



Individual-specific and shared representations during episodic memory encoding and retrieval



Xiaoqian Xiao^{a,b,1}, Yu Zhou^{a,1}, Jing Liu^a, Zhifang Ye^a, Li Yao^b, Jiakai Zhang^b, Chuansheng Chen^c, Gui Xue^{a,*}

^a State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute of Brain Research, Beijing Normal University, Beijing, 100875, PR China

^b School of Informatics, Beijing Normal University, Beijing, 100875, PR China

^c Department of Psychological Science, University of California, Irvine, 92697, USA

ABSTRACT

Although human memories seem unique to each individual, they are shared to a great extent across individuals. Previous studies have examined, separately, subject-specific and cross-subject shared representations during memory encoding and retrieval, but how shared memories are formed from individually encoded representations is not clearly understood. Using a unique fMRI design involving memory encoding and retrieval, and representational similarity analysis to link representations from different individuals, brain regions, and processing stages, the current study revealed that distributed brain regions showed both subject-specific and shared neural representations during both memory encoding and retrieval. Furthermore, different brain regions showed stage-specific representational strength, with the visual cortex showing greater unique and shared representations during encoding, whereas the left angular gyrus showing greater unique and shared representations during retrieval. The neural representations during encoding were transformed during retrieval, as shown by smaller cross-subject encoding-retrieval similarity (ERS) than cross-subject similarity either during encoding or during retrieval. This cross-subject and cross-stage similarity was found both within and across regions, with strong pattern similarity between the encoded representation in VVC and the retrieved representation in the angular gyrus. Simulation analysis further suggested that these patterns could be achieved by incorporating stage-specific representational strength, and cross-region reinstatement from encoding to retrieval, but not by a common transformation from encoding to retrieval across subjects. Together, our results shed light on how memory representations are encoded and transformed to maintain individual characteristics and at the same time to create shared representations to facilitate interpersonal communication.

1. Introduction

Theoretical discussions and behavioral studies have long considered memory as personal experiences that are unique to each individual. This intuitive perspective is also supported by empirical evidence. Previous studies, including neural imaging studies, have found that in response to the same stimuli and task instruction, subjects showed significant individual variances regarding how many items, which items, and the precise details associated with each item that can be remembered (Kirchhoff, 2009; Loftus et al., 1992; Miller et al., 2002; Munday, 1985; Shapiro and Penrod, 1986).

Nevertheless, memory is also shared across individuals, forming strong collective memory at different levels (Hasson et al., 2004, 2008; Jääskeläinen et al., 2008; Wilson et al., 2008). Supporting the notion of shared memory experience, existing studies revealed shared neural representations (i.e., the distributed pattern of neural activities associated with given stimuli (Vilarroya, 2017)) during memory retrieval. For example, Chen et al. (2017) reported that after viewing the same movie,

the cross-subject similarity of the neural activity during free recall of the movie details was significantly above the baseline (Chen et al., 2017). In another study, Zadbood et al. (2017) examined how memories transmitted from one person to another person and found shared response patterns across participants when watching, recalling, and listening to spoken descriptions of movie scenes (Zadbood et al., 2017).

Other studies have also reported unique and shared representations in perception (Golland et al., 2007; Hasson et al., 2004, 2010; Jääskeläinen et al., 2008; Kauppi et al., 2010) and semantic processing (Chadwick et al., 2016; Charest et al., 2014; Huth et al., 2016; Wilson et al., 2008). Moreover, individual participants' neural organization of semantic memory was found to better predict their episodic memory performance than would the group-averaged memory representations (Chadwick et al., 2016).

The above studies, however, did not simultaneously examine the unique and shared representations within and across processing stages and brain regions, leaving several important questions unaddressed. First, are the retrieved representations more likely to be shared than are

* Corresponding author.

E-mail address: gxue@bnu.edu.cn (G. Xue).

¹ These two authors contributed equally to this paper.

the encoded representations, or vice versa? It is conceivable that the encoded representations might be shared to a greater extent because subjects perceive the same stimulus whereas the retrieved representations are more subjective and uncontrollable. Alternatively, it is also possible that the encoded representations are more unique because of their high fidelity, whereas the retrieved representations are more abstract and hence are probably more likely to be shared by subjects.

Second, although memory retrieval involves the reinstatement of perceived representations, this pattern of reinstatement is a constructive process (Schacter et al., 1998; for a review, see Xue, 2018). The question thus is how representations are transformed from encoding to retrieval and is the transformation process similar across subjects? Memory transformation can be characterized by the change in representational content or format, as reflected by the reduced pattern similarity within brain regions between encoding and retrieval, as well as by the shift in brain regions that showing item-specific representations between the two memory stages, given that different brain regions contain different aspects of the representations. For example, in our previous study (Xiao et al., 2017), we found that during encoding the visual cortex showed strong item-specific representations, but during retrieval higher-level brain regions such as the angular gyrus showed strong item-specific representations. Existing studies have shown that the representation in the VVC contains perceptual details, whereas that in the AG is identity-specific and invariant to viewpoints (Jeong and Xu, 2016), and is modulated by semantic similarity (Ye et al., 2016). These results thus suggest a change in representational format between encoding and retrieval.

Furthermore, we compared the encoding-retrieval similarity (i.e., ERS) with pattern similarity during encoding (i.e., encoding-encoding similarity, EES) and retrieval (retrieval-retrieval similarity, RRS) (Xiao et al., 2017). The results revealed that the ERS was smaller than EES in the ventral visual cortex (VVC) and RRS in angular gyrus (AG), suggesting that the neural representational patterns during retrieval were transformed from those during encoding. In light of this stage-specific representations, we further tested and verified the hypothesis that the encoded representation in the VVC could be transformed and reinstated

in associative regions such as the angular gyrus, suggesting possible cross-region reinstatement of memory representations (Xiao et al., 2017).

Besides the within-subject transformation, another study by Chen et al. (2017) compared the cross-subject similarity in retrieval (i.e., RRS) with cross-subject encoding-retrieval similarity (ERS). They found greater cross-subject RRS than cross-subject ERS, again suggesting that the encoded representation was transformed into common representations shared by different individuals during retrieval. Their simulation analysis suggested that this pattern could be achieved via a common transformation from encoding to retrieval across subjects (Chen et al., 2017). Nevertheless, this shared transformation can conceivably result in greater shared representation during retrieval than encoding, but this possibility was not examined in that study. In addition, the mechanism for this shared representation during retrieval is not clear.

The present study examined subject-specific and shared episodic memory representations during encoding and retrieval. Furthermore, we examined whether the two types of memory transformation, i.e., within-region memory transformation and cross-region memory reinstatement, were shared by all subjects or were specific to each individual. We adopted a design where subjects were required to study and retrieve word-picture associations. Importantly, each picture was associated with two different cues in separate runs. To capture the individual-specific and shared episodic memory representations, we generated for each subject two representational similarity matrices (RSMs) (Kriegeskorte et al., 2008) for encoding, and two matrices for retrieval, each containing one repetition of the pictures. These RSMs allowed us to compare the representations across processing stages, brain regions, and subjects (Fig. 1), and at the same time to overcome the anatomical differences between subjects and brain regions (Chen et al., 2017). To examine the within-region memory transformation, we compared cross-subject ERS with cross-subject EES and cross-subject RRS. To examine the cross-region memory transformation, we examined whether there was significant above-chance cross-subject, cross-region ERS. Finally, we conducted simulations to investigate whether the shared representations during retrieval were a result of a common transformation, or through stage-specific representation and cross-region memory transformation.

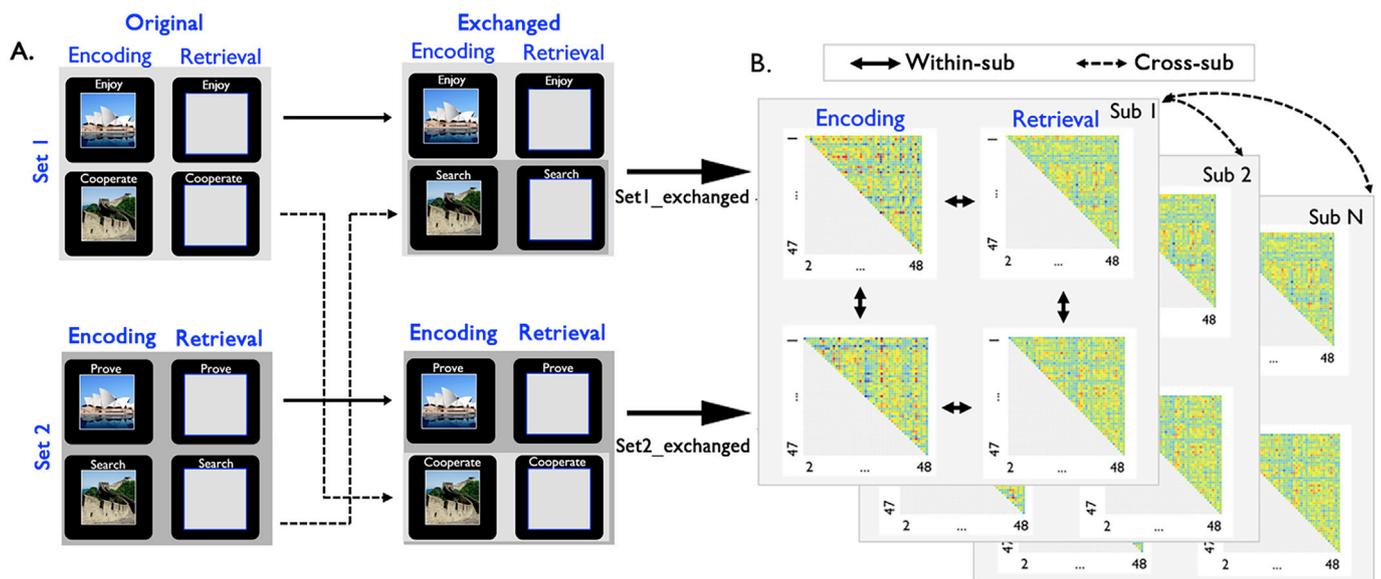


Fig. 1. Experiment paradigm and diagram of analytical strategies. A. Each picture was used twice and paired with two different word cues (set 1 and set 2). Half of the pictures in set 1 were exchanged with the same pictures in set 2 to form two new picture sets containing the same pictures arranged in the same order. B. Calculation of representational space; For each subject, we generated the representational spaces separately for the two data sets by processing stage (encoding and retrieval). The representational space was calculated as the correlations between each target's neural spatial pattern and that for each of other targets from the same data set. Only the correlation coefficients in the upper triangle were used so only one correlation coefficient for each pair of materials was used. The similarity (Pearson correlation) for representational spaces was calculated within and across subjects. Using the averaged representational similarity matrix, we calculated the similarity across processing stages, brain regions, and individuals.

