NeuroImage 277 (2023) 120222

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Mapping multidimensional content representations to neural and behavioral expressions of episodic memory

Yingying Wang[a,b], Hongmi Lee[c], Brice A. Kuhl[b,*]

[a] Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou 310028, China
[b] Department of Psychology, University of Oregon, Eugene, OR 97403, USA
[c] Department of Psychological and Brain Sciences, Johns Hopkins University, Baltimore, MD 21218, USA

ARTICLE INFO

Keywords:
Lateral parietal cortex
Visual cortex
Inverted encoding model
Reconstruction
Episodic retrieval

ABSTRACT

Human neuroimaging studies have shown that the contents of episodic memories are represented in distributed patterns of neural activity. However, these studies have mostly been limited to decoding simple, unidimensional properties of stimuli. Semantic encoding models, in contrast, offer a means for characterizing the rich, multidimensional information that comprises episodic memories. Here, we extensively sampled four human fMRI subjects to build semantic encoding models and then applied these models to reconstruct content from natural scene images as they were viewed and recalled from memory. First, we found that multidimensional semantic information was successfully reconstructed from activity patterns across visual and lateral parietal cortices, both when viewing scenes and when recalling them from memory. Second, whereas visual cortical reconstructions were much more accurate when images were viewed versus recalled from memory, lateral parietal reconstructions were comparably accurate across visual perception and memory. Third, by applying natural language processing methods to verbal recall data, we showed that fMRI-based reconstructions reliably matched subjects' verbal descriptions of their memories. In fact, reconstructions from ventral temporal cortex more closely matched subjects' own verbal recall than other subjects' verbal recall of the same images. Fourth, encoding models reliably transferred across subjects: memories were successfully reconstructed using encoding models trained on data from entirely independent subjects. Together, these findings provide evidence for successful reconstructions of multidimensional and idiosyncratic memory representations and highlight the differential sensitivity of visual cortical and lateral parietal regions to information derived from the external visual environment versus internally-generated memories.

1. Introduction

Neuroimaging studies of human episodic memory have found that the contents of memory retrieval are reflected in broadly distributed patterns of neural activity (Danker and Anderson 2010; Rissman and Wagner 2012). While initial fMRI decoding studies of memory focused on relatively coarse information such as the visual category to which a stimulus belongs (Kuhl et al., 2011; Polyn et al., 2005), more recent studies have demonstrated item- or event-specific representations (Favila et al., 2018; Lee et al., 2019; St-Laurent et al., 2015; Xiao et al., 2017). However, these studies have overwhelmingly focused on decoding simple, unidimensional, and objective properties of stimuli. In contrast, real-world episodic memories are complex, multidimensional, and subjective (Cooper and Ritchey 2019; Richter et al., 2016). Notably, this limitation is often paralleled in behavioral measures of memory where simple, categorical expressions of retrieval success or accuracy are more common than the kinds of complex and idiosyncratic descriptions humans *actually use* to describe memories (Chen et al., 2017; Gilmore et al., 2021; Heusser et al., 2021). Thus, an important challenge for the field is to develop neuroimaging measures that capture the richness and complexity of episodic memory retrieval and to directly *align* these measures with behavioral expressions of memory that have similar richness and complexity.

A handful of recent fMRI studies have moved closer toward capturing the richness of memories by using voxel-wise encoding/decoding models (Kay et al., 2008; Naselaris et al., 2011) to map fMRI activity patterns to multidimensional measures of memory content. For example, Naselaris et al. (2015) demonstrated that low-level visual features can be successfully reconstructed during mental imagery. Specifically, they extracted low-level visual features from complex natural images and trained algorithms to predict these features from fMRI activity patterns elicited during visual perception. This mapping was then used to predict features of independent natural images based on activity patterns evoked during mental imagery. Using a similar approach applied

* Corresponding author.
E-mail address: bkuhl@uoregon.edu (B.A. Kuhl).

https://doi.org/10.1016/j.neuroimage.2023.120222.
Received 6 January 2023; Received in revised form 6 June 2023; Accepted 8 June 2023
Available online 14 June 2023.

to higher-level visual information, Lee and Kuhl (2016) mapped distinct face components to patterns of fMRI activity and then used these mappings to reconstruct faces held in working memory. In another study, Bone et al. (2020) applied deep convolutional neural networks to complex natural images in order to extract content information and map that information to patterns of fMRI activity. They demonstrated that, during memory recall, fMRI activity patterns reflect content information across multiple levels: from low-level visual information to high-level semantic concepts. Collectively, these studies provide important evidence that multidimensional content representations can be mapped to patterns of neural activity evoked during memory retrieval. Notably, however, none of these studies used behavioral measures of memory that matched the richness of the neural measures.

Complementing the studies described above, other fMRI studies have embraced more complex *behavioral measures* of verbal recall (Chen et al., 2017; Gilmore et al., 2021; Heusser et al., 2021; Nguyen et al., 2019). For example, Chen et al. (2017) and Nguyen et al. (2019) applied latent semantic analysis (LSA) to verbal recall of movies and Heusser et al. (2021) used topic models to measure changes in verbal recall content over time. Each of these studies found that subject-specific measures of verbal recall content were related to measures of fMRI activity. For example, in Chen et al. (2017) and Nguyen et al. (2019), subjects with more similar behavioral expressions of recall—or more similar interpretations of the stimuli—showed greater fMRI pattern similarity. In Heusser et al. (2021), the specific time course of content changes during verbal recall was predicted by changes in fMRI activity. These studies strongly attest to the feasibility and value of measuring subject-specific verbal recall and relating these behavioral expressions to patterns of neural activity. However, it is important to note that these studies did not directly measure content information within fMRI data and, therefore, they did not directly *align* the content of behavioral recall with the content of fMRI activity patterns.

To the extent that multidimensional memory representations are captured by patterns of neural activity, an additional question is how these representations are distributed across cortical areas. Traditionally, memory-based content representations have been measured within (or decoded from) sensory cortical regions involved during initial perceptual experience (Danker and Anderson 2010). In particular, much of this work has focused on ventral temporal cortical areas which represent high-level visual category information (Kuhl et al., 2011; Polyn et al., 2005). However, there is now substantial and accumulating evidence that the contents of memory retrieval are also robustly reflected in lateral parietal cortex (LPC) (Kuhl and Chun 2014; St-Laurent et al., 2015; Xiao et al., 2017). Much of this work has focused on the angular gyrus, which is not only a core component of the episodic memory network (Gilmore et al., 2015; Rugg and Vilberg 2013) but is heavily involved in semantic processing (Humphreys et al., 2021). Indeed, several recent findings specifically suggest that LPC—and angular gyrus, in particular—contains the kinds of rich, multidimensional information that is critical for episodic remembering (Bonnici et al., 2016; Cowen et al., 2014; Favila et al., 2018; Huth et al., 2016; Lee et al., 2019; Lee and Kuhl 2016; Yu and Shim 2017). Interestingly, there is also emerging evidence for a potential dissociation in content representations across LPC and sensory cortices. Namely, whereas content representations in sensory cortex are generally *weaker* during memory retrieval compared to perception, content representations in LPC may be *as strong or stronger* during memory retrieval compared to perception (Favila et al., 2018, 2020; Long and Kuhl 2021; Xiao et al., 2017). Thus, understanding how memory representations are distributed across LPC and ventral temporal cortical areas remains an important objective with implications for theories of memory (Favila et al., 2020; Rugg and King 2018).

Here, we sought to bridge neuroimaging methods for measuring multidimensional content representations with behavioral methods for measuring complex, subjective, and idiosyncratic expressions of memory. To this end, we used semantic encoding models (Kay et al., 2008) and an extensive-sampling fMRI design (thousands of trials per subject) to map multidimensional semantic information from natural scene images to fMRI activity patterns. We then inverted these encoding models (Ester et al., 2015; Kok et al., 2020; Sprague et al., 2016) to reconstruct semantic information as subjects viewed and recalled images from memory. These fMRI-based content reconstructions were directly compared to subjects' verbal recall of the scenes using natural language processing methods. This allowed us to test not only whether fMRI-based reconstructions captured the objective content within scene images, but whether reconstructions matched subjective—and potentially idiosyncratic (subject-specific)—details of how scenes were remembered. Additionally, by comparing reconstructions generated from different regions of visual cortex and LPC, we tested whether these regions differentially expressed content information during image viewing versus image recall. Finally, we tested whether semantic encoding models successfully generalized across subjects—a question that has important implications for leveraging data-rich models from extensively-sampled subjects.

2. Materials and methods

2.1. Subjects

Nineteen experimental sessions were collected from four human subjects (two females, age 23–30 years) from the University of Oregon community. Three subjects completed five sessions each; one subject only completed four sessions due to unavailability for a 5th session. The sample size was modeled after Naselaris et al. (2015), which used a similar encoding model procedure for memory-based reconstructions with a sample size of 3 subjects and 5–6 sessions per subject. Despite our small sample size, each subject was sampled extensively across a large number of stimuli, a procedure which may have distinct advantages compared to sampling many individuals across a more limited number of stimuli (Naselaris et al., 2021). All subjects were right-handed and reported normal or corrected-to-normal vision. Informed consent was obtained in accordance with procedures approved by the University of Oregon Institutional Review Board.

2.2. Stimuli

Two sets of image stimuli were prepared: one for use in a recognition memory task and one for use in a recall memory task. The recognition set contained a total of 5000 complex scene images, which were selected from the Microsoft COCO dataset (http://cocodataset.org/, Lin et al., 2015). These images depict complex everyday scenes of common objects from 91 categories in their natural context. Each image in the dataset is annotated with five written descriptions from independent human subjects. These descriptions capture the main content of the images and were used, in the present study, as information channels for the inverted encoding model. For each subject and each session, 680 images were randomly selected (without replacement) from the recognition set. Of these, 600 were studied prior to the fMRI session and served as 'old' items in the recognition test. The remaining 80 images served as novel foils ('new' items) in the recognition test. The recall set consisted of a total of 100 images, also drawn from the Microsoft COCO dataset. Each image was randomly paired with a texture (taken from the internet), creating a set of paired associates. The textures served as cues during the cued recall task (described below). For each session, 20 of the paired associates were studied and tested. The same 20 pairs were used in each session for each subject in order to facilitate across-subject analyses.

2.3. Experimental design and procedures

Overview of paradigm. Each session of the experiment consisted of two separate parts which were conducted across consecutive days (Fig. 1A).

