

# UC Irvine

## UC Irvine Previously Published Works

### Title

Intersubject similarity in neural representations underlies shared episodic memory content.

### Permalink

<https://escholarship.org/uc/item/38p5p1q6>

### Journal

Proceedings of the National Academy of Sciences of USA, 120(35)

### Authors

Sheng, Jintao

Wang, Sisi

Zhang, Liang

et al.

### Publication Date

2023-08-29

### DOI

10.1073/pnas.2308951120

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives License, available at

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Peer reviewed



# Intersubject similarity in neural representations underlies shared episodic memory content

Jintao Sheng<sup>a</sup> , Sisi Wang<sup>a</sup>, Liang Zhang<sup>a</sup> , Chuqi Liu<sup>a</sup>, Liang Shi<sup>a</sup>, Yu Zhou<sup>a</sup>, Huinan Hu<sup>a</sup>, Chuansheng Chen<sup>b</sup>, and Gui Xue<sup>a,c,1</sup>

Edited by Michael B. Miller, University of California, Santa Barbara, CA; received May 28, 2023; accepted July 5, 2023 by Editorial Board Member Michael S. Gazzaniga

Individuals generally form their unique memories from shared experiences, yet the neural representational mechanisms underlying this subjectiveness of memory are poorly understood. The current study addressed this important question from the cross-subject neural representational perspective, leveraging a large functional magnetic resonance imaging dataset ( $n = 415$ ) of a face–name associative memory task. We found that individuals' memory abilities were predicted by their synchronization to the group-averaged, canonical trial-by-trial activation level and, to a lesser degree, by their similarity to the group-averaged representational patterns during encoding. More importantly, the memory content shared between pairs of participants could be predicted by their shared local neural activation pattern, particularly in the angular gyrus and ventromedial prefrontal cortex, even after controlling for differences in memory abilities. These results uncover neural representational mechanisms for individualized memory and underscore the constructive nature of episodic memory.

episodic memory | memory encoding | neural representation | intersubject analysis | fMRI

When presented with the same stimuli and instructions for a memory task, participants vary significantly in how many items they can remember (i.e., memory ability) and which items/details they remember (i.e., memory content) (1–5). Individual differences in memory abilities have been linked to many factors, such as the processing strategies (3), cognitive factors (6), representational dimensions (4), and neural anatomies (7) and activities (8, 9). Nevertheless, little is known about factors involved in individual differences in memory content (e.g., why participants of similar memory ability may remember different items).

Using within-subject subsequent memory paradigm (i.e., contrasting remembered vs. forgotten items) to examine the brain activity (10) and neural pattern similarity (11) during encoding, extant studies have revealed two important factors that determine which materials would be remembered. One factor is the fluctuation of the attentional state during the experiment (12). It has been argued that a higher level of attention would increase both the readiness to learn (13) and the fidelity of stimulus representations (14–16), which in turn would lead to better memory. Extending the attention account of the subsequent memory effect to individual differences in memory content, we would expect that participants with different patterns of attention fluctuation would remember different items.

In addition to attention (and its consequences for readiness to learn and the fidelity of representations), emerging studies have revealed another important factor influencing memory, i.e., the constructive neural transformation of the learning materials (17–19). For example, one recent study revealed that greater constructive neural transformation from perceptual to semantic representations during encoding was associated with better subsequent memory (20). Since this reconstruction involves an effective interaction between the learning materials and each participant's existing long-term knowledge (21), we expect each participant to conduct their own unique neural transformation of the stimuli and consequently form their unique memory content. Consistently, it has been shown that individuals' unique neural organization of semantic memory predicted their false memory better than did the group-averaged memory representations (2).

In sum, existing studies suggest that individualized memory content might be due to each individual's spontaneous and desynchronized fluctuation of attention and/or their unique constructive neural transformation. One promising approach to examining these two factors' roles in individualized memory content is to use a cross-participant analysis of shared memory content (SMC) and neural responses. This approach involves two steps. First, the degree of SMC is quantified for each pair of participants, which theoretically ranges from 0% (when two participants remember none of the same items) to 100% (when two participants remember exactly the same items). Second, the neural

## Significance

One common observation in memory research is the significant individual differences in memory. Such differences have traditionally been attributed to the spontaneous fluctuation of attention during encoding. Applying the cross-subject neural representational analytical approach to simple and static stimuli, we found several factors that contributed differentially to individuals' memory ability and content. Specifically, the individual-to-group synchronization in brain activations, which may reflect the fluctuation of attention, predicted individuals' memory ability. In contrast, the common activation patterns between participants, which reflect their shared reconstructions of learning materials, predicted their shared memory content. These results have both theoretical and methodological implications for our understanding of the nature of human episodic memory.