2. Materials and methods

2.1. Subjects and experiment design

Detailed information regarding the subjects, experimental design, behavioral data analysis, and functional magnetic resonance imaging (fMRI) data preprocessing can be found in our previously published paper (Xiao et al., 2017). Briefly, 20 healthy subjects (11 males, mean age = 20.95 ± 1.96 years, range of 18–25 years) participated in this study. Informed written consent was obtained from the participants before the experiments. The fMRI study was approved by the institutional review board of Peking University and the State Key Laboratory of Cognitive Neuroscience and Learning at Beijing Normal University in China.

Subjects studied and retrieved 96 word-picture pairs. Pictures were 48 well-known scenes, including 32 architectures (half from China and the other half from abroad) and 16 natural landscapes (half depicting water landscapes and the other half depicting terrestrial landscapes). The cues were 96 two-character Chinese verbs. Each picture was associated with two different word cues (cue 1 and cue 2). Words and pictures were randomly paired across subjects.

One day before the fMRI scan, subjects were trained to be familiar with all pictures and then to memorize all 96 word-picture associations. The overtraining paradigm was used to make sure subjects could recall the visual details of the pictures during retrieval. It included three stages: picture familiarization, self-paced word-picture association learning, and recall. The training ended once subjects could correctly report the category and four details of the picture associated with each cue. Subjects spent about 2 h in this session.

During the fMRI scan, subjects were asked to restudy the word cue-picture pairs (the encoding run) and then to recall the visual details of the pictures associated with the word cues (the retrieval run) (Fig. 1A). A slow event-related design (16 s for each trial) was used to obtain better estimates of single-trial BOLD responses for both encoding and retrieval. During encoding, each trial started with 4 s presentation of the word cue-picture association, and subjects were asked to try to remember as many details as they could (i.e., the encoding stage). The frame of the picture then turned green for 2 s (i.e., the category judgment stage), during which subjects were asked to judge the category of the picture but held their response until the frame turned red and the response labels showed on the screen. The response labels representing the four possible picture categories, i.e., domestic architectures, foreign architectures, water landscapes, and terrestrial landscapes, were introduced to prevent subjects from planning motor response during the category judgment stage. Specifically, each response key/button corresponded to one of the four label locations (lined up from left to right) instead of the picture category, and the order of the four category labels presented on the screen was randomized across trials. Subjects had another 2 s to make the response (i.e., the response stage) according to the response labels. To prevent further processing of the word cue-picture association, subjects were asked to do a perceptual orientation judgment task for 8 s. During this task, an arrow pointing either left or right was presented on the screen, and subjects were asked to judge the orientation of the arrow as quickly as possible. A self-paced procedure was used to make the task engaging.

The retrieval stage was similar to the encoding stage, except that, for the first 4 s, only the retrieval cue was presented, and subjects were asked to retrieve the visual details of the associated picture (Fig. 1A). For both the encoding and retrieval stages, subjects were told explicitly to focus on the details rather than simply the category of the pictures. The 48 pictures were divided into two groups (to keep the scanning time of each run within 7 min). For each group, each picture was paired with cue set 1 in one encoding-retrieval session and then with cue set 2 in the next session. In each run, the word-picture pairs were presented in random order. In total, there were four encoding-retrieval sessions.

After the scan, subjects finished an oral test outside the scanner to report the details of the picture associated with each cue. Pictures correctly recalled with more than four different kinds of details (e.g.,

objects in the scene, color, structure, name, and so on) were scored as remembered with details.

2.2. fMRI image data acquisition and preprocessing

Scanning was carried out in the MRI Center at Peking University, using a MAGNETOM Prisma 3.0T MRI scanner (Siemens Healthcare, Erlangen, Germany) with a 64-channel head-neck coil. High-resolution functional images were acquired using a prototypical simultaneous multi-slice EPI sequence (FOV = $224 \text{ mm} \times 224 \text{ mm}$; matrix = 112×112 ; slice thickness = 2 mm; TR/TE/ θ = 2000 ms/30 ms/90°, slice acceleration factor = 2). Sixty-four contiguous axial slices parallel to the AC-PC line were obtained to cover the whole cerebrum and partial cerebellum. High-resolution structural images using a three-dimensional, T1-weighted, magnetization-prepared rapid gradient-echo (MPRAGE) sequence were acquired for the whole brain (FOV = $256 \text{ mm} \times 256 \text{ mm}$; matrix = 256×256 ; slice thickness = 1 mm; TR/TE/ θ = 2530 ms/2.98 ms/7°). A high-resolution T2-weighted image was also acquired using a T2-SPACE sequence for use in MTL segmentation. The image plane was perpendicular to the main hippocampal axis and covered the whole MTL region (FOV = $220 \text{ mm} \times 220 \text{ mm}$; matrix = 512×512 ; slice thickness = 1.5 mm; TR/TE/ θ = 13150 ms/82 ms/150°, 60 slices).

Image preprocessing and statistical analyses were carried out using FEAT (fMRI Expert Analysis Tool) version 5.98. The first 10 images from each run were automatically discarded by the scanner to allow scanner equilibrium. Functional images were realigned, and temporally filtered (nonlinear high pass filter with a 90s cut-off). The EPI images were first registered to the first volume of the fifth run and then registered to the MPRAGE structural volume using Advanced Normalization Tools (ANTs) (Avants et al., 2011). Registration from structural images to the standard space was further refined using ANTs nonlinear registration SyN (Klein et al., 2009). All fMRI analyses were performed in each subject's native space and then transformed into standard space for group analysis.

2.3. Single-trial response estimate

The GLM models were separately created for each of the 96 encoding and retrieval trials to estimate the single-trial response. A least-square single method was used, where the target trial was modeled as one experimental variable (EV), and all other trials were modeled as another EV (Mumford et al., 2012). The trial was modeled at its presentation time, convolved with a canonical hemodynamic response function (double gamma). The whole first 4 s were modeled during both encoding and retrieval for each trial. The t-statistic was used for representation similarity analysis to increase the reliability by noise normalization (Walther et al., 2016).

2.4. Region of interests (ROI) definition

Following previous studies and our previous findings, 8 ROIs were defined based on the Harvard-Oxford probabilistic atlas (threshold at 25% probability), including the bilateral ventral visual cortex (VVC, containing ventral lateral occipital cortex, occipital fusiform, occipital temporal fusiform, and parahippocampus) (Danker et al., 2017; Wing et al., 2015), angular gyrus (AG) (Kuhl and Chun, 2014), the inferior frontal cortex (IFG) (Ritchey et al., 2013), medial prefrontal cortex (mPFC) (Guise and Shapiro, 2017), and posterior medial cortex (PMC) (Chen et al., 2017) (Fig. 2A). The hippocampus was not included, as we did not find item-specific representation in this region in our previous study (Xiao et al., 2017). These regions have been consistently involved in memory encoding and retrieval. All ROIs were defined in MNI space, and then were realigned to each subject's native anatomical space.

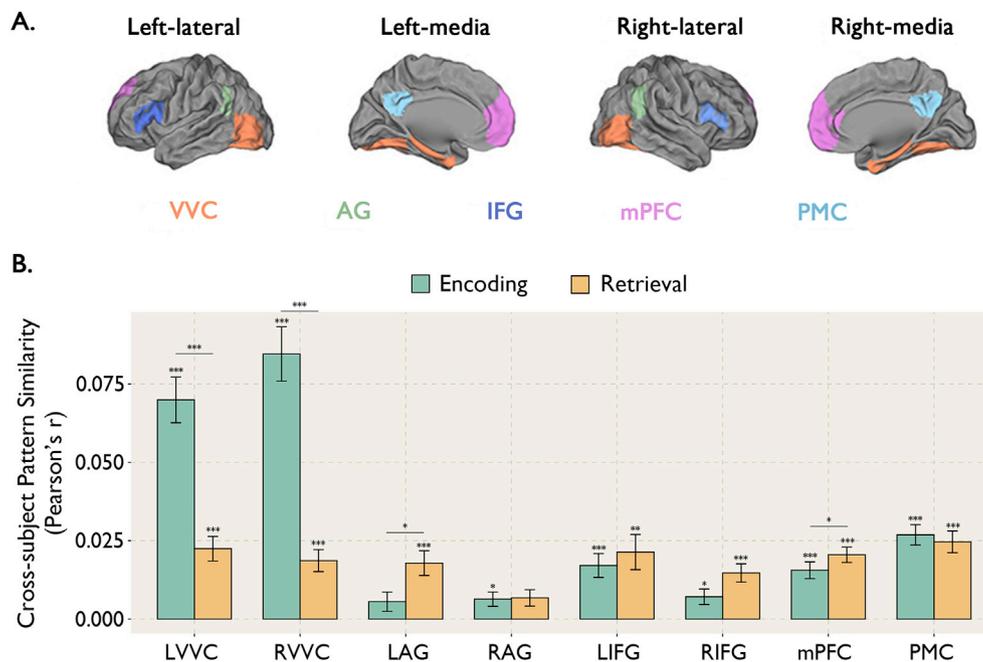


Fig. 2. Shared representational patterns. A. The location of the pre-defined anatomic ROIs. B. Cross-subject representational similarity for encoding and retrieval, using averaged RSMs. Error bars indicate the within-subject error. *** $q < 0.001$, ** $q < 0.01$, * $q < 0.05$.