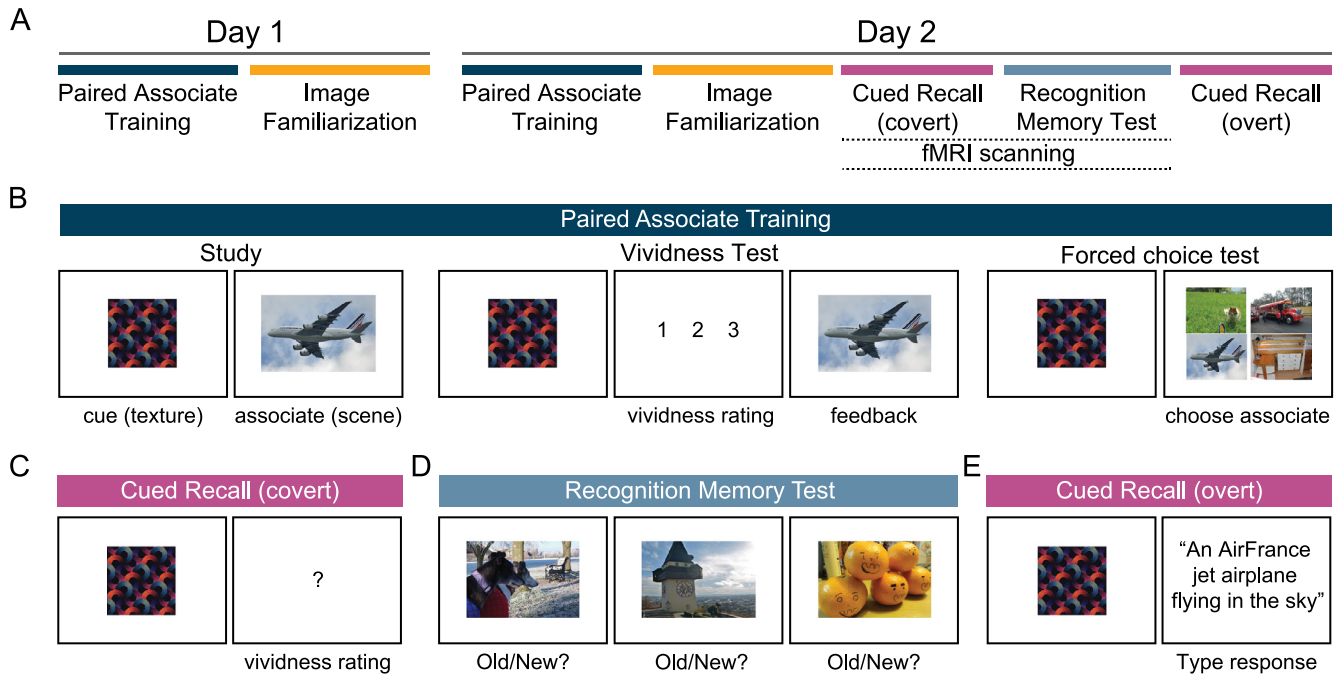

Fig. 1. Experimental procedure. **A**. Overview of experimental phases. Each subject completed 4 to 5 experimental sessions. Each experimental session involved two consecutive days of tasks. On Day 1, subjects learned 20 associations between cues (textures) and associates (scenes) via a paired associate training procedure and were also familiarized with 600 additional scene images (image familiarization). No fMRI scanning was conducted on Day 1. On Day 2, subjects completed additional paired associate training and image familiarization before entering the scanner. During scanning, subjects completed a covert cued-recall test of the cue-associate pairs followed by a recognition memory test. After exiting the scanner, subjects completed an overt cued-recall test for the cue-associate pairs. **B**. The paired associate training included three phases: study, vividness test, and forced choice test. During study, subjects were shown textures followed by scenes and attempted to learn these associations. During the vividness test, subjects were shown textures and then indicated the vividness with which they were able to recall the corresponding scene. The correct scene was then shown as feedback. During the forced choice test, subjects were shown a texture followed by four previously-studied scenes and were asked to select the corresponding scene. **C**. In the scanned (covert) cued-recall phase, subjects were shown textures and rated the vividness with which they could recall the corresponding scene (as in the vividness test, but without feedback). **D**. In the scanned recognition memory test, subjects made old/new judgements for scenes (that did not include the scenes from the paired associate training). The sole purpose of the recognition memory test was to train the semantic encoding models. **E**. In the final (overt) recall test, subjects were shown textures and typed a sentence to describe the content of the corresponding scene.

The two-day protocol was intended to minimize fatigue given the overall length of the tasks. On Day 1 of each session, subjects were overtrained on 20 paired associates (textures + scene images) and were familiarized with a separate 600 scene images (from the recognition set). On Day 2 of each session, subjects first completed additional training on the paired associates from Day 1 and additional familiarization with images from the recognition set. Then, during fMRI scanning subjects completed two phases: (1) a covert cued recall phase in which the 20 textures were repeatedly used to test memory for the associated scenes, and (2) a recognition memory phase which included the 600 familiarized images + 80 lures. The rationale for conducting the cued recall phase before the recognition memory phase was to minimize interference/forgetting during the cued recall phase. Finally, subjects exited the scanner and completed an overt (verbal) cued recall test for the 20 paired associates. This two-day procedure constituted a single session and each subject completed 4–5 sessions. In order to minimize across-session memory interference, there was a delay of at least 7 days between sessions, for each subject. Each subject completed all of their sessions within a 2-month window.

Paired associate training. Paired associate training was conducted at the beginning of Day 1 (4–5 rounds) and the beginning of Day 2 (1 round). Each round of paired associate training consisted of three distinct phases in the following sequence: study, vividness test, study, vividness test, forced choice associative memory test. In the study phases, subjects saw and deliberately encoded each of the 20 paired associates, one pair at a time. On each trial, the cue (texture) was first presented for 1 s, followed by a fixation cross for 0.5 s, and then the target image (scene) for 2 s. Another fixation cross was presented for 1 s at the end of each trial (before the start of the next trial). In the vividness test phases, each cue was presented for 1 s followed by a 3-point vividness scale ("1 2 3") and subjects reported, via button press, the vividness with which they could recall the target image (1-*Can't remember,* 2-*Remember,* and 3-*Vividly remember*). The rationale for using a vividness report was that (1) it encouraged participants to recall the images in detail, (2) it provided a measure of task vigilance/compliance, and (3) it did not explicitly orient subjects to specific semantic dimensions. The vividness report was self-paced. After responding, feedback was given by presenting the target image alone on the screen for 1.5 s. A fixation cross was presented for 0.5 s in between trials. In the forced choice associative memory test, a cue image was first presented for 1 s and then, after a fixation cross (0.5 s), four scene images appeared on the screen. The four images included the target (correct) scene along with three scenes randomly selected from the remaining 19 scenes in the set of paired associates studied in the current session. Subjects were instructed to select the scene image that had been paired with the cue by pressing one of four keys. There was no time limit to respond. After subjects made a selection, feedback was provided. If the correct image was selected, a green fixation cross was presented (0.5 s) followed by the correct image presented in the center of the screen (1 s). If an incorrect image was selected, a red fixation cross was presented (0.5 s) followed by the correct image (1 s) presented in the center of the screen. Finally, a black fixation cross was presented for 1 s (until the start of the next trial). For each session, subjects were required to reach at least 95% accuracy for two consecutive rounds on Day 1 before proceeding to the Image Familiarization Phase. Using this performance criterion, all subjects completed 5 paired associate training rounds on Day 1 for each session, with the

exception of one subject that reached the criterion by the 4th round for one of the sessions.

Image Familiarization. For each session, Image Familiarization was conducted on Day 1 and Day 2, immediately after the paired associate training rounds. During each familiarization phase subjects saw all 600 scene images presented in the center of the screen, one at a time and in random order, and distributed across five blocks (120 images/block). Subjects were instructed to try their best to remember each image for a later memory test (the recognition memory test). No behavioral responses were made. On Day 1, each image was presented for 1 s with a 0.5 s fixation cross in between trials. On Day 2, each image was presented for 0.6 s with a 0.4 s fixation point in between trials.

Scanned Cued Recall. For each session, subjects completed two fMRI runs of a covert cued recall task (Fig. 1B, left), each lasting 6 min and 16 s. As during the paired associate training rounds, subjects were shown cues (textures) and indicated the vividness with which they could recall the corresponding scene image using a 3-point scale. Each run consisted of 40 recall trials (2 trials per association per run), with the order of trials in each run pseudorandomized with the constraint that the same association was not tested consecutively. Every trial started with a cue image centrally presented over a white background for 0.5 s. Next, a question mark appeared in the center of the screen (3.5 s), prompting subjects to make their vividness response using a button box. Finally, a fixation cross was presented either for 4 s (75% of trials) or 8 s (25% of trials).

Scanned Recognition Memory Test. Following the cued recall task, subjects completed the recognition memory test (Fig. 1B) which consisted of eight runs, each lasting 6 min and 20 s. Each run contained 75 old images and 10 novel images presented in random order, for a total of 680 images across the 8 runs. Each trial began with the presentation of a scene in the center of the screen (1 s). Next, a question mark (3 s) prompted subjects to make an old/new decision by pressing one of two keys on a button box. After a small number of the recognition trials (6/85), a fixation cross was presented for 4 s. The rationale for including a disproportionate number of old images (600 out of 680) was because fMRI data from the recognition memory test was used to train encoding models applied to the cued recall task and we sought to increase the extent to which these models were trained on 'memory data' (old trials). Specifically, recent evidence indicates systematic differences in the spatial activity patterns associated with memory-based content representations compared to perception-based content representations (Favila et al., 2020). Thus, our intuition—though not a point we directly tested—was that transfer from the recognition to recall trials might benefit from the recognition trials having a high percentage of old trials. Additionally, the recognition memory test served as a cover task to help keep subjects engaged while viewing hundreds of images per session.

Post-scan Cued Recall. After subjects exited the scanner, they completed a final cued recall test (Fig. 1B). However, in contrast to the prior cued recall tests which recorded covert (vividness) memory judgments, here subjects were asked to explicitly describe their memories. On each trial, subjects were shown a cue (texture) and were asked to type a sentence to describe the content of the associated scene image. Specifically, the instructions asked subjects to "write a complete but simple sentence" that should "include adjectives if possible, describe the main characters, the setting, or the relation of the objects in the image, and try to be concise". After subjects typed their response on the computer screen, they pressed enter to advance to the next trial. No time limit was given and each of the 20 associations from the current session was tested once, in random order.

2.4. fMRI data acquisition

fMRI scanning was conducted on a Siemens 3 T Skyra scanner at the Robert and Beverly Lewis Center for NeuroImaging at the University of Oregon. Before the functional imaging, a whole-brain high-resolution anatomical image was collected for each subject and each session using a T1-weighted protocol (grid size 256 × 256; 176 sagittal slices; voxel size 1 × 1 × 1 mm). Whole-brain functional images were collected using a T2*-weighted multi-band accelerated EPI sequence (TR = 2 s; TE = 25 ms; flip angle = 90°; 72 horizontal slices; grid size 104 × 104; voxel size 2 × 2 × 2 mm). Each cued recall scan consisted of 188 volumes. Each recognition memory test scan consisted of 190 volumes.

2.5. fMRI data preprocessing

MRI data were first converted to Brain Imaging Data Structure (BIDS) format using in-house scripts. MRIQC v0.15.1 (Esteban et al., 2017) was used for preliminary data quality assessment. We applied a threshold that no more than 20% of TRs in any scan run could exceed a framewise displacement of 0.3 mm; however, no scan runs were excluded using this threshold. Preprocessing was performed using FMRIPrep v1.5.4 (RRID:SCR_016216) (Esteban et al., 2019), a Nipype (RRID:SCR_002502) based tool, with the default processing steps. Each structural image was corrected for intensity non-uniformity and skull-stripped. Brain surfaces were reconstructed using recon-all from FreeSurfer v6.0.1. Spatial normalization to the ICBM 152 Nonlinear Asymmetrical template version 2009c was performed through nonlinear registration with the antsRegistration tool of ANTs v2.1.0, using brain-extracted versions of both T1w volume and template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using FAST (FSL v5.0.9).

Functional data were slice time corrected, motion corrected, and corrected for field distortion. This was followed by co-registration to the corresponding T1w using boundary-based registration with six degrees of freedom using bbregister (FreeSurfer v6.0.1). Motion correcting transformations, BOLD-to-T1w transformation and T1w-to-template (MNI) warp were concatenated and applied in a single step using antsApplyTransforms (ANTs v2.1.0) using Lanczos interpolation. We then applied a high pass filter using a cutoff period of 100 s. Finally, the preprocessed fMRI data were smoothed by a 1.6 mm full-width-half-maximum Gaussian kernel with FSL's SUSAN (Smoothing over Univalue Segment Assimilating Nucleus) (Smith and Brady 1997). Grand-mean intensity normalization of each functional image volume was performed by a single multiplicative factor. Confounding regressors including framewise displacement (FD), global signal, white matter, and cerebrospinal fluid signals were generated for each volume. Within-subject reconstructions were conducted in subjects' native EPI space, and across-subject reconstructions were conducted in the standard space.