Author contributions: J.S. and G.X. designed research; J.S. and G.X. performed research; J.S., S.W., L.Z., C.L., L.S., Y.Z., and H.H. contributed new reagents/analytic tools; J.S. and G.X. analyzed data; C.C. and G.X. review and editing; and J.S., C.C., and G.X. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission. M.B.M. is a guest editor invited by the Editorial Board.

Copyright © 2023 the Author(s). Published by PNAS. This article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/) (CC BY-NC-ND).

<sup>1</sup>To whom correspondence may be addressed. Email: [gxue@bnu.edu.cn](mailto:gxue@bnu.edu.cn).

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2308951120/-/DCSupplemental>.

Published August 21, 2023.

mechanisms are examined by linking the similarity in memory content to the similarity in neural responses, including the inter-subject synchronization of neural activities (22) and shared neural representations (23–25).

Using cross-participant analyses of neural responses, existing studies have revealed novel cross-participant neural indices (neural synchronization and shared neural representations) that predict subsequent memory and the engagement of attention. Greater synchronization in the brain responses (22, 26–28) and more shared neural representations (23, 25) have been found for subsequently remembered items than for forgotten items. It has been argued that synchronized activity and shared representations indicate a higher level of attention (24, 28, 29). Consistently, research has shown that the degree of intersubject synchrony is positively correlated to measures of attention such as the self-reported level of task engagement (28) and interest in the stimuli (26). The synchrony is reduced when participants shift their attention away from external stimuli toward an internal task (27).

The current study used the cross-participant approach to examine individual differences in memory content as well as those in memory ability. In addition to examining the group-averaged subsequent memory effect, we further used an individual difference approach to link the intersubject (for each pair of participants) neural measures to their behavioral performance. Importantly, from the constructive perspective of memory, the pair-level (but not the group-averaged) measure of intersubject neural synchronization and shared neural representations may reflect either 1) their engagement of common cognitive processes and neural representations or 2) their shared reconstruction of learning materials that differs from that of other pairs of participants. One way to disambiguate these mechanisms is to separately examine individual-to-group and intersubject similarities. Specifically, as the group-averaged neural activity and representations are supposed to reflect the canonical stimulus processing, the individual-to-group similarity could reflect overall engagement of attention and the fidelity of stimulus representations, thus would predict individuals' memory abilities (30). In contrast, the intersubject similarity, particularly after controlling the group-averaged neural activity and representations, would reflect each pair of participants' shared reconstruction of learning materials and predict their SMC (2).

Using a large fMRI dataset ( $n = 415$ ) of a face–name associative memory task, the current study tested these hypotheses by systematically examining the roles of individual-to-group similarity and intersubject similarity in accounting for individual differences in memory content and ability. Our results revealed that memory ability was primarily predicted by individual-to-group similarity, whereas SMC was primarily predicted by intersubject similarity. These results uncover dissociated intersubject neural mechanisms for individualized memory ability and memory content and contribute to a deeper mechanistic understanding of individual differences in episodic memory.

## Results

### Individual Differences in Memory Ability and Memory Content.

In this study, participants ( $n = 415$ ) were asked to remember 30 unfamiliar face–name pairs. Each face–name association was studied twice within one scanning run with an interrepetition interval ranging from 8 to 15 trials. To better measure individual differences, the stimuli were presented in the same order, and participants were asked to perform the same encoding task (i.e., making a subjective judgment on the fitness of face–name associations). We used a slow event-related design (12 s for

each trial) to better estimate the single-trial blood oxygen level-dependent (BOLD) responses (Fig. 1A). To prevent participants from further processing the face–name pairs after encoding, they were asked to perform a self-paced perceptual orientation judgment task for 7.5 s before the subsequent trial started.