2.5. Representational space construction

To compare the within-vs. cross-subject memory representations, it is not sufficient to directly contrast a stimulus's within-subject similarity with its cross-subject similarity due to finer differences in anatomical structure (Chen et al., 2017), as well as the differences in structure-function mapping between subjects (i.e., the same function might be localized in different voxels within a region for different individuals). This issue can be ameliorated by comparing the similarity in the representational space, which relies on the second-order similarity. The representational space, namely representational similarity matrix (RSM), is the pairwise pattern correlations for all stimuli in a given brain region. Use the RSM, we could calculate the representational similarities within and between memory stages (e.g., encoding-encoding similarity, retrieval-retrieval similarity, and encoding-retrieval similarity), between brain regions (e.g., cross-region similarity), and across individuals (e.g., cross-subject similarity).

Since each of the 48 pictures was associated with two cues during encoding and retrieval, this design enabled us to construct, for each subject, two RSMs for both encoding and retrieval. This enabled us to compare the within-subject similarity and cross-subject similarity during encoding and retrieval (Fig. 1). Due to the autocorrelation of BOLD signal, the similarity of activation pattern between two stimuli might be affected by their temporal distance within a scan session. As a result, the RSM for a given set would be affected by the temporal sequences of the stimuli, which varied between the two sets within a subject as well as across subjects. To overcome this issue, we did a resampling procedure by randomly switching half of the stimuli between set 1 and set 2. Specifically, each time we randomly selected half of the stimuli from set 1 and the other half from set 2 to form one RSM, and used the remaining stimuli to form another RSM. Each "exchanged" RSM included a full set of the pictures, which were arranged in the same order. The within-subject similarity was calculated by correlating the two exchanged RSMs from the same subject. The cross-subject similarity was calculated by correlating a given subject's RSM with the RSMs from each of the remaining 19 subjects. These 19 similarities were averaged to represent one subject's cross-subject similarity (Fig. 1). The procedure was repeated 1000 times and the results were averaged to obtain more

reliable results. We then compared the within-vs. cross-subject pattern similarity.

Meanwhile, as each exchanged RSM was noisy, we also averaged 1000 RSMs to form an averaged RSM, separately for each individual at each memory stage (encoding and retrieval). As indicated by Fig S2, within-subject similarity increased after averaging the original noisy RSMs, and reached the maximum similarity of 1 after averaging around 400 noisy RSMs. As a result, we could only use the unaveraged, noisy RSMs to compare within- and cross-subject similarity during encoding and retrieval. When examining the shared representation during encoding and retrieval, and comparing the cross-stage and cross-region similarity between within-vs. cross-subject, the averaged RSMs were used for three reasons. First, compared to the individual RSMs, the averaged RSMs were more reliable and less noisy and were less affected by stimulus sequence and BOLD autocorrelation. Second, the averaged RSMs had a higher statistical power to detect the possible shared representation across subjects during encoding and retrieval, if there was any. Third, it was more feasible to do the permutation test on the averaged RSMs than on the individual RSMs from 1000 permutations on each of the 1000 shuffles when the noisy RSMs were used.

2.6. Representational similarities analysis

Three types of similarity (Pearson correlation) in RSM were calculated within each subject: between the two sets of stimuli during encoding, between the same two sets of stimuli during retrieval, and between encoding and retrieval for both sets of stimuli (Fig. 1B). Corresponding to the three types of within-subject similarity in RSM, we calculated three types of cross-subject similarity in RSM, i.e., cross-subject similarity during encoding, during retrieval, and between encoding and retrieval (i.e., ERS). The cross-subject representational similarity for a given subject was obtained by calculating the similarity between the subject's RSM and each of the 19 other subjects' RSM (Fig. 1B), which was then averaged. As stated above, to compare within- and cross-subject representational similarity, the correlation analysis was done on each of the 1000 unaveraged noisy RSMs, and the results were obtained by averaging the 1000 correlations. To examine the degree of shared representations, we used the averaged RSM from the 1000

shuffles. All r values were then transformed into Fisher's Z scores for further analysis.

To examine the cross-region and cross-stage pattern reinstatement (Xiao et al., 2017), we also calculated the similarity between RSM in VVC during encoding and RSM in the mPFC and AG during retrieval. This was done both within and across subjects, using the averaged RSM.

2.7. Permutation test of significance

We used a permutation test to determine statistical significance. In particular, we examined (1) whether there were significant shared representations during encoding and retrieval, and between encoding and retrieval, and (2) whether there was significant within-subject pattern reinstatement during retrieval. To test the significance of shared representations (cross-subject similarity) during encoding or retrieval, we randomly shuffled the labels of the pictures and generated an RSM based on the shuffled picture labels, which was then correlated with the original RSM from each of the remaining 19 subjects. The results were averaged to generate one cross-subject similarity for each shuffle. This was done 1000 times, and the mean was used as the baseline for this subject. When this was done for all subjects, a paired sample t -test was conducted between the similarities of the original RSM and the baseline. To test the within-subject pattern reinstatement, we examined the significance of within-subject ERS. The baseline was obtained by the correlation between the shuffled RSM during encoding and the original RSM during retrieval from the same subject.

We also used the permutation test to examine (3) whether there was subject-specific representation, by comparing within-vs. cross-subject representations. In this analysis, we randomly permuted the labels of within/cross-subject similarity for 1000 times. For each permutation, a difference value between within- and cross-subject similarity was calculated. We then constructed a null distribution of the differences based on the 1000 permutations, and used this distribution to determine the p -value of within-vs. cross-subject comparison. The statistical significance was also corrected for multiple comparisons across all comparisons using FDR (threshold $q = .05$).

2.8. Simulation study

To further investigate the nature of memory transformation, we used simulation to examine the relative contributions of the following factors to memory transformation: (1) stage-specific strength of representation, because different brain regions are involved to a different extent during encoding and retrieval (e.g. greater representational strength in the VVC during encoding and greater representational strength in the AG during retrieval) (Xiao et al., 2017); (2) shared memory transformation (by adding a common random pattern across subjects to their encoding patterns) (Chen et al., 2017); and (3) cross-region reinstatement (by adding the encoded pattern from another brain region). Please note that this encoded pattern from another brain region was both idiosyncratic to each subject and shared across subjects according to our experimental results.

Following Chen et al., we simulated neural activation patterns (125 features each) for five subjects, each containing two brain regions, using the following equation:

$$P_{\text{encoding}} = (\text{subject-shared pattern} + \text{subject-specific pattern} * k_{\text{specific}} + \text{region-specific pattern} * k_{\text{region}}) * k_{\text{strength_encoding}} + \text{random noise}$$

Subject-shared patterns were constructed by creating five identical 125-element/voxel random vectors, whereas the subject-specific patterns were five unique 125-voxel random patterns. The strength of the subject-specific pattern was controlled by k_{specific} . The region-specific pattern was a 125-voxel random pattern shared by subjects, whose strength was controlled by k_{region} . The $k_{\text{strength_encoding}}$ indicates stage-specific representational strength during encoding. The random noise is also a

125-voxel random pattern unique to each subject. The strength of the random noise was controlled by the "noise strength factor" (k_{noise}).

During retrieval, we assumed two mechanisms, including a common transformation pattern shared by subjects, and cross-region transformation, so the pattern during retrieval was:

$$P_{\text{retrieval}} = (P_{\text{encoding}} + \text{common_pattern} * k_{\text{trans_shared}} + P_{\text{encoding}}(\text{other_region}) * k_{\text{CR}}) * k_{\text{strength_retrieval}} + \text{random noise}$$

Where $k_{\text{trans_shared}}$ indicates common transformation patterns shared by subjects and k_{CR} indicates cross-region transformation. $k_{\text{strength_retrieval}}$ here indicates stage-specific representational strength during retrieval.

We systematically varied the following parameters: (1) stage-specific strength of representation during encoding and retrieval ($k_{\text{strength_encoding/retrieval}}$, ranged from 0 to 1, 4 steps, with smaller values indicating weaker representation, e.g., the AG during encoding and VVC during retrieval); (2) the extent of common transformation shared by subjects ($k_{\text{trans_shared}}$, ranged from 0 to 1, 4 steps, with bigger values indicating stronger shared transformation and 0 representing no shared transformation); (3) the extent of cross-region reinstatement (k_{CR} , ranging from 0 to 1, 4 steps, with bigger values representing stronger cross-region reinstatement and 0 representing no cross-region reinstatement); (4) the strength of the random noise (k_{noise} , ranging from 1 to 4, 3 steps, with bigger values indicating stronger random noise).

Please note that for some parameters, such as subject-specific and region-specific representations and noise strength factor, we initially tried a wide range of values, and narrowed down to one value (i.e., 0.6 for subject-specific and region-specific representation, and 4 for noise strength factor) for each parameter to achieve the best match with the experimental data. To reduce the complexity of the results, we only reported results using that value to demonstrate the effect of these factors. As shown in supplementary results, the selection of different values only affected the absolute similarity values, not the overall pattern of results across conditions. For the four critical parameters, stage-specific strength of representation during encoding and retrieval ($k_{\text{strength_encoding/retrieval}}$), strength of common ($k_{\text{trans_shared}}$) and cross-stage (k_{CR}) transformation, we chose relatively finer steps. The simulation was done 1000 times. For each simulation, we calculated within- and cross-subject similarity during encoding and retrieval, within- and cross-subject encoding-retrieval similarity (ERS), and cross-region ERS.

Code and data availability. The code and data supporting the findings of this study are available from the corresponding author upon request, with a formal data-sharing agreement.

3. Results

3.1. Item-level vs. representational space level similarity

Subjects performed very well during the memory test in the scanner (hits $94.3 \pm 5.1\%$). The post-scan test further showed that subjects could correctly report more than four details associated with each retrieved picture. Together, the behavioral results suggest that overtraining before the scan was effective.