2.6. Regions of interest

Regions of interest (ROIs) included four subregions of the posterior lateral parietal cortex (LPC), the ventral temporal cortex (VTC), and the occipital temporal cortex (OTC) (Fig. 4A). Post-hoc analyses for additional brain regions, including two control ROIs (primary auditory cortex, primary motor cortex) are reported in Table S1. ROIs were defined using FreeSurfer's Destrieux atlas (the following label numbers refer to Simple_surface_labels2009.txt). The subregions of LPC consisted of the angular gyrus (ANG, #25), supramarginal gyrus (SMG, #26), superior parietal lobule (SPL, #27), and intraparietal sulcus (IPS, #57). The VTC ROI was comprised of regions 21, 23, 51, 52, 61, and 62. The OTC ROI was comprised of regions 2, 19, 43, 58, and 60. The ROIs were co-registered to the functional images and further masked by subject-specific whole-brain masks generated from functional images to exclude areas where signal dropout occurred. All ROIs contained brain regions from both hemispheres (mean number of voxels for each ROI: 1816 in ANG; 1942 in SMG; 1594 in SPL; 1318 in IPS; 4459 in VTC; 3854 in OTC).

2.7. Single-item response estimation

For each session, separate general linear models (GLM) were created for each of the 20 images during the cued recall task and each of the 680 images during the recognition memory test. A least-square single method was used for each item, where the given item was modeled with a single regressor and all the remaining items were modeled with another regressor. The presentation of each stimulus was modeled as an impulse and convolved with a canonical hemodynamic response function (double gamma). The GLM included head-motion parameters (six rotation and translation head movement estimates) and nuisance regressors marking outlier TRs (FD > 0.3 mm from previous TR) as confounding regressors. The *t*-statistic values associated with each image were used in the semantic encoding model to increase reliability by noise normalization (Walther et al., 2016).

2.8. Image content reconstruction

To represent the content of each scene image, we used the Word2vec embedding algorithm. This algorithm transforms single words into 300-dimensional vectors (word embeddings). Similarities/distances between these vectors reflect the similarity of the corresponding words. In our analysis, we took advantage of the annotation captions (five captions for each image) from the COCO dataset. After a standard preprocessing procedure that included filtering of stop words and tokenization, we obtained the word embeddings for the critical words in the annotation captions. Notably, no corrections were applied for negations given their very low frequency. We calculated the mean vector, across the five captions, to represent the content in each image (Fig. 2A). We then applied principal component analysis (PCA) on the entire pool of 300-dimensional word embeddings for the 5100 images (i.e., the full set of recognition + recall images). The first 30 components, which explained 68.59% of the total variance (Fig. S1), were used as information channels in the semantic encoding model. We refer to these 30 components as *semantic component scores.* The goal of reconstruction analyses was to accurately predict the semantic component scores.

Reconstructions of semantic component scores were generated using a cross-validation approach. fMRI activation patterns evoked during the recognition trials for 'old' images were used as training patterns to estimate the relationship between fMRI activity patterns and semantic component scores (Fig. 2B). Data from 'new' trials during the recognition memory test were not used for training (or testing) of the encoding models. We modeled the response in each voxel as a weighted sum of the information channels (i.e., the 30 semantic components):

$$B_1 = WC_1$$

where B_1 (n images × m voxels) is the activation patterns of voxels (*t* maps) during the recognition memory test, C_1 (n images × k components) is the modeled response of each component, or information channel, on each trained image, and W (k components × m voxels) is a weight matrix quantifying the contribution of each information channel to each voxel (Fig. 2B). We can solve for W using the following ordinary least-squares linear regression:

$$\hat{W} = B_1 C_1^T \left(C_1 C_1^T\right)^{-1}$$

Given the estimated weights within an ROI ($\hat{W}$) and a novel pattern of activation (B_2) from the recognition trials (recognition-based reconstruction) or the recall trials (recall-based reconstruction), we can compute an estimate of the semantic component scores by inverting the model (Fig. 2C):

$$\hat{C}_2 = \left(\hat{W}^T \hat{W}\right)^{-1} \hat{W}^T B_2$$

Recognition-based reconstruction. Separately for each subject, N-fold cross-validation was performed on recognition data where N equals the number of scanning runs pooled across all of the sessions that each subject completed (i.e., 40 runs for the three subjects that completed 5 sessions each and 32 runs for the remaining subject that completed 4 sessions). For each fold, the activation patterns from N-1 runs (i.e., 39 or 31 runs) were used as training patterns and those of the remaining run served as the testing set (i.e., the trials for which the semantic component scores were predicted). In this manner, all trials iteratively contributed to both model training and model testing.

Recall-based reconstruction. To predict semantic component scores during recall trials, activation patterns from recognition runs were used as training data and the estimated weights based on the recognition trials (training data) were then applied to the recall trials (testing data) to predict the semantic component scores for each of the recalled images. Recall-based reconstructions were performed in two ways: within-subjects and across-subjects. For within-subject reconstructions, all of the recognition runs across all sessions for a given subject were used as the training data and the testing data were all of the recall runs across all sessions for that same subject. For across-subject reconstructions, all of the recognition runs across all sessions from N-1 subjects were used as the training data and the testing data were all of the recall runs across all sessions from the held-out subject.

2.9. Reconstruction accuracy

For all reconstruction analyses, accuracy was based on Fisher-transformed Pearson correlations between predicted (reconstructed) and 'actual' semantic component scores. 'Actual' semantic component scores were either based on COCO annotations (which were derived from an independent set of subjects) or from subjects' verbal recall responses (which were collected in the final cued recall test, after fMRI scanning). Unless otherwise noted, successful reconstruction accuracy was defined as greater *within-item correlations* than *across-item correlations* (Fig. 2D). Within-item correlations refer to correlations between reconstructed and actual semantic component scores corresponding to the same image. Across-item correlations refer to the mean of correlations between reconstructed and actual semantic component scores corresponding to different images [e.g., *r*(reconstructed scores for image 1, actual scores for image 2)]. Across-item correlations were always restricted to images from the same fMRI session. Additionally, within-item correlations for recognition-based reconstructions were only compared against across-item correlations for other recognition-based reconstructions; likewise, within-item correlations for recall-based reconstructions were only compared against across-item correlations for other recall-based reconstructions. Group-level results were obtained by first averaging correlations within sessions for each subject and then across sessions and subjects. In addition to correlation values, for each ROI we also report the mean percentile rank of reconstructions (i.e., the rank of the within-item correlation among the distribution of all across-item correlations). Note: the mean rank values were not used for statistical analyses; rather, they are included to provide a more intuitive measure of reconstruction accuracy.

2.10. Statistical analysis

Unless otherwise stated, mixed-effects models were used to test the reconstruction accuracy of correlation difference measures. Linear mixed-effects models were implemented with lme4 in R 3.6.3, fitted using restricted maximum likelihood. To determine whether within-item correlations differed from across-item correlations, we used the likelihood ratio test to compare models with (full model) and without (null model) the predictor of interest (i.e., correlation type: within-item correlation or across-item correlation). Subject and session numbers were included as random factors. For statistical tests of reconstruction accuracy within individual ROIs, uncorrected *p* values are reported. In tests that compared reconstruction accuracies across ROIs or conditions, mixed-effects models were used with the subject number and session number included as random factors.

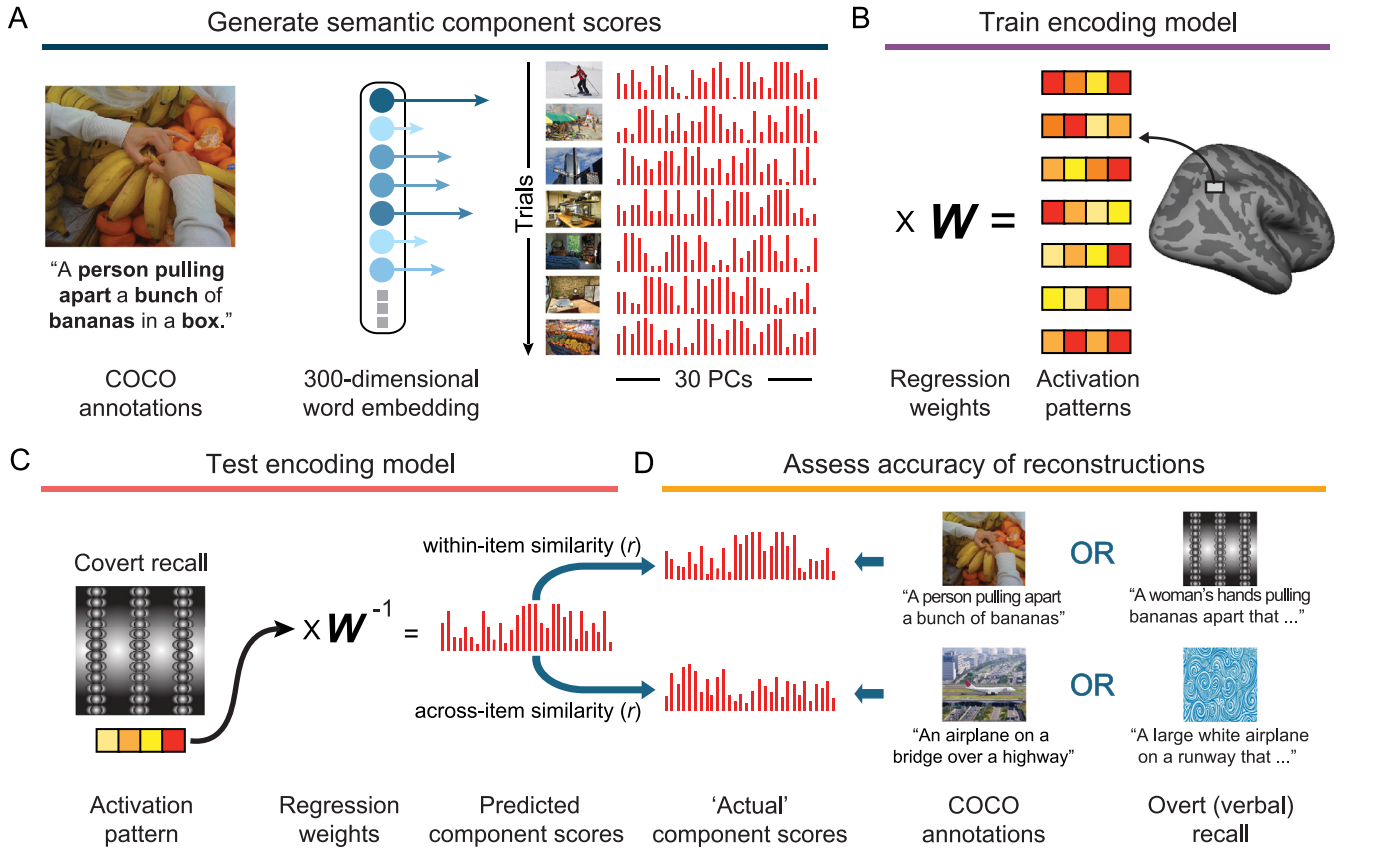

Fig. 2. Schematic overview of the semantic content reconstruction analysis. **A**. Generating semantic component scores. Annotations from the COCO image dataset were used as semantic descriptions of the images. After filtering out the stop words, the captions were transformed into 300-dimensional vectors using the Word2Vec embedding method. PCA was run on all of the 5100 candidate images, and the first 30 principle components (hereinafter, semantic components) were extracted so that the content of each image could be expressed as a weighted sum of these components. **B**. Training the encoding model. Linear regression was used to estimate a model that learned the mapping between the semantic component scores of the trained images (i.e., the training set) and the fMRI activation patterns they evoked. **C**. Testing the encoding model. The regression weights obtained from the training set were applied to an independent set of images (i.e., the testing set) to predict semantic component scores. Encoding models were tested using cued recall trials (shown) or recognition trials (not shown). **D**. Assessing reconstruction accuracy. The accuracy of reconstruction for each image was determined by computing the Pearson correlations between the predicted semantic component scores and the actual semantic component scores. Actual semantic component scores were either based on the COCO dataset captions (left side of boxes) or the verbal recall responses subjects generated in the final cued-recall test (right side of boxes). Correlations were separately computed for 'matching' images (within-item similarity) and non-matching images (across-item similarity). Reconstructions were considered accurate if within-item similarity was higher than across-item similarity.