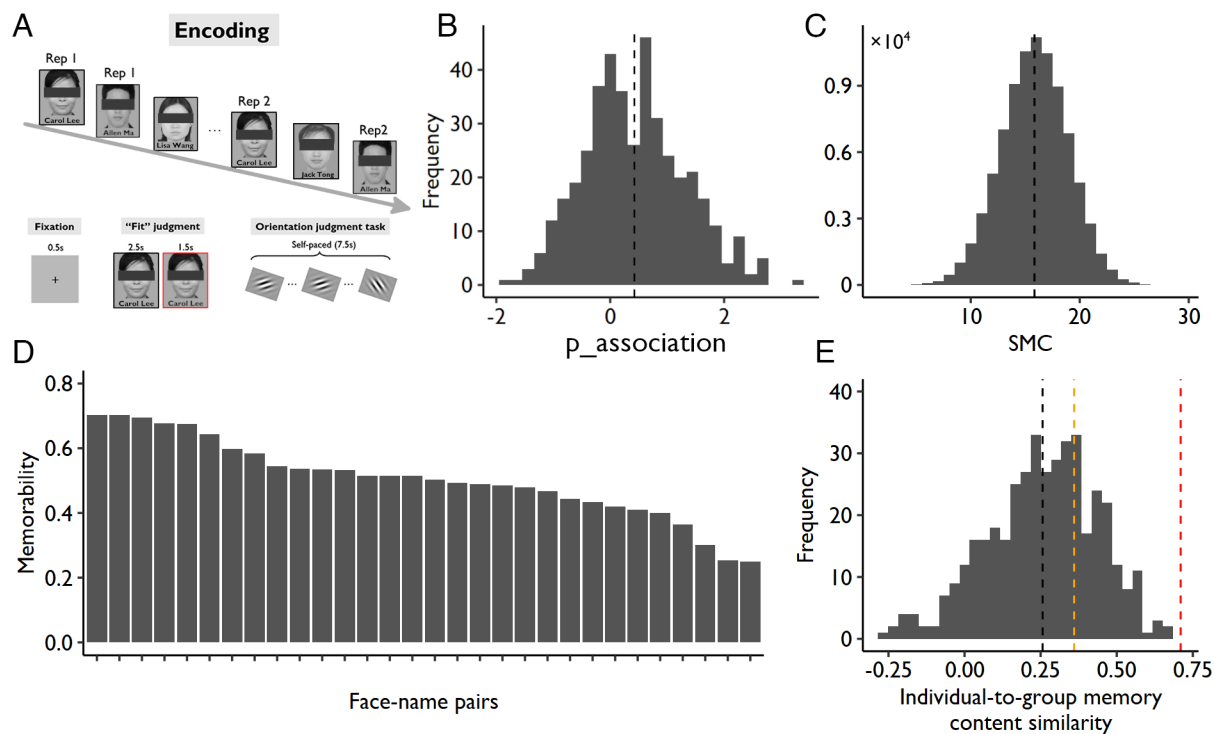
Subjects took a memory test approximately 24 min after learning and were asked to choose the correct given name for each face from three candidate names or indicate “new” if they had not studied the face before. To quantify memory ability, we calculated the associative recognition rate ( $p_{\text{association}}$ ) by using the following formula:  $p_{\text{association}} = Z(\text{associative hit rate}) - Z(\text{associative miss rate})$ , which was the proportion of recognized faces associated with correct names (i.e., associative hit: choosing the correct name for studied face) subtracted by the proportion of recognized faces associated with incorrect names (i.e., associative miss: choosing the incorrect name for studied face). As shown in Fig. 1B, memory ability varied significantly across participants, ranging from  $-1.83$  to  $3.32$ , with a mean  $\pm$  SD of  $0.42 \pm 0.89$ , in a near-normal distribution (skewness =  $0.30$ , kurtosis =  $-0.12$ ).

To quantify the degree of SMC across participants, we first defined the old faces recognized with correct names as remembered (i.e., scored 1) and all other old faces as forgotten (i.e., scored 0) and then calculated the Manhattan distance (MD) of the memory score across all items ( $n = 30$  items) for each pair of participants. The MD is the sum of the absolute difference of each item's memory score between two participants, which is calculated using the following formula:  $MD = |x_1 - y_1| + |x_2 - y_2| + \dots + |x_n - y_n|$ , in which  $x, y$  represent two different participants and  $n$  represents the number of items. MD in memory content indicates the level of nonSMC, so we reversed this index by subtracting it from the theoretical maximal distance (i.e., 30) and created an intuitive index of SMC (higher numbers indicate more overlapping memory contents between two participants). As shown in Fig. 1C, SMC varied significantly across participant pairs (ranging from 2 to 29 with a mean  $\pm$  SD of  $15.848 \pm 3.067$ ), and the data also followed a normal distribution (skewness =  $0.073$ , kurtosis =  $-0.048$ ). Finally, we calculated the averaged memory score across participants for each item (*Methods*), which reflects the memorability of the materials (Fig. 1D). We found that the individuals' memory content did not resemble the pattern of memorability (Fig. 1E), again suggesting significant variabilities in individuals' memory content.