In our previous study (Xiao et al., 2017), item-specific memory representation was calculated by examining the pattern similarity between each pair of stimuli sharing the same visual pictures but different cues (i.e., first-order similarity). In the current study, to facilitate within-vs. cross-subject comparison, we examined the item-specific representation by examining the similarity between two representational similarity matrices (RSMs) consisting of identical set of pictures (i.e., second-order similarity). Due to the autocorrelation of BOLD signal, the similarity of activation patterns between two stimuli was affected by their temporal distance within a scan session. We thus shuffled the data by randomly switching half of the stimuli between set 1 and set 2. Within- and cross-subject similarities were then calculated using these exchanged,

and noisy RSMS. As shown in Fig S1, after around 800 exchanges, the averaged within-subject similarity became stable. To be conservative, we used the results based on the average of 1000 exchanges for subsequent statistical analyses. To validate this second-order analysis, we directly compared the results of first-order similarity and second-order similarity. We found that the overall patterns were quite similar, suggesting that both measures could reliably capture the item-specific neural representations (Fig S3).

In our first analysis, we tested whether within-subject similarity (the second-order similarity) was significant above baseline during encoding and retrieval, using the noisy, unaveraged RSMS. We found significant within-subject similarity in the LVVC ($t(19) = 4.823$, $q < 0.001$, with FDR correction), RVVC ($t(19) = 5.032$, $q < 0.001$), LIFG ($t(19) = 3.010$, $q = 0.012$), mPFC ($t(19) = 3.074$, $q = 0.008$) and PMC ($t(19) = 2.871$, $q = 0.020$) during encoding, and in the LAG ($t(19) = 3.202$, $q = 0.008$), LIFG ($t(19) = 4.274$, $q < 0.001$), mPFC ($t(19) = 4.409$, $q < 0.001$) and PMC ($t(19) = 3.294$, $q = 0.005$) during retrieval (Fig S3).

3.2. Shared representations during encoding and retrieval

We then examined whether there were shared representations during memory encoding and retrieval, using the stable, averaged RSMS (see Methods, Fig S2). During encoding, cross-subject representational similarity was significantly greater than baseline in the LVVC ($t(19) = 9.601$, $q < 0.001$), RVVC ($t(19) = 9.789$, $q < 0.001$), RAG ($t(19) = 2.853$, $q = 0.041$), LIFG ($t(19) = 4.53$, $q < 0.001$), RIFG ($t(19) = 2.954$, $q = 0.022$), mPFC ($t(19) = 5.872$, $q < 0.001$) and PMC ($t(19) = 8.27$, $q < 0.001$). During retrieval, significant cross-subject representational similarity was found in the LVVC ($t(19) = 5.745$, $q < 0.001$), RVVC ($t(19) = 5.328$, $q < 0.001$), LAG ($t(19) = 4.531$, $q < 0.001$), LIFG ($t(19) = 3.813$, $q = 0.005$), RIFG ($t(19) = 5.083$, $q < 0.001$), mPFC ($t(19) = 8.368$, $q < 0.001$) and PMC ($t(19) = 7.079$, $q < 0.001$) (Fig. 2B).

Although we found shared representations during both encoding and retrieval, different ROIs showed different patterns: the RAG and LAG only

showed shared representations during encoding and retrieval, respectively, whereas the VVC, LIFG, RIFG, mPFC, and PMC showed shared representations during both encoding and retrieval. Direct comparisons showed that cross-subject similarity was significantly greater during encoding than that during retrieval in the LVVC ($F(1,19) = 39.298$, $q < 0.001$) and RVVC ($F(1,19) = 57.047$, $q < 0.001$), whereas it was greater during retrieval than that during encoding in LAG ($F(1,19) = 7.148$, $q = 0.024$) and mPFC ($F(1,19) = 7.231$, $q = 0.019$) (Fig. 2B). These findings suggest that there were shared representations during both memory encoding and retrieval, but in different regions.

3.3. Comparing shared and individual-specific representations during encoding and retrieval

Having revealed both subject-specific and shared representations during encoding and retrieval, we then examined whether there was greater subject-specific than shared representations. In this analysis, within- and cross-subject similarities were calculated using the noisy, unaveraged RSMS. This was done on 1000 exchanged RSMS (between set 1 and set 2) and the results were averaged.

During encoding, within-subject similarity was significantly greater than cross-subject similarity in the LVVC ($t(19) = 2.828$, $q = 0.029$), RVVC ($t(19) = 2.663$, $q = 0.030$), LIFG ($t(19) = 2.453$, $q = 0.038$), mPFC ($t(19) = 3.140$, $q = 0.029$), and PMC ($t(19) = 2.803$, $q = 0.029$). During retrieval, significant subject-specific representation was found in the LAG ($t(19) = 2.891$, $q = 0.018$), LIFG ($t(19) = 3.697$, $q = 0.008$), mPFC ($t(19) = 4.388$, $q < 0.001$) and PMC ($t(19) = 3.338$, $q = 0.008$) (Fig. 3). To further examine whether the subject-specific representation was modulated by processing stage, we conducted subject-specificity (within-vs. cross-subject similarity) by stage (encoding vs. retrieval) ANOVA, separately for each ROI. This analysis only revealed a marginally significant interaction in the PMC ($F(1,19) = 3.396$, $p = .081$), indicating stronger subject-specific representation during retrieval than that during encoding in the PMC.

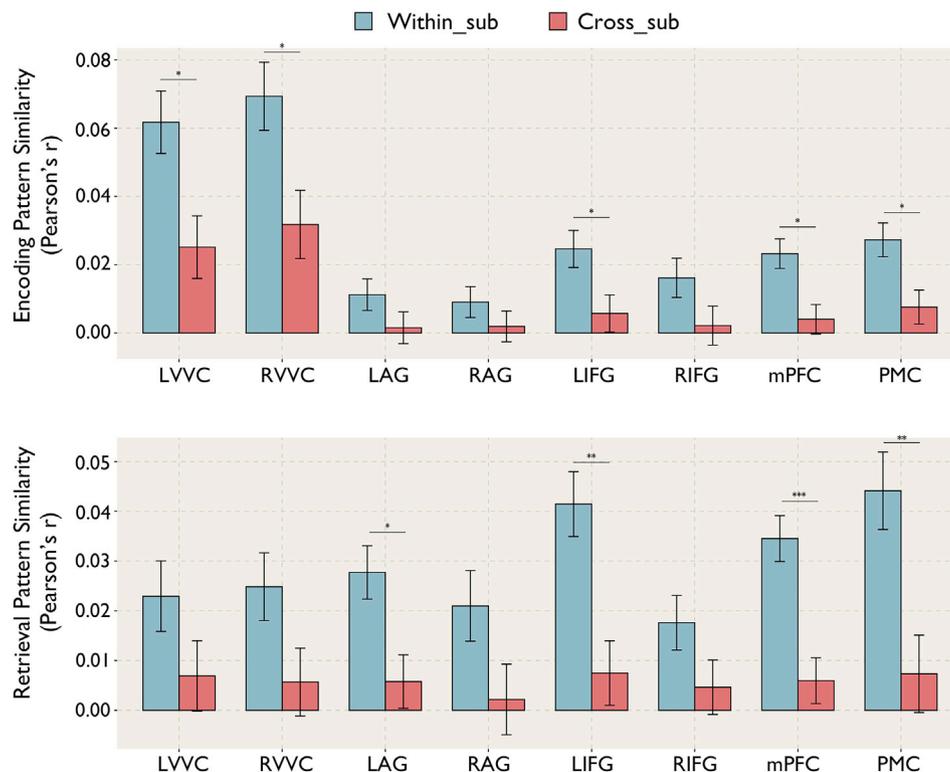


Fig. 3. Comparing the within-subject with cross-subject representational similarities during encoding and retrieval, which were calculated using the 1000 unaveraged, noisy RSMS, and then averaged across 1000 exchanges. Error bars indicate the within-subject error. *** $q < 0.001$, ** $q < 0.01$, * $q < 0.05$.

3.4. Subject-specific and shared pattern reinstatement during memory retrieval

The above analyses revealed both shared and subject-specific representations during encoding and retrieval. Existing studies suggest that the encoded representation might be reinstated during retrieval, showing significant item-specific encoding-retrieval similarity (ERS) (Xiao et al., 2017). This pattern reinstatement has also been found across subjects, as revealed by the significant cross-subject ERS (Chen et al., 2017). Nevertheless, it is still unknown whether there is a subject-specific pattern reinstatement.

To investigate this question, we calculated both within- and cross-subject ERS, using averaged RSMs. We found that all regions showed significant within- and cross-subject ERS compared with baseline (all $q_s < 0.024$, corrected for multiple comparisons), except for the RVVC, LAG, RAG and RIFG, which did not show significant within-subject ERS (RVVC: $t(19) = 2.210$, $q = 0.079$; LAG: $t(19) = 1.759$, $p = .095$; RAG: $t(19) = 2.036$, $p = .056$, uncorrected; RIFG: $t(19) = 1.136$, $p = .270$). Interestingly, although within-subject ERS was numerically greater than cross-subject ERS, the direct comparison revealed no significant differences (all $p_s > .08$, without correction for multiple comparisons) (Fig. 4). These results thus suggest strong shared pattern reinstatement across subjects, but no evidence for subject-specific reinstatement, which is consistent with a previous study which revealed that the within-subject ERS was only slightly greater than cross-subject ERS (with two voxels in temporoparietal junction surviving correction) (Chen et al., 2017).

3.5. Shared within-region memory transformation from encoding to retrieval

The lack of subject-specific reinstatement might be due to overall weak ERS, as a result of memory transformation (Xiao et al., 2017), as well as shared memory transformation across-subject (Chen et al., 2017). For example, there was greater pattern similarity during encoding (i.e., encoding-encoding similarity, EES) and retrieval (i.e., retrieval-retrieval similarity, RRS) than the similarity between encoding and retrieval (ERS) (Xiao et al., 2017). Furthermore, there was greater a cross-subject RRS than cross-subject ERS (Chen et al., 2017).

To examine the shared representational transformation from encoding to retrieval in our study, we compared cross-subject ERS with cross-subject similarity during encoding and retrieval. We found that cross-subject ERS was smaller than cross-subject EES in the VVC (LVVC: $F(1,19) = 42.097$, $q < 0.001$; RVVC: $F(1,19) = 46.876$, $q < 0.001$). Meanwhile, cross-subject ERS was also marginally smaller than cross-subject RRS in the LAG ($F(1,19) = 3.07$, $q = 0.096$) (Fig. 4). This

evidence indicates that the representational patterns might be transformed in a systematic way across subjects from encoding to retrieval in VVC and LAG. Together, although using very different designs and materials, our results showed a similar pattern with previous studies (Chen et al., 2017; Zadbood et al., 2017).