3. Results

3.1. Behavioral performance

Group-level results were obtained by first averaging data within sessions for each subject and then across sessions and subjects. On Day 1 of each session, subjects studied 20 paired associates (textures with scenes) across 4–5 rounds. For each round, memory was tested via cued recall and forced-choice associative memory tasks. In the cued recall tasks, subjects reported how vividly they could recall the target image on a 3-point scale. The mean percentage of "Vividly remember" responses (the highest rating) was 49.74% ± 22.96% (SD) in round 1, 93.16% ± 10.23% in round 2, 99.34% ± 1.83% in round 3, 99.74% ± 0.79% in round 4, and 100% ± 0.00% in round 5 (Fig. 3A). In the forced-choice associative memory test, subjects were asked to select the target image for each cue from a set of three image choices. Performance was high across all rounds (Round 1: 97.89% ± 4.19%; Round 2: 98.95% ± 2.09%; Round 3: 98.68% ± 2.27%; Round 4: 98.68% ± 2.81%; Round 5: 97.81% ± 3.15%). Critically, performance remained high at Day 2 as evidenced by the rate of "Vividly remember" responses during the cued recall tasks (pre-scan cued recall: 98.16% ± 3.89%; scanned cued recall task: 98.55% ± 6.00%; Fig. 3A) and performance on the forced-choice associate memory test, which occurred prior to scanning (98.95% ± 2.09%; Fig. 3A). Note: given the extremely high success rate of recall, fMRI analyses comparing successful vs. failed recall trials were not feasible.

After exiting the scanner, subjects completed a final post-scan cued recall task during which they generated a sentence to describe the content in each target image. These subject-specific recall-based descriptions were transformed to 300 dimensional vectors (word embeddings) using Word2Vec. The COCO annotations for each of these images were also transformed to word embeddings using Word2Vec. We then calculated the Pearson correlations between the word embeddings from subjects' verbal recall and those from the independent COCO annotations. As shown in Fig. 3B, each subject exhibited markedly higher within-item correlations (i.e., correlations between verbal recall vectors and COCO annotation vectors corresponding to the same image) than across-item correlations (i.e., correlations between recall vectors and annotation vectors corresponding to different images). These results confirm that subjects were able to accurately describe images from memory and also validate our approach of characterizing verbal recall through word embeddings.

For the recognition memory test conducted during scanning, mean recognition sensitivity (d') across sessions and subjects was 1.98 ± 0.52. The mean hit rate for studied images was 84.63% ± 11.51%, and the mean correct rejection rate for new images was 76.97% ± 12.48%. A

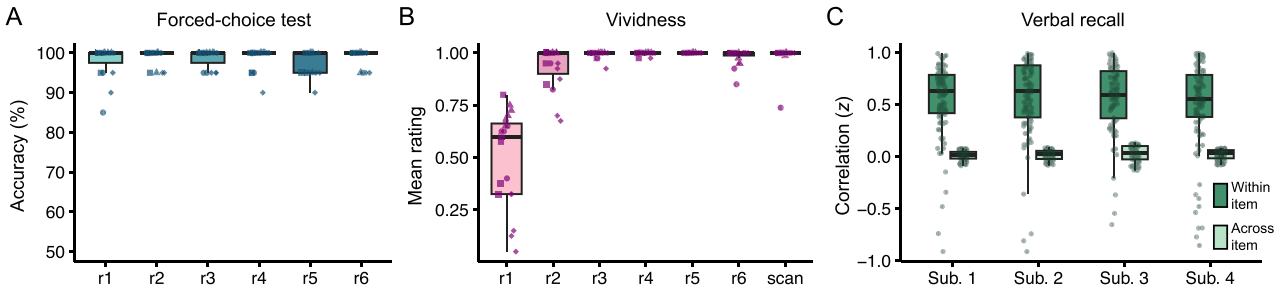

Fig. 3. Behavioral performance across the entire experimental procedure. **A**. Forced-choice test accuracy was measured during the paired associate training rounds on Days 1 and 2. The first five rounds (r1–r5) were completed during Day 1. The 6th round (r6) was completed during Day 2 (just prior to fMRI scanning). Chance accuracy = 25%. **B**. Vividness ratings were made during the first five paired associate training rounds on Day 1 (r1–r5), during the 6th paired associate training round on Day 2 (r6), and during the covert cued recall test during fMRI scanning on Day 2 (scan). Ratings were rescaled from 1, 2, 3 to 0, 0.5, 1.0 with 0 corresponding to the lowest vividness rating and 1.0 to the highest vividness rating. For **A** and **B**, data are represented by boxplots with dots representing data from individual sessions with each subject represented by a different shape. Note: for many of the rounds performance was at ceiling and boxplots are therefore compressed. **C**. Verbal recall performance from the overt cued recall test following scanning on Day 2. For each subject (Sub. 1–4) and each recalled image, Pearson correlations were computed between the 30 semantic components generated from subjects' verbal responses and semantic components generated from the independent COCO annotations of (i) the same images (within-item similarity) or (ii) other images from the recall set (across-item similarity). For within-item similarity, each dot represents the within-item correlation for a single recall trial to its corresponding COCO annotations. For across-item similarity, each dot represents the mean *z*-transformed correlation between a single recall trial and all non-corresponding COCO annotations.

mixed effects model including subject and session numbers as random factors showed that the hit rate was significantly higher than the false alarm rate ($\chi^2_1 = 89.96$, $p < 0.0001$, $\beta = 0.616$, SE = 0.030).

3.2. Reconstruction of content from viewed images

For fMRI analyses, we first tested for successful reconstruction of image content from activity patterns evoked in visual and lateral parietal cortices during the recognition memory task (when images were visually presented on the screen). Image content was defined by 30 semantic component scores derived from the 300-dimensional Word2Vec vectors from the COCO annotations (Fig. 2). The 30 semantic components explained 68.59% of the variance in COCO annotations for the images used in the study. As with the behavioral analyses above, we assessed reconstruction accuracy by comparing within-item vs. across-item similarity. Here, however, within-item similarity was defined as the Fisher-transformed Pearson correlation between the reconstructed component scores for a given image (as predicted from the inverted fMRI encoding model) and the 'actual' semantic component scores for that image (derived from COCO annotations). Across-item similarity was defined as the mean Fisher-transformed Pearson correlation between predicted component scores for a given image and actual component scores for *different images* (from the same session). Reconstruction of content information was determined to be successful if within-item similarity was greater than across-item similarity, as determined by mixed-effects linear models which included subject and session numbers as random factors. Consistent with our previous studies (Cowen et al., 2014; Lee and Kuhl 2016), robust reconstruction accuracies were obtained from visual regions (VTC: rank = 64.52%; $\chi^2_1 = 3024.1$, $p < 0.0001$, $\beta = 0.136$, SE = 0.002; OTC: rank = 66.39%; $\chi^2_1 = 3785.6$, $p < 0.0001$, $\beta = 0.146$, SE = 0.002) as well as ANG and other lateral parietal ROIs (ANG: rank = 57.15%; $\chi^2_1 = 754.4$, $\beta = 0.063$, SE = 0.002; SMG: rank = 54.31%; $\chi^2_1 = 277.8$, $\beta = 0.040$, SE = 0.002; SPL: rank = 56.58%; $\chi^2_1 = 624.0$, $\beta = 0.059$, SE = 0.002; IPS: rank = 57.23%; $\chi^2_1 = 774.5$, $\beta = 0.066$, SE = 0.002; p values < 0.0001) (Fig. 4A,B). However, reconstruction accuracies sharply varied across ROIs (main effect of ROI from repeated-measures ANOVA: $F_{5,90} = 350.69$, $p < 0.0001$, $\eta_p^2 = 0.95$), with higher accuracies in the visual ROIs compared to the parietal ROIs (p's < 0.0001 for all paired-samples t-tests comparing the VTC and OTC ROIs to each of the lateral parietal ROIs). For results from additional brain regions—including prefrontal cortex, lateral temporal cortex, primary motor cortex, and primary auditory cortex—see Table S1.

3.3. Reconstruction of image content from cued recall task

Next, we extended our method to test for content reconstruction for images retrieved from memory during the cued recall task. Critically, and in contrast to recognition-based reconstructions for which the to-be-reconstructed image was visually present, here the to-be-reconstructed image was visually absent (subjects were only shown the texture cues) thus requiring top-down retrieval of the target image. For this analysis, we trained the semantic encoding model with 'old' images from the recognition set (exploiting the large number of recognition trials) but tested it on images from the cued recall task. As with the recognition-based reconstructions, evidence for successful recall-based reconstructions was obtained if within-item similarity (correlations between the semantic component scores predicted from the inverted encoding model and the 'target' semantic component scores) were reliably higher than across-item correlations. As described in the following sections, we used several approaches for defining the 'target' semantic component scores.

As a first step, we defined target semantic component scores based on the COCO annotations (as in the preceding section which tested for reconstruction accuracy during the recognition memory task). Successful recall-based content reconstruction was observed across each of the visual and parietal regions (VTC: rank = 60.08%; $\chi^2_1 = 83.9$, $p < 0.0001$, $\beta = 0.083$, SE = 0.009; OTC: rank = 55.24%; $\chi^2_1 = 24.6$, $p < 0.0001$, $\beta = 0.043$, SE = 0.009; ANG: rank = 55.29%; $\chi^2_1 = 26.6$, $p < 0.0001$, $\beta = 0.049$, SE = 0.009; SMG: rank = 55.22%; $\chi^2_1 = 26.0$, $p = 0.0001$, $\beta = 0.048$, SE = 0.009; SPL: rank = 56.12%; $\chi^2_1 = 24.5$, $p < 0.0001$, $\beta = 0.047$, SE = 0.009; IPS: rank = 57.08%; $\chi^2_1 = 32.6$, $p < 0.0001$, $\beta = 0.056$, SE = 0.010; Fig. 4C). Accuracy significantly varied across ROIs (main effect of ROI: $F_{5,90} = 3.46$, $p = 0.007$, $\eta_p^2 = 0.16$), with accuracy numerically highest in VTC. To provide a sense of the subjective accuracy of recall-based reconstructions, we used the "most_similar" function of Word2Vec to generate examples of words that were most similar to the reconstructed semantic components. The "most_similar" function generates these words by computing the cosine similarity between the mean of the projection weight vectors (derived from the encoding model) and the vectors for each word in the Word2Vec model. Fig. 5 shows examples for images with varying degrees of recall-based reconstruction accuracy. Specifically, we pooled the reconstructed semantic component scores in VTC across all subjects and sessions, and then rank ordered these reconstructed scores by accuracy (match to the target scores). Examples of the "most similar" words are included for reconstructions that were in the top 1%, top 25%, and top 50%.

