, we combined three runs together, as detailed below, before applying GLM modeling to solve the beta weight for each of the 240 FIR functions (30 TRs x 8 categories). Because the trial onset times were jittered with respect to TR onsets, trial onset times were rounded to the nearest TR before GLM modeling. To obtain independent training and testing data for pattern decoding, we split the runs into odd and even halves, each containing 7 runs. For each half of the data, we applied an GLM to each 3-run combination from the 7 runs of each half of the data. This resulted in 35 beta weight estimates for each FIR function in each half of the data for each surface vertex. The beta weights of all the vertices in a given ROI formed our fMRI response pattern for that ROI. For each ROI, we thus had 35 patterns for each TR and each category for each half of the data. Note that these 35 patterns were not all independent from each other as some of them were estimated from shared runs; nonetheless, because some of the patterns were derived from nonoverlapping runs, when we used all 35 patterns from each condition to train an SVM decoder, we still had some protection against data overfitting. As detailed later, pattern decoding training was done within the 35 patterns in one-half of the data and then tested on the 35 patterns in the other half and vice versa. Training and testing were thus performed on independent data.

To generate a response amplitude time course from each ROI for each category, we averaged all the beta weights across all the surface vertices within an ROI and from both halves of the data. The results are shown in Figure S1D. Based on the peak responses from all the ROIs and categories, we defined an encoding period (from 4.8 to 8 s), a delay period (from 12 to 15.2 s) and a probe period (from 17.6 to 20.8 s). We then averaged the four beta weights within each period to generate an averaged response for each of these three VWM processing stages.

Note that to minimize signal contamination from the encoding period and given that encoding-related fMRI responses may still be present 10–12 s post stimulus onset,[50] in the present study the delay period was defined as the flat period from 12 to 15.2 s right before the signal rose again with the presentation of the probe (see Figures S1D and S2A). It thus included the lowest and stable time points of the BOLD responses and category decoding before the rise of both following probe presentation. Although this period partially overlapped with probe presentation at 13 s, because neither response amplitude nor category decoding rose till a little over 2 s after the onset of the probe (see the responses from LOT and VOT in Figures S1D and S2A) and the same was also true during the encoding period, it is reasonable to assume that the signals measured during this period largely came from delay period activity. If we examine an earlier delay period with no overlap with probe presentation (from 11.2 to 12.8 s), very similar but slightly weaker results were obtained (see later), consistent with potential signal contamination from encoding.

Prior to our decoding analysis, to remove amplitude differences across categories, ROIs and VWM processing stages, following our previous studies,[10,15,16,35] we z-normalized each fMRI response pattern. For a given ROI and a pair of conditions, we used SVM for pattern decoding (LibSVM[38]). Specifically, during decoder training, using all the response patterns from one-half of the data, we trained a linear discriminant function to discriminate the fMRI response patterns between the two stimulus conditions. The discriminant function could be expressed as follows: $g(x) = w_i x_i + w_o$, where x_i was the response amplitude of vertex i, w_i was the weight of that vertex, and w_o was the overall bias. After training, all the response patterns from the other half of the data were used to test the classification performance of the discriminant function, ensuring that training and testing were done on independent datasets. Based on the output (either $g(x) > 0$ or $g(x) < 0$), the test data were assigned to one of the two stimulus conditions. We performed training and testing across the two-halves of the data in both directions, with each half serving as the training data and the other half as the testing data. Classification accuracies were averaged across the two training and testing directions.

Training and testing were performed separately for each pair of object categories and the decoding results were averaged across all pairs to derive the average decoding performance. This was done separately in each ROI within each of the three VWM stages (Figure 1D) as well as within each TR (see Figure S2A). To streamline the analysis, based on the similarity of the results obtained and the anatomical location of the ROIs, we averaged the decoding results from multiple ROIs to form four sectors: V1–V4, LOT/VOT, V3ab/IPS0, and IPS1–IPS4 (Figure 2A). Our main comparisons focused on the differences between LOT/VOT and IPS1–IPS4.