3.6. Subject-specific and shared cross-region pattern reinstatement

Representational reinstatement not only occurs within a brain region, but also across regions, such that the encoded representation in one region is reinstated in a different region during retrieval (Xiao et al., 2017), due to their differential roles in memory encoding and retrieval (See Introduction). It is still unclear whether cross-region reinstatement can be found across subjects.

When examining the shared representation, the current study found interesting dissociations in brain regions showing shared representations during encoding and retrieval: the bilateral VVC showed stronger shared representations during encoding, while the mPFC and LAG showed stronger shared representations during retrieval. These results led to the following questions: Was the encoded representation in VVC reinstated in mPFC and LAG during retrieval? If so, was this cross-region reinstatement unique to each individual or shared across subjects?

By examining the cross-subject, cross-region, and cross-stage similarity (e.g., sub1, VVC-encoding with sub2, mPFC-retrieval), we found significant cross-subject correlations between encoded RSM in VVC and retrieved RSM in mPFC (LVVC-mPFC: $t(19) = 3.882$, $q = 0.001$; RVVC-mPFC: $t(19) = 3.941$, $q = 0.001$), as well in LAG (LVVC-LAG: $t(19) = 2.963$, $q = 0.016$; RVVC-LAG: $t(19) = 2.867$, $q = 0.040$) (Fig. 5), which indicates that the information in bilateral VVC during encoding might be reinstated in LAG and mPFC during retrieval. Interestingly, we found that the within-region ERS in LVVC and RVVC was comparable to the VVC-LAG ERS, both within (LVVC: $t(19) = -0.256$, $p = .800$; RVVC: $t(19) = 0.489$, $p = .631$) and across subjects (LVVC: $t(19) = -1.396$, $p = .179$; RVVC: $t(19) = -1.689$, $p = .107$), which suggests comparable strength of within-region and cross-region reinstatement. The within-subject RVVC-LAG reinstatement was greater than that for across subjects ($t(19) = 2.066$, $q = 0.05$), suggesting individual-specific cross-region reinstatement (Fig. 5).

3.7. Simulation analysis to examine the mechanisms of memory transformation

The above analysis revealed several important features regarding shared and subject-specific representations, as well as memory transformations across memory stages and brain regions. First, we found that

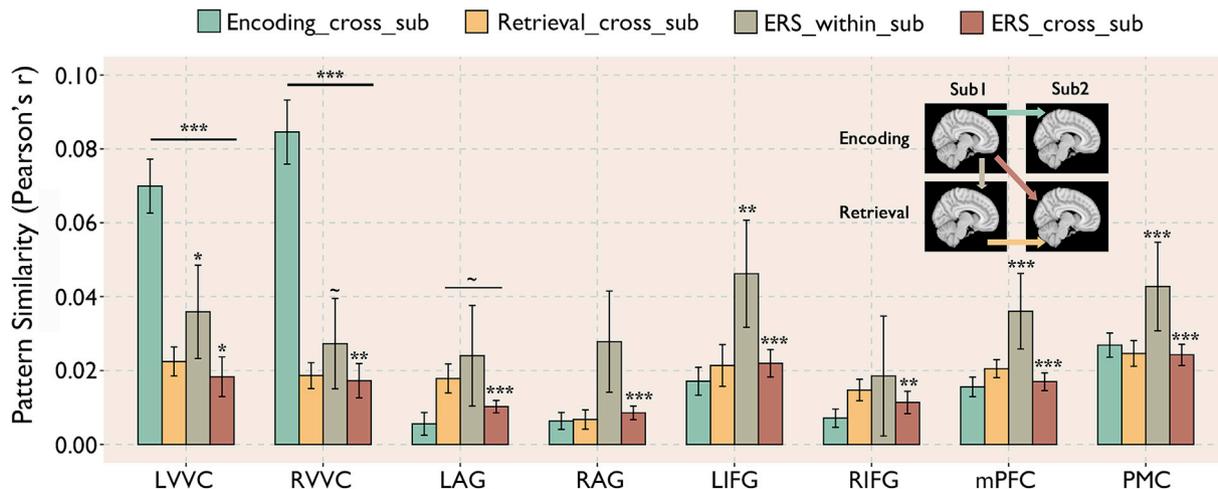


Fig. 4. Comparison of cross-subject ERS, encoding similarity, retrieval similarity, and within-subject ERS, using the averaged RSMs. Error bars indicate standard error. *** $q < 0.001$, ** $q < 0.01$, * $q < 0.05$.

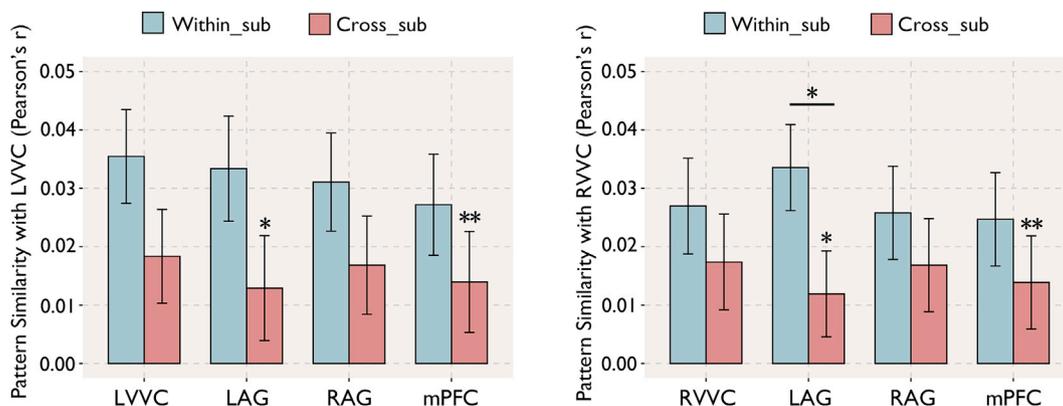


Fig. 5. Comparing within-vs. cross-subject reinstatement, from LVVC (left panel) and RVVC (right panel) to ipsilateral VVC, LAG, RAG and mPFC. Error bars indicate the within-subject error. *** $q < 0.001$, ** $q < 0.01$, * $q \leq 0.05$.

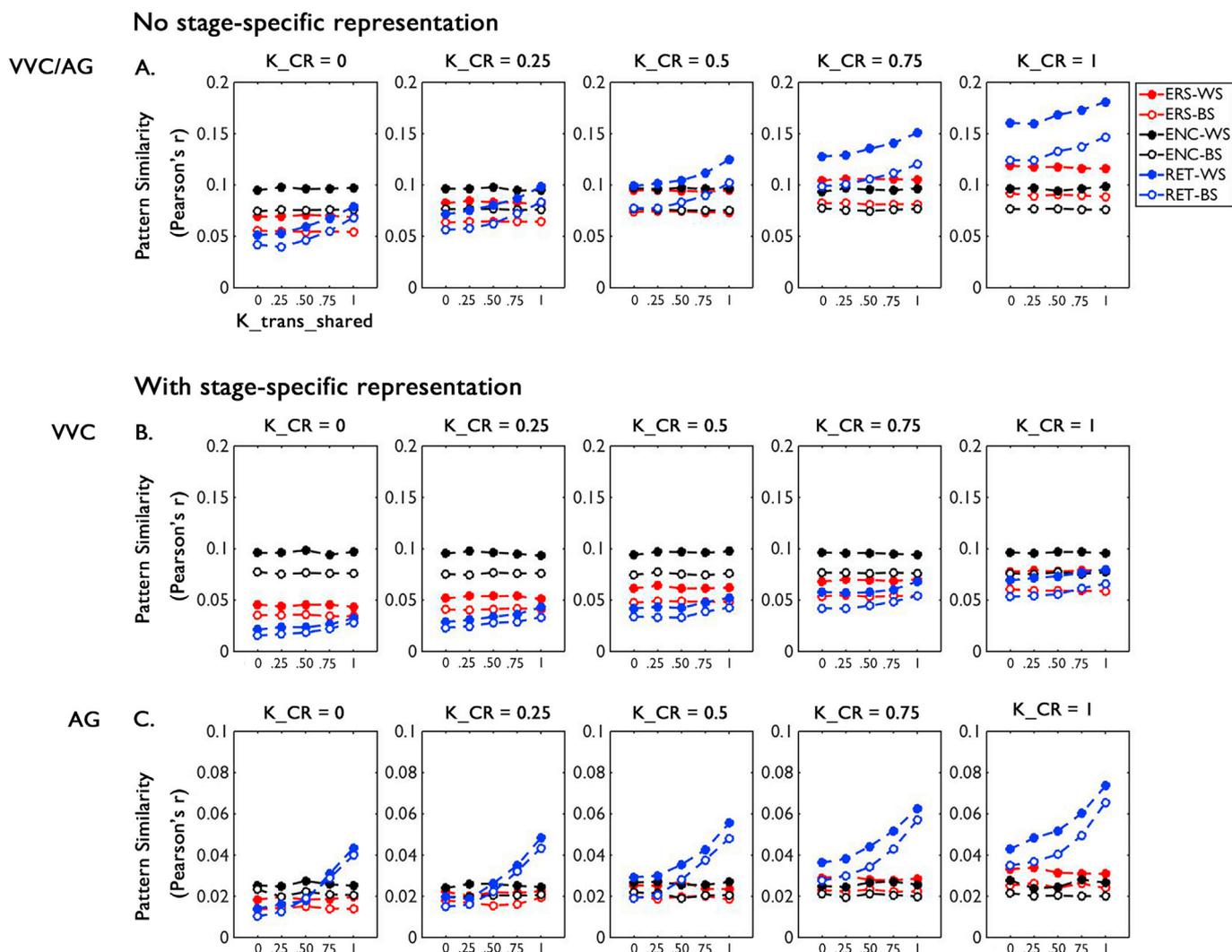


Fig. 6. Simulation results. The top row (A) shows the result when no stage-specific representation strength was introduced for VVC and AG, in this case, the representation patterns of VVC and AG were the same. The bottom two rows (B & C) show the results when stage-specific representation strength was included (for AG: $k_{strength_encoding} = 0.5$, $k_{strength_retrieval} = 1$; for VVC: $k_{strength_encoding} = 1$, $k_{strength_retrieval} = 0.5$). Each plot shows pattern similarity as a function of shared within-region transformation across subjects (k_{trans_shared} , varying from 0 to 1, with 4 steps). The five panels from left to right show the results with different levels of cross-region reinstatement (k_{CR} , varying from 0 to 1, with 4 steps). It is clear that the stage-specific representation is necessary to generate greater pattern similarity during retrieval than during encoding (e.g., VVC, Fig. 6A vs. Fig. 6B), and stage-specific representation strength together with cross-region reinstatement could generate greater pattern similarity during retrieval than during encoding (e.g., AG, Fig. 6A vs. Fig. 6C). ERS: encoding-retrieval similarity; ENC: encoding; RET: retrieval; WS: within-subject; BS: between-subject.