### Intersubject Pattern Similarity (ISPS) Predicted Subsequent Memory.

We then examined whether intersubject neural representations could predict individual differences in memory ability and content. As a first step, we examined whether we could replicate the subsequent memory effect from an intersubject perspective. Previous studies using naturalistic stimuli have found that subsequently remembered events showed a higher intersubject neural signal synchronization (29, 31) and ISPS (24) than did forgotten events. Since we used faces as stimuli and each item was only presented for a few seconds, the current study examined the representational patterns instead of event-level time series. Specifically, we examined whether subsequently remembered items showed greater ISPS than did forgotten items.

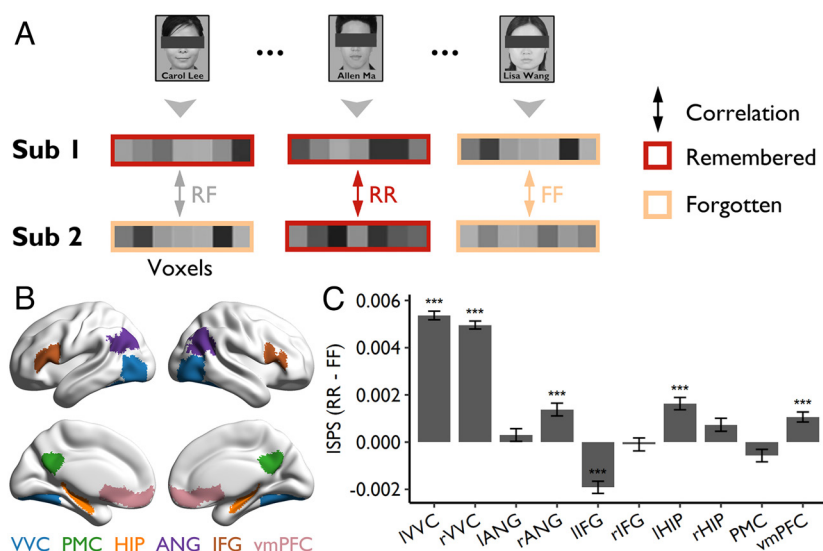
For each item, we first averaged brain activity across the two repetitions and then calculated the ISPS (Pearson  $r$ ) between two participants across voxels in a given region of interest (ROI). The correlations were then Z-transformed and averaged by type of items: those remembered by both participants (RR), those forgotten by both participants (FF), and those remembered by one participant but forgotten by the other (RF) (Fig. 2A). For the 85,002 participant pairs (out of 85,905) who have at least one



**Fig. 1.** Experimental design and behavioral results of Exp 1. (A) Experimental design of the face-name associative memory task. (B) Distribution of memory ability (i.e.,  $p_{\text{association}}$ , associative recognition rate); the black vertical dotted line represents the mean of all participants. (C) Distribution of the SMC; the black vertical dotted line represents the mean of all participant pairs. (D) The averaged memory performance across participants (i.e., memorability) of each face-name pair (30 pairs in total). (E) The distribution of correlations between each individual's memory profile and the group-averaged profile (i.e., memorability); the black dotted line represents the mean correlation across participants (i.e.,  $R = 0.26$ ); the orange dotted line represents the correlation coefficient that is statistically significant at  $P < 0.05$  (i.e.,  $R = 0.36$ ); the red dotted line represents the correlation coefficient that explains 50% of the variance (i.e.,  $R = 0.71$ ).

RR and FF trial, there were on average  $8.06 (\pm 3.96)$ , ranging from 1 to 25) RR trials and  $7.79 (\pm 3.81)$ , ranging from 1 to 24) FF trials. Since differences in the number of items between participants might impact the reliability of ISPS estimation, we excluded the participant pairs with fewer than 6 RR or FF items, and then matched the number of trials in each condition by

random sampling for the remaining 35,072 participant pairs. Following previous studies (23, 25, 32–35), we focused our analyses on several predefined ROIs implicated for neural representations of episodic memory, including bilateral ventral visual cortex (VVC), bilateral hippocampus (HIP), bilateral angular gyrus (ANG), bilateral inferior frontal gyrus (IFG), the posterior



**Fig. 2.** ISPS and subsequent memory. (A) Schematic diagram depicting the calculation of ISPS for each item. (B) The location of the predefined ROIs. These regions were mapped onto the cortical surface using BrainNet Viewer (36). (C) The subsequent memory effect [higher ISPS for items remembered by both participants (RR) than that for items forgotten by both participants (FF)] was found in bilateral VVC, right ANG, left HIP, and vmPFC, whereas an opposite effect was found in the left IFG. Error bars represent SEs of the mean across participant pairs;  $*P_{\text{Holm}} < 0.05$ ,  $**P_{\text{Holm}} < 0.01$ ,  $***P_{\text{Holm}} < 0.001$ . lVVC, left ventral visual cortex; rVVC, right ventral visual cortex; lANG, left angular gyrus; rANG, right angular gyrus; lIFG, left inferior frontal gyrus; rIFG, right inferior frontal gyrus; lHIP, left hippocampus; rHIP, right hippocampus; PMC, posterior medial cortex; vmPFC, ventromedial prefrontal cortex.

medial cortex (PMC), and the ventromedial prefrontal cortex (vmPFC) (Fig. 2*B* and *Methods*).