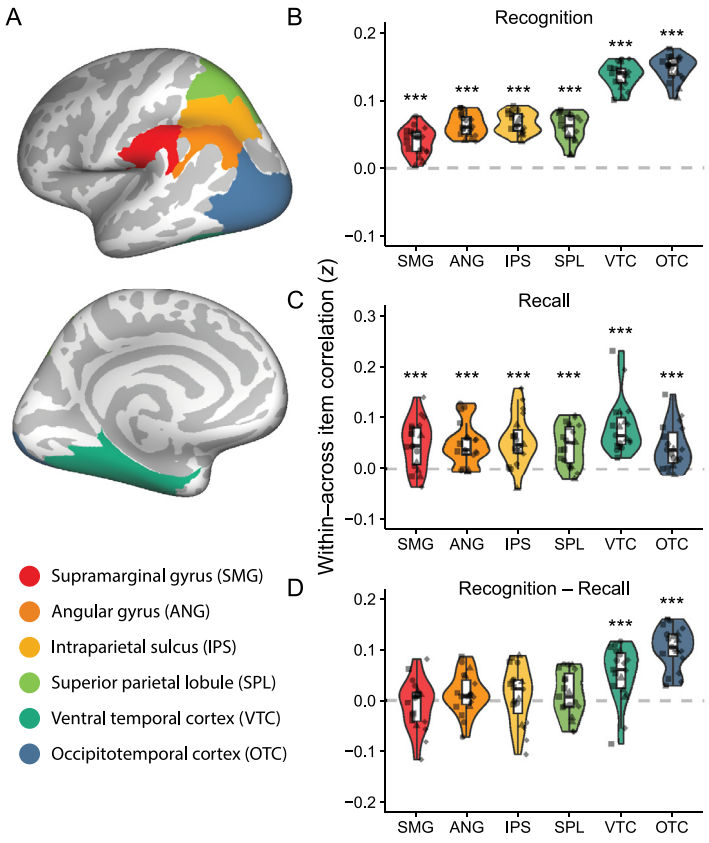

Fig. 4. Accuracy for fMRI-based reconstructions of semantic component scores. **A**. Anatomical regions of interest (ROIs), visualized on the inflated surface of an averaged template brain (from FreeSurfer). Top: left lateral view. Bottom: left medial view. **B**. Mean reconstruction accuracies of semantic component scores for each ROI based on encoding models trained and tested on recognition trials. Independent COCO annotations were used to define the 'actual' content of each image and semantic component scores from these annotations were then compared to semantic component scores reconstructed from fMRI activity patterns during the covert cued recall phase. Accuracy is expressed as within-item correlations – across-item correlations, with positive values (i.e., > 0) reflecting successful (item-specific) reconstructions. Accuracy was significantly above chance for all ROIs. **C**. As in **B**, but based on encoding models trained on recognition trials and tested on recall trials. Accuracy was significantly above chance for all ROIs. **D**. Difference in reconstruction accuracy for recognition vs. recall trials (**B** vs. **C**). Positive values reflect higher accuracy for recognition trials than recall trials. Only VTC and OTC exhibited significantly greater accuracy for recognition-based reconstructions than recall-based reconstructions. Also see Fig. S2 for similarity matrices of semantic component scores reconstructed from each ROI, separately for recognition and recall trials. Notes: dots represent data from individual sessions with each subject represented by a different shape; *** $p < 0.001$.

While all ROIs exhibited above-chance content reconstruction for both recognition-based and recall-based reconstructions, the difference between recognition- versus recall-based reconstructions markedly varied across ROIs, as reflected by an interaction between trial type (recognition, recall) and ROI ($F_{5,90} = 24.88$, $p < 0.0001$, $\eta_p^2 = 0.58$). Whereas content reconstruction accuracy was much higher for recognition than recall in VTC ($\chi_1^2 = 15.6$, $p < 0.0001$, $\beta = 0.052$, SE = 0.012) and OTC ($\chi_1^2 = 50.7$, $p < 0.0001$, $\beta = 0.103$, SE = 0.010), reconstruction accuracy in parietal regions did not significantly differ for recognition versus recall trials (ANG: $\chi_1^2 = 2.16$, $\beta = 0.014$, SE = 0.009; SMG: $\chi_1^2 = 0.58$, $\beta = -0.008$, SE = 0.011; SPL: $\chi_1^2 = 1.49$, $\beta = 0.012$, SE = 0.010; IPS: $\chi_1^2 = 0.74$, $\beta = 0.010$, SE = 0.011; p values > 0.140) (Fig. 4D). Thus, whereas VTC and OTC exhibited a clear 'preference' for images that were visually present (recognition trials), reconstructions from parietal regions were of comparable success when images were visually present (recognition trials) or entirely driven by memory (recall trials).

Given that the visual cortical ROIs (VTC and OTC) contained many more voxels than the parietal ROIs, one concern is that main effects of ROI and/or interactions by ROI may have been driven by differences in the number of voxels. To address this concern, we randomly subsampled voxels from the VTC and OTC ROIs for each subject so that they matched the mean size of the ANG ROI. Critically, the interaction between trial type (recognition, recall) and ROI remained significant ($F_{5,90} = 23.17$, $p < 0.001$, $\eta_p^2 = 0.56$). For recognition trials alone, the main effect of ROI was also significant ($F_{5,90} = 286.5$, $p < 0.0001$, $\eta_p^2 = 0.94$), driven by markedly higher accuracies for the visual ROIs. For recall trials alone, the main effect of ROI was no longer significant ($F_{5,90} = 1.05$, $p = 0.394$, $\eta_p^2 = 0.06$).

3.4. Similarity between reconstructed content and verbal descriptions of memories

In the preceding analyses, the target semantic content of each image was defined by image annotations that are part of the COCO image dataset. We next tested the degree to which semantic component scores reconstructed from the inverted fMRI encoding models (measured during the scanned cued recall task) matched the semantic component scores derived from subjects' *own verbal memory* of each image (measured during the post test) (Fig. 6A). As described for behavioral analysis of the verbal recall data (Fig. 3B), each subject's verbal recall of each image was translated into 30 semantic component scores. These target component scores could then be readily compared to (correlated with) the semantic component scores predicted from the inverted fMRI encoding models. Again, we found higher within- than across-item correlations in each of the visual and parietal ROIs (Fig. 6B) (VTC: rank = 58.18%; $\chi_1^2 = 49.7$, $p < 0.0001$, $\beta = 0.087$, SE = 0.012; OTC: rank = 54.70%; $\chi_1^2 = 21.2$, $p < 0.0001$, $\beta = 0.055$, SE = 0.011; ANG: rank = 55.22%; $\chi_1^2 = 14.1$, $p < 0.0001$, $\beta = 0.050$, SE = 0.013; SMG: rank = 54.20%; $\chi_1^2 = 12.3$, $p = 0.0004$, $\beta = 0.047$, SE = 0.001; SPL: rank = 53.23%; $\chi_1^2 = 7.4$, $p = 0.007$, $\beta = 0.038$, SE = 0.014; IPS: rank = 53.93%; $\chi_1^2 = 7.4$, $p = 0.007$, $\beta = 0.036$, SE = 0.013). Accuracy varied across ROIs (main effect of ROI: $F_{5,90} = 2.75$, $p = 0.024$, $\eta_p^2 = 0.13$), with accuracy numerically highest in VTC. These results confirm that the reconstructed semantic information from LPC and visual regions matched subjects' verbal descriptions of their memories.

While the preceding analysis confirms a match between verbal recall and reconstructed semantic component scores, an even stricter test is whether the semantic component scores reconstructed from a given subject's fMRI data more closely resembled the semantic component scores from *that subject's verbal recall* compared to semantic component scores from other subjects' verbal recall of the exact same images. To test this, we first calculated the Pearson correlations between the semantic component scores reconstructed from a given subject's inverted fMRI encoding model and the corresponding semantic component scores derived from that same subject's verbal recall (within-subject similarity). We then compared this within-subject similarity to across-subject similarity: the correlations between a given subject's reconstructed semantic component scores and the corresponding semantic component scores

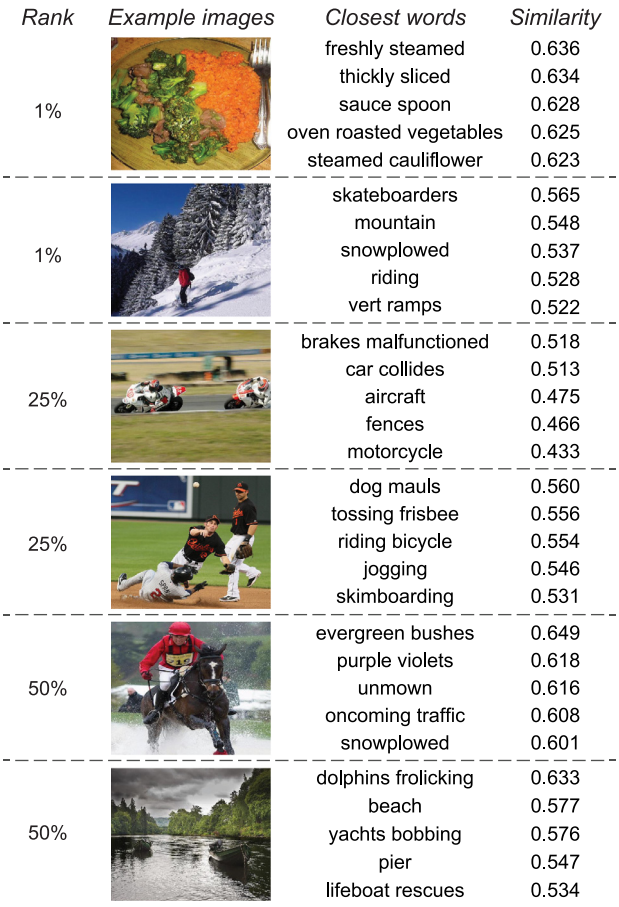

Fig. 5. Examples of reconstructed image content from VTC. (Left) Rank of the reconstruction accuracy pooled over all subjects and sessions. (Middle left) Example images being recalled. (Middle right) The top 5 most similar words and word combinations describing the semantic component scores reconstructed from VTC. The words were generated by the Word2Vec default 'most_similar' function. (right) Similarity scores between vectors corresponding to the content reconstructed from VTC and vectors of the Word2Vec most similar words.

derived from *different subjects'* verbal recall of the same images. It is important to emphasize that both of these measures were *within-item* correlations (i.e., they relate to the exact same images). If within-subject similarity exceeds across-subject similarity, this provides evidence for a subject-specific correspondence between fMRI-based reconstructions and verbal recall.

For each subject, session, and ROI we compared within-subject similarity to across-subject similarity in order to generate an accuracy score for each image. This image-specific accuracy score reflected the percentage of comparisons for which within-subject correlations were greater than across-subject correlations. For example, for a given image recalled by subject 1, the fMRI-based reconstructed semantic component scores would be correlated with the semantic component scores derived from verbal recall from subject 1 (within-subject similarity) and with the semantic component scores derived from verbal recall from subjects 2, 3 and 4 (across-subject similarity). If, for example, the within-subject correlation [$r(1,1)$] was greater than two of the three possible across-subject correlations [$r(1,2)$, $r(1,3)$, $r(1,4)$], this would correspond to an accuracy of 66.66% for that image. In this manner, the mean accuracy was computed for each subject, session, and ROI. Chance-level accuracy was 50% (i.e., by chance, within-subject similarity should exceed across-subject similarity 50% of the time). Strikingly, we observed above-chance accuracy—i.e., subject-specific reconstructions—in VTC (54.39%, $t_{18} = 2.90$, $p = 0.009$, Cohen's $d = 0.66$)—which was also the ROI that exhibited the highest recall-based reconstruction accuracy in each of the preceding analyses. Accuracy did not exceed chance in any of the other ROIs [OTC: $M = 49.21\%$, $t_{18} = -0.36$, $p = 0.720$, Cohen's $d = -0.08$; ANG: $M = 47.76\%$, $t_{18} = -0.96$, $p = 0.348$, Cohen's $d = -0.22$; SMG: $M = 51.36\%$, $t_{18} = 0.67$, $p = 0.512$, Cohen's $d = 0.20$; SPL: $M = 50.40\%$, $t_{18} = 0.24$, $p = 0.816$, Cohen's $d = 0.05$; IPS: $M = 48.11\%$, $t_{18} = -1.04$, $p = 0.310$, Cohen's $d = -0.24$].