the same brain regions contained both shared and subject-specific representations, and that there was greater within-subject similarity than cross-subject similarity during both encoding and retrieval, in both the visual cortex and higher-order brain regions, such as the LAG. Second, the representational strength was stage-specific across regions, with the VVC and LAG showing stronger representational strength during encoding and retrieval, respectively. Third, cross-subject ERS was smaller than cross-subject similarity during encoding and retrieval, suggesting that the representation was transformed. Fourth, within-subject ERS was numerically bigger but statistically comparable to cross-subject ERS, suggesting some common transformation shared by subjects. Finally, there existed cross-stage and cross-region reinstatement both within and across subjects, with stronger within-subject reinstatement in the LAG.

A critical question then is how these memory transformations are achieved. One possibility, as suggested by Chen et al. (2017), is that a common transformation could be introduced from encoding to retrieval, by adding a common memory pattern to all subjects at retrieval. Alternatively, enlightened by the findings of stage-specific representational strength, as well as cross-region transformation in the current study, the comparable cross- and within-subject ERS might be due to the overall small ERS as a result of memory transformation, and the common transformation might result from the cross-region transformation (e.g., from VVC to AG), given the shared representation in VVC during encoding.

In this section, we used simulation to adjudicate these hypotheses. In particular, we focused on three factors: (1) the stage-specific strength of memory representation in different brain regions ($k_{\text{strength_encoding/retrieval}}$), (2) the extent of shared cross-stage transformation ($k_{\text{trans_shared}}$), and (3) the extent of cross-region reinstatement (k_{CR}) (see Methods).

Our simulation showed that these three factors affected the pattern similarities in different ways. Specifically, the shared transformation alone increased retrieval similarities (blue lines) (Fig. 6A, left panel), consistent with Chen et al. (2017). A weaker stage-specific representation strength decreased the pattern similarities during encoding and retrieval, as well as the ERS (Fig. 6B and C). Finally, the cross-region reinstatement alone increased the retrieval similarities enough to surpass encoding similarities and ERS (Fig. 6A).

We then asked which mechanism(s) could better account for the empirical data. Our simulation showed that although the shared transformation mechanism increased cross-subject retrieval similarity (blue lines) (Fig. 6), itself alone did not lead to greater cross-subject retrieval similarity than cross-subject encoding similarity in AG (Fig. 6A, left panel). Also, the strength of shared transformation did not affect the cross-region pattern similarity (Fig. 7, left panel). In contrast, with stage-

specific representation (especially the weak representation during encoding for AG), we simulated the greater pattern similarity during retrieval than encoding (Fig. 6C, right panel). Combined with increasing strength of cross-region transformation, the cross-subject retrieval similarity surpassed cross-subject ERS (Fig. 6), and the between-region ERS (black lines) was comparable to the within-region ERS (red lines) (Fig. 7).

Taken together, the simulation results suggest that stage-specific representational strength and cross-region reinstatement can fit the observed data very well, whereas the shared transformation mechanisms alone cannot.

4. Discussion

Combining a unique experimental design with cross-subject representational similarity analysis, the current study for the first time systematically examined within-vs. cross-subject representations during memory encoding and retrieval, as well as the transformation of representations between the two processing stages. Our results found both individual-specific and shared memory representations in distributed brain regions, with the visual cortex showing greater item-specific representation (both subject-specific and shared) during encoding and the angular gyrus showing greater item-specific representation (both subject-specific and shared) during retrieval. Importantly, we found that memory representation was transformed systematically across regions and stages. Our simulation captured well the response pattern by incorporating stage-specific representational strength and cross-region pattern reinstatement. Together, these results help to advance our understanding of the nature of common and unique memory representations among different individuals and their dynamic changes.

Given the personalized nature of individuals' memories, it is thus of great surprise and significance to find common neural representations during memory retrieval (Charest et al., 2014; Chen et al., 2017). However, the degree of commonality and uniqueness of memory representations across individuals remains unclear. One previous study found that individuals' semantic representation patterns could predict their false memory responses better than could the group-averaged responses (Chadwick et al., 2016). Our study represents a major extension of existing studies by providing a comprehensive picture of the unique and shared representations during memory encoding and retrieval. In particular, we found that almost all regions showed significant shared representations across subjects, but there were also much greater within-than cross-subject similarity, during both encoding and retrieval. This provides clear neural evidence to support the behavioral observations that each brain has distinct representational patterns in perceiving, remembering, and recalling an event, although there are also

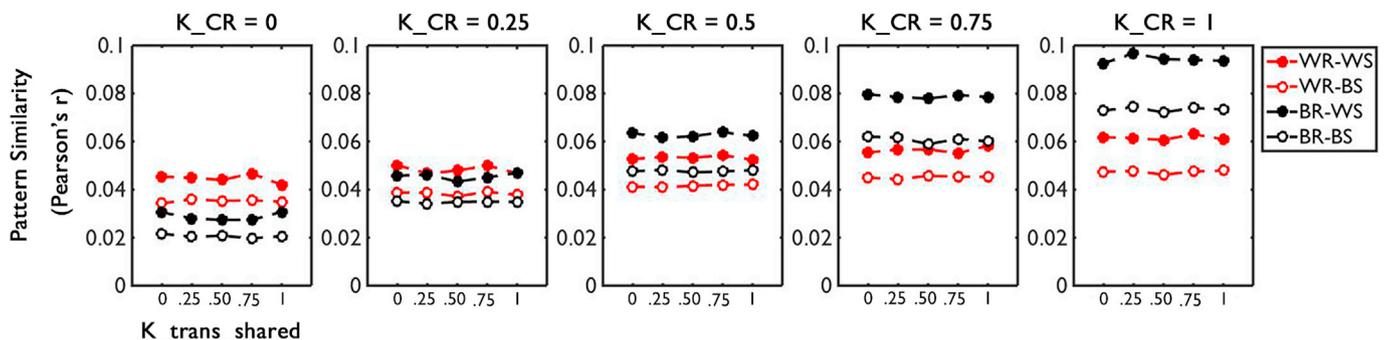


Fig. 7. Simulation of cross-region pattern similarity for VVC to AG. Each plot shows pattern similarity as a function of shared within-region transformation across subjects ($k_{\text{trans_shared}}$, varying from 0 to 1, with 4 steps). The five panels from left to right show the results with different levels of cross-region reinstatement (k_{CR} , varying from 0 to 1, with 4 steps). With increasing strength of cross-region transformation, cross-region similarity (black lines) increased to match and even surpass within-region similarity (red lines). The strength of shared within-region transformation had no significant effect on the cross-region pattern similarity. WR: within-region; BR: between-region; WS: within-subject; BS: between-subject.

cross-subject similarities.

Interestingly, we found that the degree of shared representations was modulated by processing stages: Whereas the bilateral VVC showed greater shared representations during encoding, the AG and mPFC showed greater shared representations during retrieval. This is consistent with a previous study of shared representations during movie watching and recall (Chen et al., 2017). Using different analytical approaches, the current study and a previous study (Xiao et al., 2017) found the same pattern of results that emphasizes the role of VVC and AG in representing high-fidelity item-level information during encoding and retrieval, respectively. Consistently, another study only found category-level but not item-level reinstatement in the VVC during retrieval (Kuhl and Chun, 2014), and it was not possible to reconstruct the perceived images from the occipitotemporal representations during working memory (Lee and Kuhl, 2016). Given these regions' distinct roles in memory encoding and retrieval, it might be reasonable to argue that the shared memory representations during recall in the high-level cortical regions are mainly due to their representational role in memory retrieval. Our simulation analysis further emphasizes the importance of stage-specific memory representations.

These results suggest that retrieval might not involve a faithful reinstatement of encoded representations, but rather may engage additional abstraction processes to enable the formation of conceptual knowledge from perceptual experience (Binder and Desai, 2011). Indeed, the AG and VVC have been posited to represent distinct information: the representation in the VVC contains perceptual details, whereas the representation in the AG is identity-specific and invariant to viewpoints (Jeong and Xu, 2016), and is modulated by semantic similarity (Ye et al., 2016). Besides, the cross-subject encoding-retrieval similarity was smaller than that during encoding and retrieval, suggesting the representation in the same regions might be altered from encoding to retrieval. Combining both lines of evidence, we further hypothesized and verified that this abstraction might be achieved by cross-region reinstatement, i.e., the encoded representation in VVC was reinstated in the mPFC and AG. In particular, we found the VVC-LAG ERS was comparable to LAG-LAG ERS, suggesting some cross-region information transformation.