In an effort to obtain comparable decoding performance in LOT and VOT during the three VWM stages and then examine the resulting object representational structure formed (see below), instead of examining peak encoding and probe responses in LOT and VOT, we also examined the off-peak responses in these regions. By taking the rising part of the encoding responses (from 3.2 to 4.8 s) and falling part of the probe responses (from 21.6 to 24 s), we were able to obtain in these ROIs comparable category decoding accuracies during encoding, probe and delay and similar to those of IPS1–IPS4 during all three VWM stages (Figure 1E).

Following the same decoding procedure, we also performed cross-decoding across the different time points (in which a classifier trained at one time point was asked to decode the responses across all times points, including the trained time point; Figure S2B). We also performed cross-decoding across the three VWM stages and reported the averaged results for LOT/VOT (peak and offpeak) and IPS1–IPS4 (Figure 2C).

It should be noted that using run combination and a split-half approach in the present pattern decoding analysis were necessary due to the large number of conditions included in each run with very few repetitions of each condition and the significant overlap between adjacent trials due to hemodynamic response lag. With this approach, as reported in the main results, we were able to replicate our previously published results using block design and show the 2-pathway structure across OTC and PPC areas as well as show the role of feedback in shaping VWM representation in OTC in the present study. This indicates the utility of such an approach as a

viable analysis approach to analyze fMRI response patterns to overcome the challenges faced in the present data analysis. That being said, a more thorough and systematic documentation of this approach is needed by testing both simulated and real fMRI data to help us understand how decoding performance may differ should response patterns from the individual runs were available and the data were analyzed using the standard leave-one-out cross-validation procedure. Doing so would allow us to more effectively use the split-half method developed here in situations in which only a smaller amount of fMRI data could be collected due to either the nature of the design (e.g., in a slow event-related design with large number of conditions and fewer trials as in the present study) or the limited time available for data collection (e.g., in a patient or child study).

RDM analysis

As we did in a prior study,[10] to directly visualize how the similarities between object categories were captured by a given brain region, we performed MDS analysis[51] on the object representational structures (i.e., RDMs). MDS takes as input either distance measures or correlation measures. From the group average pairwise SVM category decoding accuracy, we constructed an 8 by 8 category-wise RDM with the diagonal being 0.5 and the other values ranging from around 0.5 to 1 (Figure 1F). To turn these values into distance measures that started from 0 in order to perform MDS, we subtracted 0.5 from all cell values and then multiplied all values by 2 to obtain an RDM with the diagonal set to zero. To avoid negative values (as distance measures cannot be negative), we replaced all values below zero with zero. We then projected the modified category-wise RDM for a given ROI or sector onto a 2D space with the distance between the categories denoting their relative representational similarities to each other in that brain region or sector. This was done separately for the three VWM processing stages for LOT/VOT and IPS1–IPS4 as shown in Figure 3E.

Note that instead of using pairwise SVM category decoding accuracies to construct the category-wise RDM, other measures such as Pearson's *r*, Euclidean distance, and Mahalanobis distance may also be used. We used SVM decoding performance here as it provides a sensitive way to measure VWM representations during delay and has been used extensively in previous VWM studies.[24,34,35] The present study followed this established practice in the VWM literature so that results obtained here can be compared with those reported previously. SVM decoding results have also been used to construct the RDMs as we did in our previously published studies and have generated robust, meaningful and replicable results.[10,15–17]

To determine the extent to which object category representations in one ROI were similar to those in another one and across the three VWM processing stages, we performed pairwise correlation of the unmodified category-wise RDMs (i.e., without subtracting 0.5 from the diagonal or replacing the values below zero with zero) across ROIs and VWM stages to form a region-processing-wise RDM. This was done by concatenating all the off-diagonal values of the unmodified category-wise RDM for a given ROI at a given VWM stage to create an object category similarity vector for that ROI in that stage and then performing pairwise correlation of these vectors across ROIs and VWM stages. To visualize the representational similarity of the different ROIs in the different VWM stages, we projected the region-processing-wise RDM onto a 2D space using MDS, with the distance between the ROIs in particular VWM stages denoting their relative similarities to each other (Figure 1G).