To more explicitly emphasize subject-unique information expressed during verbal recall, we also repeated the analysis described above but only after subtracting out the semantic component scores from the COCO annotations from each subject's verbal recall component scores. In other words, we subtracted out 'normative' information from each subject's recall. With this approach, we again observed above-chance accuracy—i.e., subject-specific reconstructions—in VTC ($M = 55.26\%$, $t_{18} = 2.58$, $p = 0.019$, Cohen's $d = 0.59$), and also in SMG ($M = 53.95\%$, $t_{18} = 2.17$, $p = 0.043$, Cohen's $d = 0.50$). Accuracy was not above chance for any of the other ROIs (p's > 0.05).

Finally, to directly establish the degree to which subject-specific reconstructions depended on variability in verbal recall across subjects, we computed the mean correlations in verbal recall for each pair of subjects recalling the same images (see Tables S2 and S3). For each subject and session, we then median split the images in the recall session according to whether they were associated with high or low across-subject variability (i.e., low vs. high correlations). We then computed subject-specific reconstruction accuracy, as described above. Across ROIs, subject-specific reconstruction accuracy was significantly greater for high-variability images than low-variability images (main effect of variability: $F_{1,18} = 5.56$, $p = 0.030$, $\eta_p^2 = 0.24$; Fig. S3). Thus, the ability to measure subject-specific reconstructions benefitted from variability in how different subjects recalled the same image.

3.5. Across-subject reconstruction of recalled memories

Finally, we tested whether information 'learned' by the semantic encoding models (i.e., the mappings between voxel activity patterns and semantic component scores) successfully transferred across individuals. More specifically, we tested whether the contents of memory recall for each subject could be reconstructed using encoding models trained on data from independent subjects. To test this, we iteratively trained semantic encoding models using the recognition data from three of the four subjects and tested the model on recall trials from the held-out subject. That is, the weight matrix that was applied to each subject's fMRI activity patterns from the recall trials was entirely derived from independent subjects. We first tested content reconstruction accuracy by correlating the reconstructed component scores with component scores derived from the COCO annotations (as in Fig. 4C). Again, within-item similarity was compared against across-item similarity. Successful reconstruction (greater within-item similarity than across-item similarity) was observed in ANG (rank = 53.23%; $\chi_1^2 = 13.6$, $p = 0.0002$, $\beta = 0.033$, SE = 0.009), SPL (rank = 52.39%; $\chi_1^2 = 5.5$, $p = 0.020$, $\beta = 0.022$, SE = 0.010), IPS (rank = 52.15%; $\chi_1^2 = 5.4$, $p = 0.020$, $\beta = 0.022$, SE = 0.009), VTC (rank = 56.90%; $\chi_1^2 = 37.8$, $p < 0.0001$, $\beta = 0.056$, SE = 0.009), and OTC (rank = 53.98%; $\chi_1^2 = 10.5$, $p = 0.001$, $\beta = 0.027$, SE = 0.008) (Fig. 7A).

We next replicated this analysis with the only difference being that reconstructed component scores were correlated with component scores derived from each subject's (own) verbal recall (as in Fig. 6B). Again, within-item similarity was greater than across-item similarity in ANG (rank = 54.39%; $\chi_1^2 = 15.7$, $p < 0.0001$, $\beta = 0.049$, SE = 0.012), SPL (rank = 52.11%; $\chi_1^2 = 6.4$, $p = 0.011$, $\beta = 0.032$, SE = 0.013), IPS (rank = 52.96%; $\chi_1^2 = 7.7$, $p = 0.005$, $\beta = 0.034$, SE = 0.012), VTC (rank =55.06%; $\chi_1^2 = 24.1$, $p < 0.0001$, $\beta = 0.056$, SE = 0.011), and OTC (rank = 53.10%; $\chi_1^2 = 6.4$, $p = 0.011$, $\beta = 0.031$, SE = 0.012) (Fig. 7B). Interestingly, reconstruction of verbal recall content was not significantly different when the encoding models were trained/tested across subjects (Fig. 7B) vs. trained/tested within-subjects (Fig. 6B) (main effect of within- vs. across-subject encoding model: $F_{1,18} = 2.06$, $p = 0.168$,

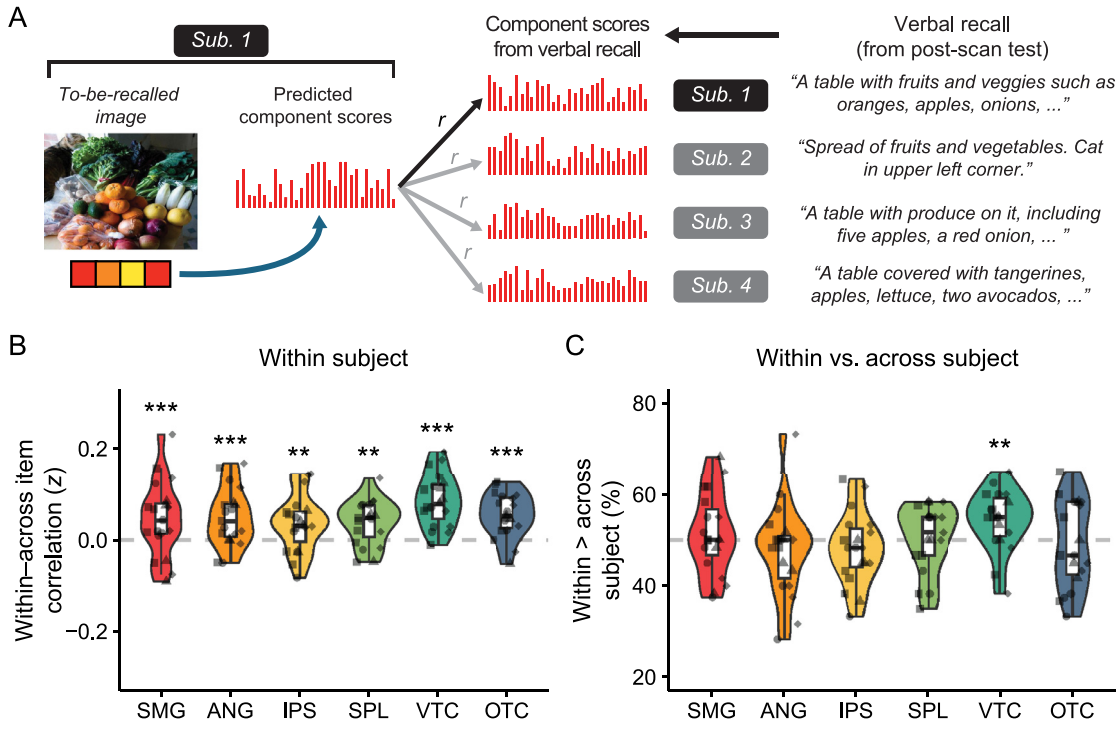

Fig. 6. Correspondence between semantic component scores reconstructed from fMRI vs. derived from verbal recall. **A**. Schematic of the analysis. For each to-be-recalled image for each subject, semantic component scores were reconstructed (predicted) from fMRI activity patterns using semantic encoding models trained on the recognition trials and tested on the recall trials. These reconstructed semantic component scores were then correlated with semantic component scores derived from subjects' verbal recall of the same image (measured during the post-scan overt cued recall test). **B**. Reconstruction accuracy as reflected by the difference between within-item vs. across-item correlations, with all correlations performed within-subject. Reconstruction accuracy was significantly above chance for all ROIs. **C**. Subject-specific reconstructions. To test for subject-specific (idiosyncratic) reconstructions, the semantic component scores reconstructed from one subject's fMRI data were correlated with semantic component scores generated from (i) the same subject's verbal recall data (e.g., Sub. 1 -> Sub. 1, black arrow, in **A**) and (ii) other subjects' verbal recall data of the exact same images (e.g., Sub- 1 -> Sub. 2, gray arrows, in **A**). Reconstructions were considered to contain subject-specific information when within-subject correlations were higher than the across-subject correlations. Data shown reflect the mean percentage of within-subject correlations that exceeded across-subject correlations. Accuracy was significantly above chance (dash line, 50%) only for VTC. Notes: dots represent data from individual sessions with each subject represented by a different shape; ** $p < 0.01$, *** $p < 0.001$.

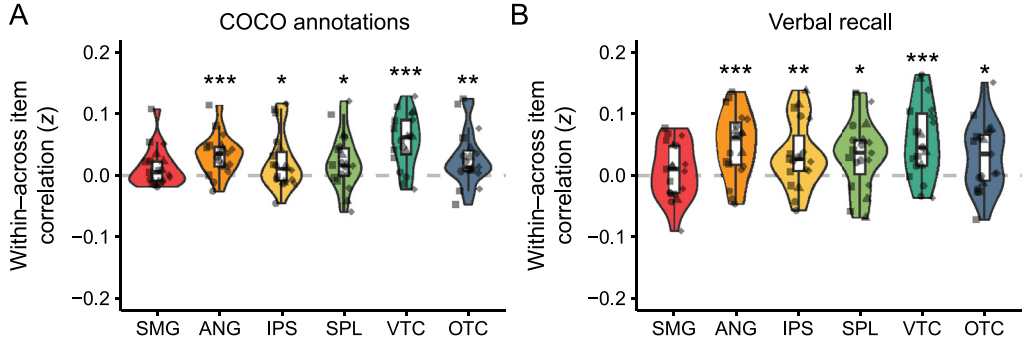

Fig. 7. Across-subject application of the semantic encoding models. For these analyses, the semantic encoding model was iteratively trained on recognition trials from 3 of the 4 subjects and then tested on recall trials from the held-out subject. **A**. Mean accuracy of reconstructed semantic component scores for each ROI based on comparison to semantic component scores derived from COCO annotations (within-item correlations – across-item correlations). **B**. Mean reconstruction accuracy for each ROI based on comparison to semantic component scores derived from verbal recall (within-item correlations – across-item correlations). For **B**, although the training/testing of the encoding models was performed across subjects, the covert recall trials used for reconstructing the semantic component scores and the verbal recall trials used for testing accuracy were always within the same subject. Notes: dots represent data from individual sessions with each subject represented by a different shape; ** $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, two tailed.

$\eta_p^2 = 0.10$). These findings provide evidence that, across subjects, the mappings between semantic content and fMRI activity patterns were shared to a degree that encoding models could be transferred to independent subjects to reconstruct the contents of memory recall.

4. Discussion

In the current study, we extracted high-level semantic features from complex natural images and modeled relationships between these semantic features and fMRI activity patterns using voxelwise encoding models. By inverting the encoding models, we tested whether the semantic content of retrieved memories could be reconstructed from evoked fMRI activity patterns. Using a multiple-session training procedure, we show that semantic content was successfully reconstructed from fMRI activity patterns in lateral parietal and visual cortices. Notably, however, reconstruction accuracy differed across these regions according to whether images were visually present (during recognition) or cued by

arbitrarily-associated abstract images (during recall). Whereas reconstruction accuracy in visual cortex was markedly lower when images were recalled from memory (recall trials) compared to when they were visually present (recognition trials), lateral parietal regions were relatively insensitive to this difference between trial types. Separately, by applying natural language processing methods to subjects' verbal recall data and projecting these recall data into the same feature space as the fMRI reconstructions, we also established that fMRI-based reconstructions reliably matched subjects' verbal recall data. In fact, reconstructions from ventral temporal cortex reflected idiosyncratic differences in how different subjects remembered the exact same image. Finally, we show that encoding models trained on a subset of subjects reliably transferred to held-out subjects, indicating that the mapping between fMRI activity patterns and semantic content was consistent enough across subjects to allow for across-subject reconstructions. Collectively, these findings provide important evidence for multidimensional memory representations in lateral parietal and visual cortices and establish the relevance of these neural representations to complex behavioral expressions of memory recall.