It should be noted that encoding and retrieval involve different task structures and visual inputs, which might have contributed to the different activation patterns in encoding and retrieval. For example, we found a significant shift in brain regions that showed shared and subject-specific representations during encoding (e.g., VVC) and retrieval (e.g., AG), which could be due to the task differences between the two memory stages. Nevertheless, by focusing on the neural representations, our data also revealed changes in representational format between encoding and retrieval. Furthermore, the second-order pattern analysis, like the first-order analysis of item-specific representation used in the previous study (Xiao et al., 2017), could control the effect of common cognitive processes shared by items on the representation patterns. As a result, the differences between encoding and retrieval within regions mainly reflected the change of item-specific representations. One potential caveat is that this approach could not account for the possible task by stimuli interactions, which could affect the representational structure. Future studies should further examine the nature of representations during encoding and retrieval, such as the representational format, content, and dimension, to further elucidate the transformative nature of memory retrieval.

The current study revealed no significant differences between within- and cross-subject ERS, although the former was numerically greater. This is also quantitatively consistent with a previous observation that only a few voxels showed greater within-than cross-subject ERS (Chen et al., 2017). Inconsistent with the previous study (Chen et al., 2017) which found significantly greater cross-subject retrieval similarity than cross-subject ERS in the PMC, the current study only found a marginally significant difference in the angular gyrus but not PMC. This could be due to the obvious differences between the two studies in their experimental design, materials and analytical approaches. First, the materials used in

the current study were rather homogeneous, as compared to various events in a movie in Chen et al. (2017). This could reduce the overall statistical power in the representational similarity analysis. Second, we did not ask subjects to verbally describe the visual scenes during memory retrieval, which probably reduced the encoding-to-retrieval transformation. Third, the visual representation might be less likely to be shared across individuals than the abstract verbal description, reducing the cross-subject similarity during retrieval. Future studies should examine the shared and unique memory representations using different memory paradigms.

What could lead to the stronger cross-subject similarity during retrieval than ERS, and also the comparable within- and cross-subject ERS? As suggested by Chen and colleagues (Chen et al., 2017), this observation could suggest a common memory transformation. Our simulation results suggest that adding a strong new pattern shared by all subjects during retrieval could simulate the greater retrieval similarity than cross-subject ERS (Chen et al., 2017), but this common transformation alone could not generate the stage-specific representation strength, the comparable within-vs. cross-subject ERS, or the comparable cross-vs. within-region ERS in the LAG. Still, it is unclear how this common transform could be achieved. Unlike other studies (Hasson, Ghazanfar, Galantucci, Garrod and Keysers, 2012; Zadbood et al., 2017), the experimental paradigm used in the current study and Chen et al. (2017) did not engage social interaction or social exchange.

In contrast, by incorporating stage-specific representational strength and cross-region reinstatement, as revealed by our data, we were able to simulate the greater cross-subject retrieval similarity than cross-subject ERS. Since the encoded representation in VVC contained both shared and subject-specific features, this cross-region transformation thus is both subject-specific and shared, which could well account for the subject-specific cross-region transformation. It should be noted that in Chen et al.'s study, a subject-specific random pattern was added to the common transformation pattern, which was termed noise in memory transformation. Without this noise, it is impossible to obtain greater within-than cross-subject similarity during retrieval. Together, our results suggest that the cross-region transformation mechanism could well account for the empirical data, as well as provide a good mechanism for the origin of subject-specific and shared patterns of cross-stage memory transformation.

How the cross-region pattern reinstatement is achieved in the brain is still unclear. One possibility is that different brain regions contain different aspects of representations of the stimuli, which share the same representational structures, allowing the consistent mapping from one brain region to another brain region during information transformation. It has been shown that neural information flow is reversed between object perception and object reconstruction from memory (Linde-Domingo et al., 2019). In particular, when seeing an object, low-level perceptual features were discriminated faster and could be decoded from brain activity earlier, than high-level conceptual features. During retrieval, the encoded representation was actively reconstructed and represented in the angular gyrus. This is achieved through the hippocampal pattern reinstatement mechanisms (Estefan et al., 2019; Lohnas et al., 2018; Xiao et al., 2017), by integrating the information from the visual cortex as well as long-term knowledge. Future studies should examine the retrieval processes with higher spatiotemporal measures of brain activation patterns.

The capacity to share memories is essential for our ability to interact with others and form social groups, and the current study provides a practical methodological framework to study cross-subject representation. Social interaction has been primarily studied using neural synchronization across subjects in the temporal domain (Nummenmaa et al., 2018), but cross-subject pattern similarity analysis could provide a more direct measure of the shared mental representation. So far, this type of analysis has been mainly conducted on fMRI data in terms of spatial patterns, but it could be readily extended to the spatiotemporal domain (Feng et al., 2019; Lu et al., 2015). Several methodological caveats

should be mentioned here. First, as emphasized in our analysis, the autocorrelation of the BOLD signal would affect the activation pattern, which could create artificial patterns of similarities when the trial orders are matched between two representation similarity matrices (RSM), or decrease the similarity results if the orders are not matched. The random exchange could help to solve this issue, but increase the computational demand and make the whole-brain searchlight analysis less feasible. Second, the similarity between two RSMs might not suggest the same representational features, but rather some higher-order similarity. Third, the regions were defined based on the group-based anatomical template. Recent studies have shown with dense sampling of resting-state data that we can achieve precise parcellation of individual brains that are very well aligned with the functional architecture (Braga and Buckner, 2017; Choe et al., 2015; Gordon et al., 2017; Laumann et al., 2015; O'Connor et al., 2017). It remains to be seen whether we can reveal stronger shared representations using individualized brain parcellation.

Finally, the current study overtrained the subjects to achieve high memory performance during retrieval. This is critical for us to examine the representation during retrieval both within and across subjects. This overtraining, however, might have changed the encoding processes, stabilize the representation (Huang et al., 2013; Wiestler and Die-drichsen, 2013), make the encoding and retrieval processes more similar (i.e., the restudy of the same item might involve the retrieval of a prior representation, termed study-phase retrieval). Although these factors might reduce the differences between encoding and retrieval, the fact that we still found low encoding-retrieval similarity provides strong evidence that memory is transformed. We would expect even greater transformation for one-shot learning, an interesting hypothesis that could be examined in future studies.

Taken together, this study tested the nature of the individual-specific and shared memory representations simultaneously, and also tested their transformation patterns from encoding to retrieval. Our results underscore the stage-specific representation and reconstructive nature of memory. Future research should further investigate how social interaction shapes the subject-specific and shared memory representation.

Declaration of competing interest

The authors declare no conflict of interest.

CRediT authorship contribution statement

Xiaoqian Xiao: Data curation, Formal analysis, Visualization, Writing - original draft. **Yu Zhou:** Formal analysis, Visualization, Writing - review & editing. **Jing Liu:** Formal analysis. **Zhifang Ye:** Formal analysis. **Li Yao:** Supervision, Writing - review & editing. **Jiacai Zhang:** Supervision, Writing - review & editing. **Chuansheng Chen:** Writing - review & editing. **Gui Xue:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing.

Acknowledgment

This work was sponsored by the National Natural Science Foundation of China (31730038), the NSFC and the Israel Science Foundation (ISF) joint project (31861143040), the NSFC and the German Research Foundation (DFG) joint project NSFC 61621136008/DFG TRR-169, and the Guangdong Pearl River Talents Plan Innovative and Entrepreneurial Team grant #2016ZT06S220

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.neuroimage.2020.116909>.