Given the higher object category decoding accuracy in OTC during VWM encoding and probe, it could be argued that fMRI decoding performance might have been at ceiling, leading to distorted OTC RDMs and a distorted region-processing-wise MDS plot. To examine this possibility, as described earlier, instead of examining peak encoding and probe responses in LOT and VOT, by taking off-peak encoding and probe responses, we were able to obtain comparable category decoding performance during encoding, probe and delay in these ROIs. These off-peak encoding and probe decoding accuracies were also similar to those of PPC ROIs. We then derived the RDMs from these off-peak decoding accuracies and projected them in the same MDS projection plot with the other ROIs (Figure 1G), allowing us to directly visualize the deviations of these off-peak RDMs from the original peak RDMs.

To provide a quantitative evaluation of the similarity between the RDMs across different regions and VWM processing stages, we examined the RDM correlation between brain sectors and between the different VWM stages. We focused our analyses on LOT/VOT and IPS1–IPS4, as V1–V4 showed overall weak VWM decoding and RDMs from all regions were clustered around higher IPS ROIs during VWM delay. Given that responses from encoding and probe showed similar decoding performance and RDMs, we further averaged RDM correlations involving encoding and probe and label it as perception. We performed four comparisons between RDM correlations: RDM correlation between LOT/VOT and IPS1–IPS4 during perception vs. VWM delay; RDM correlation between perception and VWM within LOT/VOT vs. within IPS1–IPS4; RDM correlation between VWM and perception within LOT/VOT vs. RDM correlation between LOT/VOT VWM and IPS1–IPS4 VWM; and RDM correlation between VWM and perception within LOT/VOT vs. RDM correlation between LOT/VOT VWM and IPS1–IPS4 perception (Figures 3A–3D). This allowed us to assess how RDM correlation between these two regions would change between perception and VWM, how RDM correlation between perception and VWM within a brain region would change between these two brain regions, and whether LOT/VOT VWM RDM was more similar to itself during perception than to IPS1–IPS4 during VWM or perception. To evaluate the extent to which these correlations might have been driven by noise correlation rather than the structure in the data, we randomly shuffled the cells in the RDMs and then performed the same set of correlations (Figures S3A–S3D).

We also defined a slightly earlier delay period (from 11.2 to 12.8 s) that did not overlap with the probe presentation (at 13 s) and examined the same set of correlations. Highly similar but slightly weaker results were obtained (Figures S3A–S3D). This was largely due to an increased RDM correlation between perception and VWM in LOT/VOT in the peak responses, consistent with potential signal contamination from encoding. The encoding and delay periods were temporally much closer in this case and the sluggishness of the fMRI responses may contribute to their temporal correlation. It is worth noting that even with this earlier delay period, all four correlation comparisons remained significant for LOT/VOT off-peak responses as in the main results.

Statistical analyses

To evaluate if category decoding performance was above chance for the three VWM processing stages as reported on Figures 1D, 1E and 2A, we performed t tests (1-tailed, as only effects above chance were meaningful) and corrected for multiple comparisons for the three tests performed within each ROI or sector using the Benjamini-Hochberg method.[22] For the within- and cross-decoding results shown on Figure 2C, we performed 1-tailed t tests (as only effects above zero were meaningful) to quantify the overall decoding performance and the drop in cross-decoding, and corrected for multiple comparisons for the three tests performance within each sector using the Benjamini-Hochberg method.[22] Two-tailed t tests were performed to quantify the difference in within-decoding and the cross-decoding drop between LOT/VOT offpeak and IPS1–IPS4. For the four correlation comparisons within the peak and within the off-peak conditions shown in Figures 3A–3D and S3A–S3D, we Fisher-transformed all correlation coefficients before we tested if each transformed correlation coefficient was above zero using t tests (1-tailed, as only effects above zero were meaningful); since there were a total of five unique correlations across the four comparisons, we corrected for the five tests performed using the Benjamini-Hochberg method.[22] Given that prior work has shown that feedback from associative areas like PPC and PFC can influence VWM responses in posterior sensory regions, and that the study test whether feedback during VWM simply sustains the object representations formed during perception in OTC or actively reshapes VWM representation in OTC, the study was explicitly set out to test the correlation effects between OTC and PPC in a specific direction. As such, to test whether OTC VWM representations becomes more aligned with those of PPC during VWM and perception than with itself during perception, we performed four correlation comparisons (after Fisher-transforming all correlation coefficients) using 1-tailed t tests and corrected for the four comparisons performed using the Benjamini-Hochberg method.[22] Corrections were done separately within the peak and within the off-peak correlations and comparisons.