4.1. Reconstruction and recall of multidimensional memory representations

Numerous prior fMRI studies have demonstrated content-sensitivity of fMRI activity patterns in visual and lateral parietal cortices during memory retrieval (Favila et al., 2018; Kuhl et al., 2011; Kuhl and Chun 2014; Lee et al., 2019; Polyn et al., 2005; St-Laurent et al., 2015). However, the majority of this evidence comes from studies that have measured an objective, single stimulus property or dimension. For example, many studies have tested for decoding of visual category information (Kuhl et al., 2011; Polyn et al., 2005). Others have demonstrated an item-specific 'match' between fMRI activity patterns elicited during memory encoding and those elicited during memory retrieval (Favila et al., 2018; Kuhl and Chun 2014; Lee et al., 2019; St-Laurent et al., 2015). While the current findings also constitute evidence for item-specific representations (in that our analyses revealed differences between individual scene images), the key difference in the current study is that item-specific representations were 'built' by predicting and combining constituent features (Lee and Kuhl 2016; Naselaris et al., 2011). In fact, reconstructions were based on encoding models that were not trained on the to-be-reconstructed images (Brouwer and Heeger 2009). Thus, the stimulus-specific representations observed here cannot be explained by subjects generating verbal labels or stimulus-specific tags during encoding and then re-expressing that label/tag during recall.

The motivation for establishing multidimensional neural representations of memories is that these measures have the potential to capture the richness, subjectivity, and idioscynracies with which real world memories are recalled. Critically, however, validation of these neural representations requires behavioral expressions of memory that also capture the same richness, subjectivity, and idiosyncracies. Our solution to this problem was to use natural language processing methods that allowed our fMRI and behavioral data to be described using the same feature dimensions. Considering the behavioral recall data alone, text embeddings were highly sensitive to differences between images (Fig. 3B, C) validating the use of this method to characterize verbal recall data (Heusser et al., 2021; Song et al., 2021). Moreover, across visual and lateral parietal ROIs, there was strong correspondence between fMRI-based reconstructions and subjects' verbal recall (Fig. 6B), demonstrating that the multidimensional fMRI reconstructions aligned with the multidimensional expressions of verbal recall. Most strikingly, reconstructions generated from ventral temporal cortex were significantly more similar to subjects' own verbal recall compared to other subjects' verbal recall of exactly the same images. In other words, ventral temporal cortex reconstructions reflected subjective or idiosyncratic differences in how scene images were remembered. This effect is particularly notable when considering that there were no experimental pressures for subjects to use unique language or to differentiate their responses from other subjects. Thus, these methods may be even more sensitive to subjective/idiosyncratic information in experimental contexts where there are factors that promote memory differentiation (Favila et al., 2016; Hulbert and Norman 2015; Kim et al., 2017).

4.2. Reconstructions in lateral parietal cortex versus visual cortical areas

Not surprisingly, reconstructions from visual cortical areas (VTC, OTC) were markedly higher when images were visually present (recognition trials) compared to when they were visually absent (recall trials). In contrast, this fundamental distinction between trial types did not significantly influence reconstruction accuracy in LPC regions. Notably, several recent studies have specifically shown that, in contrast to visual cortical regions, LPC representations are stronger during memory recall compared to memory encoding or perception (Akrami et al., 2018; Favila et al., 2018, 2020; Long and Kuhl 2021; Xiao et al., 2017). While a definitive account of why LPC is biased towards memory-based information is not yet clear (Favila et al., 2020), the current findings provide additional support for a relative preference toward memory-based information in LPC. Here, however, we did not observe *stronger* (more accurate) LPC reconstructions during recall compared to recognition. That said, it is important to emphasize that recognition-based reconstructions were generated from models trained and tested on recognition trials whereas recall-based reconstructions were generated from models trained on recognition trials but *tested on recall trials*. Thus, a direct comparison of reconstruction accuracy for recall versus recognition trials is not an apples-to-apples comparison. Instead, the critical statistical comparison is the *relative* sensitivity of visual versus LPC regions to the difference in trial types. Indeed, this interaction was highly significant (Fig. 4C).

An obvious question raised by the current findings is whether recall reconstructions would be significantly better if the encoding model had been trained only on recall trials (Chen et al., 2017). This is particularly relevant for LPC where transfer from visual perception (recognition) to recall may be limited (Favila et al., 2018, 2020; Long and Kuhl 2021; Xiao et al., 2017). Enhancing overall reconstruction accuracy in LPC might also have revealed greater heterogeneity across LPC ROIs. In our study, however, training the encoding model only on recall trials was not feasible because the number of recall trials was relatively small (far fewer than the number of recognition trials). At a practical level, recall trials are much harder to include in large numbers because they depend on pre-training the paired associations (e.g., we used an extensive training procedure to ensure successful, vivid recall; Fig. 1). However, in an effort to address the potential concern of poor transfer from 'pure perception' trials to recall trials, we opted to pre-expose subjects to images in the recognition set such that the images used for model training were 'old' images. The sole rationale for the pre-exposure phase was that the semantic encoding models might better transfer to recall trials if the training trials had some memory component. Specifically, we reasoned that the representational format of a recall trial might be more similar to an 'old' recognition trial than to an entirely novel stimulus. While this thinking was informed by recent findings (Akrami et al., 2018; Favila et al., 2018, 2020; Long and Kuhl 2021; Xiao et al., 2017), it was not our intention—nor are we able—to test whether this design feature actually improved model transfer. That said, it does represent an interesting question that could be tested empirically in future studies.

While we observed evidence for idiosyncratic (subject-specific) relationships between fMRI-based reconstructions and verbal recall when considering reconstructions from VTC, we did not observe significant relationships for any of the LPC ROIs. On the one hand, this null result for LPC regions is surprising in light of evidence that memory reactivation in LPC has been associated with subjective qualities of memory recall (Bone et al., 2020; Johnson et al., 2015; Kuhl and Chun 2014; Richter et al., 2016). However, across analyses, reconstruction accuracy

was higher in VTC than in LPC ROIs, meaning there simply may have been better sensitivity within VTC to subtle differences in within-subject versus across-subject comparisons. As described above, it is possible that training the encoding models on recall trials (as opposed to recognition trials) might boost performance in LPC ROIs and thereby improve sensitivity to subject-specific differences. Indeed, we view this as a very interesting and reasonable possibility. Alternatively, it is possible that LPC preferentially expresses representational formats of retrieved memories that are relatively shared across subjects (Chen et al., 2017). Given that both of these are viable possibilities, we would caution against drawing conclusions based on the absence of significant subject-specific effects in the LPC ROIs. Instead, we view the significant results in VTC as a proof of concept that our methodological approach can be used to identify subject-specific idiosyncrasies in how complex images are remembered.

4.3. Semantic encoding models generalize across subjects

Although we deliberately used an extensive-sampling procedure to maximize the amount of within-subject training data available for the encoding models (see Fig. S4 for consideration of how the amount of within-subject training data influenced model performance), we also show that encoding models transferred quite well across subjects. Specifically, training encoding models using recognition trials from N-1 subjects allowed for successful recall-based reconstruction in held out subjects (Fig. 7). In fact, recall-based reconstruction was of comparable accuracy when using within-subject encoding models (Fig. 6B) vs. across-subject encoding models (Fig. 7B). This successful transfer across subjects indicates that the mapping between semantic components and fMRI activity patterns was shared—at least to some degree—across different individuals. Importantly, this shared mapping between semantic information and fMRI activity patterns is not at odds with our finding (or the idea) of idiosyncratic memory representations. For example, consider two individuals that had breakfast together. These individuals may have a common neural representation of the concept of coffee, and each of them may have had coffee for breakfast. However, when remembering breakfast, these individuals may differ in the degree to which the concept of coffee is a salient component of their memory and, therefore, in the degree to which the neural representation of coffee is activated when they remember breakfast. Thus, leveraging shared mappings (i.e., encoding models trained across different individuals) need not come at the expense of identifying idiosyncratic ways in which individuals perceive or remember their experiences (Finn et al., 2018; Finn et al. 2020).

More generally, the success of the across-subject encoding models has two main implications. First, this finding adds to a growing body of evidence that, even for complex and naturalistic stimuli (e.g., movies and narratives), there is a surprising degree of consistency across individuals in how these stimuli are represented in patterns of neural activity (Chen et al., 2017; Finn et al., 2018; Hasson et al., 2004; Zadbood et al., 2017). Second, leveraging across-subject encoding models could have substantial practical—and theoretical—advantages. For example, as noted above, it was not feasible in our experimental paradigm for each subject to learn and recall thousands of different scenes (due to the training time it would require and the deterioration in memory performance that would be expected with such a large memory set). However, it is much more feasible to obtain thousands of recall trials *across subjects*. Thus, some analyses which are impractical—or that would be data starved—within subjects, might become feasible if across-subject models are leveraged. Moreover, a single well-powered training data set could potentially be applied to many distinct test sets. Finally, it is also notable that here, we only aligned across-subject data in anatomical space. Additional gains in across-subject transfer may well be realized by aligning data in a common high-dimensional functional space (Chen et al., 2015; Haxby et al., 2011; Haxby et al., 2020).

5. Conclusions

To summarize, we used inverted semantic encoding models applied to fMRI data to reconstruct multidimensional content in natural scene images as they were viewed and recalled from memory. We found that visual and lateral parietal cortices supported successful reconstructions both when viewing and recalling images. However, whereas lateral parietal reconstructions were relatively insensitive to whether images were viewed or recalled from memory, visual cortical reconstructions were markedly lower for recalled versus viewed images. Additionally, by applying natural language processing methods to behavioral measures of memory recall, we show that fMRI-based reconstructions of recalled content matched subjects' verbal recall and that fMRI-based reconstructions even reflected idiosyncratic qualities of subjects' recall. Finally, we show that semantic encoding models reliably transferred across individuals, allowing for successful reconstruction of a given subject's memory using encoding models trained on entirely different individuals. Collectively, these findings provide important evidence characterizing multidimensional memory representations and linking their neural and behavioral expressions.

Declaration of Competing Interest

The authors declare no competing interests.

Credit authorship contribution statement

Yingying Wang: Conceptualization, Data curation, Formal analysis, Investigation, Visualization, Writing – original draft, Writing – review & editing. **Hongmi Lee:** Conceptualization, Writing – original draft, Writing – review & editing. **Brice A. Kuhl:** Conceptualization, Funding acquisition, Project administration, Visualization, Writing – original draft, Writing – review & editing.

Data availability

Data is currently being prepared for a public repository and will be posted before manuscript publication.

Acknowledgments

The research described here was supported by NSF CAREER Award BCS -1752921 and NIH-NINDS R01 NS107727 to B.A.K.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2023.120222.

References

Akrami, A., Kopec, C.D., Diamond, M.E., Brody, C.D., 2018. Posterior parietal cortex represents sensory history and mediates its effects on behaviour. Nature 554 (7692), 368–372. doi:10.1038/nature25510.

Bone, M.B., Ahmad, F., Buchsbaum, B.R., 2020. Feature-specific neural reactivation during episodic memory. Nat. Commun. 11 (1), 1945. doi:10.1038/s41467-020-15763-2.

Bonnici, H.M., Richter, F.R., Yazar, Y., Simons, J.S., 2016. Multimodal feature integration in the angular gyrus during episodic and semantic retrieval. J. Neurosci. 36 (20), 5462–5471. doi:10.1523/JNEUROSCI.4310-15.2016.

Brouwer, G.J., Heeger, D.J., 2009. Decoding and reconstructing color from responses in human visual cortex. J. Neurosci. 29 (44), 13992–13993. doi:10.1523/JNEUROSCI.3577-09.2009.

Chen, J., Leong, Y.C., Honey, C.J., Yong, C.H., Norman, K.A., Hasson, U., 2017. Shared memories reveal shared structure in neural activity across individuals. Nat. Neurosci. 20 (1), 115–125. doi:10.1038/nn.4450.