References

- Avants, B.B., Tustison, N.J., Song, G., Cook, P.A., Klein, A., Gee, J.C., 2011. A reproducible evaluation of ANTs similarity metric performance in brain image registration. *Neuroimage* 54 (3), 2033–2044. <https://doi.org/10.1016/j.neuroimage.2010.09.025>.
- Binder, J.R., Desai, R.H., 2011. The neurobiology of semantic memory. *Trends Cognit. Sci.* 15 (11), 527–536. <https://doi.org/10.1016/j.tics.2011.10.001>.
- Braga, R.M., Buckner, R.L., 2017. Parallel interdigitated distributed networks within the individual estimated by intrinsic functional connectivity. *Neuron* 95 (2), 457–471. <https://doi.org/10.1016/j.neuron.2017.06.038> e5.
- Chadwick, M., Anjum, R., Kumaran, D., Schacter, D., Spiers, H., Hassabis, D., 2016. Semantic representations in the temporal pole predict false memories. In: *PNAS - Proceedings of the National Academy of Sciences of the United States of America*.
- Charest, I., Kievit, R.A., Schmitz, T.W., Deca, D., Kriegeskorte, N., 2014. Unique semantic space in the brain of each beholder predicts perceived similarity. *Proc. Natl. Acad. Sci. U.S.A.* 111 (40), 14565–14570. <https://doi.org/10.1073/pnas.1402594111>.
- 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. <https://doi.org/10.1038/nn.4450>.
- Choe, A.S., Jones, C.K., Joel, S.E., Muschelli, J., Belegu, V., Caffo, B.S., Lindquist, M.A., van Zijl, P.C.M., Pekar, J.J., 2015. Reproducibility and temporal structure in weekly resting-state fMRI over a period of 3.5 years. *Report* 10 (10). <https://doi.org/10.1371/journal.pone.0140134>.
- Danker, J.F., Tomparly, A., Davachi, L., 2017. Trial-by-Trial hippocampal encoding activation predicts the fidelity of cortical reinstatement during subsequent retrieval. *Cerebr. Cortex* 27 (7), 3515–3524. <https://doi.org/10.1093/cercor/bhw146>.
- Estefan, D.P., Sánchez-Fibla, M., Duff, A., Principe, A., Rocamora, R., Zhang, H., Axmacher, N., Verschuer, P.F.M.J., 2019. Coordinated representational reinstatement in the human hippocampus and lateral temporal cortex during episodic memory retrieval. *Nat. Commun.* 10 (1), 1–13. <https://doi.org/10.1038/s41467-019-09569-0>.
- Feng, K., Zhao, X., Liu, J., Cai, Y., Ye, Z., Chen, C., Xue, G., 2019. Spaced learning enhances episodic memory by increasing neural pattern similarity across repetitions. *J. Neurosci.* 39 (27), 5351–5360. <https://doi.org/10.1523/JNEUROSCI.2741-18.2019>.
- Golland, Y., Bentin, S., Gelbard, H., Benjamini, Y., Heller, R., Nir, Y., Hasson, U., Malach, R., 2007. Extrinsic and intrinsic systems in the posterior cortex of the human brain revealed during natural sensory stimulation. *Cerebr. Cortex* 17 (4), 766–777. <https://doi.org/10.1093/cercor/bhk030>.
- Gordon, E.M., Laumann, T.O., Gilmore, A.W., Newbold, D.J., Greene, D.J., Berg, J.J., Ortega, M., Hoyt-Drazen, C., Grattton, C., Sun, H., Hampton, J.M., Coalsen, R.S., Nguyen, A.L., Mcdermott, K.B., Shimony, J.S., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., Nelson, S.M., Dosenbach, N.U.F., 2017. Precision functional mapping of individual human brains. *Neuron* 95 (4), 791–807. <https://doi.org/10.1016/j.neuron.2017.07.011> e7.
- Guise, K.G., Shapiro, M.L., 2017. Medial prefrontal cortex reduces memory interference by modifying hippocampal encoding. *Neuron* 94 (1), 183–192. <https://doi.org/10.1016/j.neuron.2017.03.011> e8.
- Hasson, U., Furman, O., Clark, D., Dudai, Y., Davachi, L., 2008. Enhanced intersubject correlations during movie viewing correlate with successful episodic encoding. *Neuron* 57 (3), 452–462. <https://doi.org/10.1016/j.neuron.2007.12.009>.
- Hasson, U., Malach, R., Heeger, D.J., 2010. Reliability of cortical activity during natural stimulation. *Trends Cognit. Sci.* 14 (1), 40–48. <https://doi.org/10.1016/j.tics.2009.10.011>.
- Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., Malach, R., 2004. Intersubject synchronization of cortical activity during natural vision. *Science* 303 (5664), 1634–1640.
- Huang, Y., Zhen, Z., Song, Y., Zhu, Q., Wang, S., Liu, J., 2013. Motor training increases the stability of activation patterns in the primary motor cortex. (research article). *PLoS One* 8 (1), e53555. <https://doi.org/10.1371/journal.pone.0053555>.
- Huth, A.G., Heer, W. A. De, 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. <https://doi.org/10.1038/nature17637>.
- Jääskeläinen, I.P., Koskentalo, K., Balk, M.H., Autti, T., Kauramäki, J., Pomren, C., Sams, M., 2008. Inter-subject synchronization of prefrontal cortex hemodynamic activity during natural viewing. *Open Neuroimaging J.* 2, 14–19. <https://doi.org/10.2174/1874440000802010014>.
- Jeong, S.K., Xu, Y., 2016. Behaviorally relevant abstract object identity representation in the human parietal cortex. *J. Neurosci.* 36 (5), 1607–1619. <https://doi.org/10.1523/JNEUROSCI.1016-15.2016>.
- Kauppi, J.-P., Jääskeläinen, I.P., Sams, M., Tohka, J., 2010. Inter-subject correlation of brain hemodynamic responses during watching a movie: localization in space and frequency. *Front. Neuroinf.* 4. <https://doi.org/10.3389/fninf.2010.00005>.
- Kirchhoff, B.A., 2009. Individual differences in episodic memory: the role of self-initiated encoding strategies. *Neuroscientist* 15 (2), 166–179. <https://doi.org/10.1177/1073858408329507>.
- Klein, A., Andersson, J., Ardekani, B.A., Ashburner, J., Avants, B., Chiang, M.-C., Christensen, G.E., Collins, D.L., Gee, J., Hellier, P., Song, J.H., Jenkinson, M., Lepage, C., Rueckert, D., Thompson, P., Vercauteren, T., Woods, R.P., Mann, J.J., Parsey, R.V., 2009. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage* 46 (3), 786–802. <https://doi.org/10.1016/j.neuroimage.2008.12.037>.
- Kriegeskorte, N., Mur, M., Bandettini, P., 2008. Representational similarity analysis - connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* 2. <https://doi.org/10.3389/neuro.06.004.2008>.

- Kuhl, B.A., Chun, M.M., 2014. Successful remembering elicits event-specific activity patterns in lateral parietal cortex. *J. Neurosci.* 34 (23), 8051–8060. <https://doi.org/10.1523/JNEUROSCI.4328-13.2014>.
- Laumann, T. o, Gordon, E. m, Adeyemo, B., Snyder, A. z, Joo, S. jun, Chen, M.-Y., Gilmore, A. w, Mcdermott, K. b, Nelson, S. m, Dosenbach, N.u. ., Schlaggar, B.L., 2015. Functional system and areal organization of a highly sampled individual human brain. *Neuron* 87 (3), 657–670. <https://doi.org/10.1016/j.neuron.2015.06.037>.
- Lee, H., Kuhl, B.A., 2016. Reconstructing perceived and retrieved faces from activity patterns in lateral parietal cortex. *J. Neurosci.* 36 (22), 6069–6082. <https://doi.org/10.1523/JNEUROSCI.4286-15.2016>.
- Linde-Domingo, J., Treder, M.S., Kerrén, C., Wimber, M., 2019. Evidence that neural information flow is reversed between object perception and object reconstruction from memory. *Nat. Commun.* 10 (1), 1–13. <https://doi.org/10.1038/s41467-018-08080-2>.
- Loftus, E.F., Levidow, B., Duensing, S., 1992. Who remembers best? Individual differences in memory for events that occurred in a science museum. *Appl. Cognit. Psychol.* 6 (2), 93–107. <https://doi.org/10.1002/acp.2350060202>.
- Lohnas, L.J., Duncan, K., Doyle, W.K., Thesen, T., Devinsky, O., Davachi, L., 2018. Time-resolved neural reinstatement and pattern separation during memory decisions in human hippocampus. *Proc. Natl. Acad. Sci. U.S.A.* 115 (31), E7418–E7427. <https://doi.org/10.1073/pnas.1717088115>.
- Lu, Y., Wang, C., Chen, C., Xue, G., 2015. Spatiotemporal neural pattern similarity supports episodic memory. *Curr. Biol.* 25 (6), 780. <https://doi.org/10.1016/j.cub.2015.01.055>.
- Miller, M.B., Van Horn, J., Wolford, G., Handy, T.C., Valsangkar-Smyth, M., Inati, S., Grafton, S., Gazzaniga, M.S., 2002. Extensive individual differences in brain activations associated with episodic retrieval are reliable over time. *J. Cognit. Neurosci.* 32.
- Mumford, J.A., Turner, B.O., Ashby, F.G., Poldrack, R.A., 2012. Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *Neuroimage* 59 (3), 2636–2643. <https://doi.org/10.1016/j.neuroimage.2011.08.076>.
- Munday, R., 1985. Eyewitness testimony: psychological perspectives. *Camb. Law J.* 44 (1), 160–162. <https://doi.org/10.1017/S0008197300114564>. Edited by Gary L. Wells and Elizabeth F. Loftus. [Cambridge: Cambridge University Press. 1984. x, 335 and (Indexes) 39 pp. £25-00 net.].
- Nummenmaa, L., Lahnakoski, J.M., Glerean, E., 2018. Sharing the social world via intersubject neural synchronisation. *Current Opinion in Psychology* 24, 7–14. <https://doi.org/10.1016/j.copsyc.2018.02.021>.
- O'Connor, D., Potler, N.V., Kovacs, M., Xu, T., Ai, L., Pellman, J., Vanderwal, T., Parra, L.C., Cohen, S., Ghosh, S., Escalera, J., Grant-Villegas, N., Osman, Y., Bui, A., Craddock, R.C., Milham, M.P., 2017. The Healthy Brain Network Serial Scanning Initiative: a resource for evaluating inter-individual differences and their reliabilities across scan conditions and sessions. *GigaScience* 6 (2). <https://doi.org/10.1093/gigascience/giw011>.
- Ritche, M., Wing, E.A., LaBar, K.S., Cabeza, R., 2013. Neural similarity between encoding and retrieval is related to memory via hippocampal interactions. *Cerebr. Cortex* 23 (12), 2818–2828. <https://doi.org/10.1093/cercor/bhs258>.
- Schacter, D.L., Norman, K.A., Koutstaal, W., 1998. The cognitive neuroscience of constructive memory. *Annu. Rev. Psychol.* 49, 289–318. <https://doi.org/10.1146/annurev.psych.49.1.289>.
- Shapiro, P.N., Penrod, S., 1986. Meta-analysis of facial identification studies. *Psychol. Bull.* 100 (2), 139–156. <https://doi.org/10.1037/0033-2909.100.2.139>.
- Vilarroya, O., 2017. Neural representation. A survey-based analysis of the notion. *Front. Psychol.* 8 (1458) <https://doi.org/10.3389/fpsyg.2017.01458>.
- 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. <https://doi.org/10.1016/j.neuroimage.2015.12.012>.
- Wiestler, T., Diedrichsen, J., 2013. Skill learning strengthens cortical representations of motor sequences. *eLife* 2 (2), 1–20. <https://doi.org/10.7554/eLife.00801>.
- Wilson, S.M., Molnar-Szakacs, I., Iacoboni, M., 2008. Beyond superior temporal cortex: intersubject correlations in narrative speech comprehension. *Cerebr. Cortex* 18 (1), 230–242. <https://doi.org/10.1093/cercor/bhm049>.
- Wing, E.A., Ritche, M., Cabeza, R., 2015. Reinstatement of individual past events revealed by the similarity of distributed activation patterns during encoding and retrieval. *J. Cognit. Neurosci.* 27 (4), 679–691. https://doi.org/10.1162/jocn_a.00740.
- 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. <https://doi.org/10.1523/JNEUROSCI.2324-16.2017>.
- Xue, G., 2018. The neural representations underlying human episodic memory. *Trends Cognit. Sci.* 22 (6), 544–561. <https://doi.org/10.1016/J.TICS.2018.03.004>.
- Ye, Z., Zhu, B., Zhuang, L., Lu, Z., Chen, C., Xue, G., 2016. Neural global pattern similarity underlies true and false memories. *J. Neurosci.* 36 (25), 6792–6802. <https://doi.org/10.1523/JNEUROSCI.0425-16.2016>.
- 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. <https://doi.org/10.1093/cercor/bhx202>.