Cooper, R.A., Ritchey, M., 2019. Cortico-hippocampal network connections support the multidimensional quality of episodic memory. Elife 8, e45591. doi:10.7554/eLife.45591, doi.

Cowen, A.S., Chun, M.M., Kuhl, B.A., 2014. Neural portraits of perception: reconstructing face images from evoked brain activity. Neuroimage 94, 12–22. doi:10.1016/j.neuroimage.2014.03.018.

Danker, J.F., Anderson, J.R., 2010. The ghosts of brain states past: remembering reactivates the brain regions engaged during encoding. Psychol. Bull. 136 (1), 87–102. doi:10.1037/a0017937.

Esteban, O., Birman, D., Schaer, M., Koyejo, O.O., Poldrack, R.A., Gorgolewski, K.J., 2017. MRIQC: advancing the automatic prediction of image quality in MRI from unseen sites. PLoS One 12 (9), e0184661. doi:10.1371/journal.pone.0184661.

Esteban, O., Markiewicz, C.J., Blair, R.W., Moodie, C.A., Ilkay Isik, A., Erramuzpe, A., Kent, J.D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S.S., Wright, J., Durnez, J., Poldrack, R.A., Gorgolewski, K.J., 2019. FMRIPrep: a robust preprocessing pipeline for functional MRI. Nat. Methods 16 (1), 111–116. doi:10.1038/s41592-018-0235-4.

Ester, E.F., Sprague, T.C., Serences, J.T., 2015. Parietal and frontal cortex encode stimulus-specific mnemonic representations during visual working memory. Neuron 87 (4), 893–905. doi:10.1016/j.neuron.2015.07.013.

Favila, S.E., Chanales, A.J.H., Kuhl, B.A., 2016. Experience-dependent hippocampal pattern differentiation prevents interference during subsequent learning. Nat. Commun. 7 (1). doi:10.1038/ncomms11066.

Favila, S.E., Lee, H., Kuhl, B.A., 2020. Transforming the concept of memory reactivation. Trends Neurosci. doi:10.1016/j.tins.2020.09.006, S0166223620302137.

Favila, S.E., Samide, R., Sweigart, S.C., Kuhl, B.A., 2018. Parietal representations of stimulus features are amplified during memory retrieval and flexibly aligned with top-down goals. J. Neurosci. 38 (36), 7809–7821. doi:10.1523/JNEUROSCI.0564-18.2018.

Finn, E.S., Corlett, P.R., Chen, G., Bandettini, P.A., Constable, R.T., 2018. Trait paranoia shapes inter-subject synchrony in brain activity during an ambiguous social narrative. Nat. Commun. 9 (1), 2043. doi:10.1038/s41467-018-04387-2.

Finn, E.S., Glerean, E., Khojandi, A.Y., Nielson, D., Molfese, P.J., Handwerker, D.A., Bandettini, P.A., 2020. Idiosynchrony: from shared responses to individual differences during naturalistic neuroimaging. Neuroimage 215, 116828. doi:10.1016/j.neuroimage.2020.116828.

Gilmore, A.W., Nelson, S.M., McDermott, K.B., 2015. A parietal memory network revealed by multiple MRI methods. Trends Cogn. Sci. (Regul. Ed.) 19 (9), 534–543. doi:10.1016/j.tics.2015.07.004.

Gilmore, A.W., Quach, A., Kalinowski, S.E., Gotts, S.J., Schacter, D.L., Martin, A., 2021. Dynamic content reactivation supports naturalistic autobiographical recall in humans. J. Neurosci. 41 (1), 153–166. doi:10.1523/JNEUROSCI.1490-20.2020.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., Malach, R., 2004. Intersubject synchronization of cortical activity during natural vision. Science 303 (5664), 1634–1640. doi:10.1126/science.1089506.

Haxby, J.V., Swaroop Guntupalli, J., Connolly, A.C., Halchenko, Y.O., Conroy, B.R., Ida Gobbini, M., Hanke, M., Ramadge, P.J., 2011. A common, high-dimensional model of the representational space in human ventral temporal cortex. Neuron 72 (2), 404–416. doi:10.1016/j.neuron.2011.08.026.

Haxby, J.V., Swaroop Guntupalli, J., Nastase, S.A., Feilong, M., 2020. Hyperalignment: modeling shared information encoded in idiosyncratic cortical topographies. Elife 9, e56601. doi:10.7554/eLife.56601.

Heusser, A.., Fitzpatrick, P.C., Manning, J.R., 2021. Geometric models reveal behavioural and neural signatures of transforming experiences into memories. Nat. Hum. Behav. doi:10.1038/s41562-021-01051-6.

Hulbert, J.C., Norman, K.A., 2015. Neural differentiation tracks improved recall of competing memories following interleaved study and retrieval practice. Cerebr. Cortex 25 (10), 3994–4008. doi:10.1093/cercor/bhu284.

Humphreys, G.F., Ralph, M.A.L., Simons, J.S., 2021. A unifying account of angular gyrus contributions to episodic and semantic cognition. Trends Neurosci. 44 (6), 452–463. doi:10.1016/j.tins.2021.01.006.

Huth, A.G., Heer, W.A.d., Griffiths, T.L., Theunissen, F.E., Gallant, J.L., 2016. Natural speech reveals the semantic maps that tile human cerebral cortex. Nature 532 (7600), 453–458. doi:10.1038/nature17637.

Johnson, M.K., Kuhl, B.A., Mitchell, K.J., Ankudowich, E., Durbin, K.A., 2015. Age-related differences in the neural basis of the subjective vividness of memories: evidence from multivoxel pattern classification. Cognit., Affect. Behav. Neurosci. 15 (3), 644–661. doi:10.3758/s13415-015-0352-9.

Kay, K.N., Naselaris, T., Prenger, R.J., Gallant, J.L., 2008. Identifying natural images from human brain activity. Nature 452 (7185), 352–355. doi:10.1038/nature06713.

Kim, G., Norman, K.A., Turk-Browne, N.B., 2017. Neural differentiation of incorrectly predicted memories. J. Neurosci. 37 (8), 2022–2031. doi:10.1523/JNEUROSCI.3272-16.2017.

Kok, P., Rait, L.I., Turk-Browne, N.B., 2020. Content-based dissociation of hippocampal involvement in prediction. J. Cogn. Neurosci. 32 (3), 527–545. doi:10.1162/jocn_a_01509.

Kuhl, B.A., Chun, M.M., 2014. Successful remembering elicits event-specific activity patterns in lateral parietal cortex. J. Neurosci. 34 (23), 8051–8060. doi:10.1523/JNEUROSCI.4328-13.2014.

Kuhl, B.A., Rissman, J., Chun, M.M., Wagner, A.D., 2011. Fidelity of neural reactivation reveals competition between memories. Proc. Natl. Acad. Sci. 108 (14), 5903–5908. doi:10.1073/pnas.1016939108.

Lee, H., Kuhl, B.A., 2016. Reconstructing perceived and retrieved faces from activity patterns in lateral parietal cortex. J. Neurosci. 36 (22), 6069–6082. doi:10.1523/JNEUROSCI.4286-15.2016.

Lee, H., Samide, R., Richter, F.R., Kuhl, B.A., 2019. Decomposing parietal memory reactivation to predict consequences of remembering. Cerebr. Cortex 29 (8), 3305–3318. doi:10.1093/cercor/bhy200.

Chen, P.H.C., Chen, J., Yeshurun, Y., Hasson, U., Haxby. J., Ramadge, P.J., 2015. "A reduced-dimension FMRI shared response model." in *Advances in Neural Information Processing Systems*. Vol. 28, edited by C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett. Curran Associates, Inc.

Lin, T.-Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Zitnick, C.L., Dollár, P., 2015. "Microsoft COCO: common Objects in Context." *ArXiv:1405.0312 [Cs]*.

Long, N.M., Kuhl, B.A., 2021. Cortical representations of visual stimuli shift locations with changes in memory states. Curr. Biol. 31 (5), 1119–1126. doi:10.1016/j.cub.2021.01.004, e5.

Naselaris, T., Allen, E., Kay, K., 2021. Extensive sampling for complete models of individual brains. Curr. Opin. Behav. Sci. 40, 45–51. doi:10.1016/j.cobeha.2020.12.008.

Naselaris, T., Kay, K.N., Nishimoto, S., Gallant, J.L., 2011. Encoding and decoding in FMRI. Neuroimage 56 (2), 400–410. doi:10.1016/j.neuroimage.2010.07.073.

Naselaris, T., Olman, C.A., Stansbury, D.E., Ugurbil, K., Gallant, J.L., 2015. A voxel-wise encoding model for early visual areas decodes mental images of remembered scenes. Neuroimage 105, 215–228. doi:10.1016/j.neuroimage.2014.10.018.

Nguyen, M., Vanderwal, T., Hasson, U., 2019. Shared understanding of narratives is correlated with shared neural responses. Neuroimage 184, 161–170. doi:10.1016/j.neuroimage.2018.09.010.

Polyn, S.M., Natu, V.S., Cohen, J.D., Noman, K.A., 2005. Category-specific cortical activity precedes retrieval during memory search. Science 310 (5756), 1963–1966. doi:10.1126/science.1117645.

Richter, F.R., Cooper, R.A., Bays, P.M., Simons, J.S., 2016. Distinct neural mechanisms underlie the success, precision, and vividness of episodic memory. Elife 5, e18260. doi:10.7554/eLife.18260.

Rissman, J., Wagner, A.D., 2012. Distributed representations in memory: insights from functional brain imaging. Annu. Rev. Psychol. 63 (1), 101–128. doi:10.1146/annurev-psych-120710-100344.

Rugg, M.D., King, D.R., 2018. Ventral lateral parietal cortex and episodic memory retrieval. Cortex 107, 238–250. doi:10.1016/j.cortex.2017.07.012.

Rugg, M.D., Vilberg, K.L., 2013. Brain networks underlying episodic memory retrieval. Curr. Opin. Neurobiol. 23 (2), 255–260. doi:10.1016/j.conb.2012.11.005.

Smith, S.M., Brady, J.M., 1997. SUSAN—a new approach to low level image processing. Int. J. Comput. Vis. 23 (1), 45–78.

Song, H., Finn, E.S., Rosenberg, M.D., 2021. Neural signatures of attentional engagement during narratives and its consequences for event memory. Proc. Natl. Acad. Sci. 118 (33), e2021905118. doi:10.1073/pnas.2021905118.

Sprague, T.C., Ester, E.F., Serences, J.T., 2016. Restoring latent visual working memory representations in human cortex. Neuron 91 (3), 694–707. doi:10.1016/j.neuron.2016.07.006.

St-Laurent, M., Abdi, H., Buchsbaum, B.R., 2015. Distributed patterns of reactivation predict vividness of recollection. J. Cogn. Neurosci. 27 (10), 2000–2018. doi:10.1162/jocn_a_00839.

Walther, A., Nili, H., Ejaz, N., Alink, A., Kriegeskorte, N., Diedrichsen, J., 2016. Reliability of dissimilarity measures for multi-voxel pattern analysis. Neuroimage 137, 188–200. doi:10.1016/j.neuroimage.2015.12.012.

Xiao, X., Dong, Q., Gao, J., Men, W., Poldrack, R.A., Xue, G., 2017. Transformed neural pattern reinstatement during episodic memory retrieval. J. Neurosci. 37 (11), 2986–2998. doi:10.1523/JNEUROSCI.2324-16.2017.

Yu, Q., 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. doi:10.1016/j.neuroimage.2017.05.055.

Zadbood, A., Chen, J., Leong, Y.C., Norman, K.A., Hasson, U., 2017. How we transmit memories to other brains: constructing shared neural representations via communication. Cerebr. Cortex 27 (10), 4988–5000. doi:10.1093/cercor/bhx202.