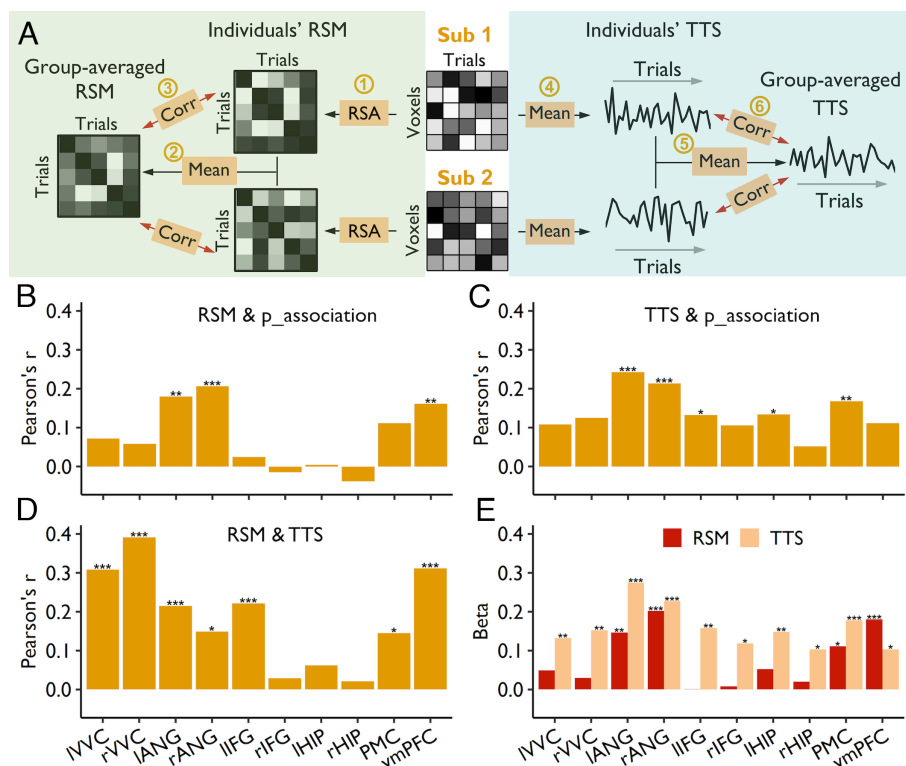
We found greater ISPS for the RR items than for the FF items in several brain regions, including the left VVC ( $\Delta_{RR-FF} = 0.0054$ ,  $P_{Holm} < 0.001$ , the  $P_{Holm}$  was determined by permutation test with 10,000 shuffles and Holm corrected for multiple comparisons, hereafter), right VVC ( $\Delta_{RR-FF} = 0.0050$ ,  $P_{Holm} < 0.001$ ), right ANG ( $\Delta_{RR-FF} = 0.0014$ ,  $P_{Holm} < 0.001$ ), left HIP ( $\Delta_{RR-FF} = 0.0016$ ,  $P_{Holm} < 0.001$ ), and vmPFC ( $\Delta_{RR-FF} = 0.0011$ ,  $P_{Holm} < 0.001$ ) (Fig. 2*C*). These results are consistent with a previous study (24) and confirm that shared representations across participants contribute to subsequent memory. A reversed effect was found in the left IFG ( $\Delta_{RR-FF} = -0.0019$ ,  $P_{Holm} < 0.001$ ; Fig. 2*C*), suggesting that remembered items have more unique neural representations in the left IFG across individuals than forgotten items.

**Individual-to-Group Similarities and Memory Ability.** Having replicated the intersubject subsequent memory effect, we examined our core hypotheses that individual-to-group similarities in neural representations and activities contributed to memory ability and content (this section), and that intersubject similarities contributed to memory content (next section).

**Individual-to-group similarity in neural representations predicted memory ability.** To examine individual-to-group similarity in neural representations, we first generated a canonical representational

matrix by averaging the neural response patterns across all participants. It should be noted that the first-order similarity (representational similarity for each item across participants, i.e., Fig. 2*A*) might be affected by the differences in the anatomy or anatomy-function alignments across participants. To mitigate this issue, we calculated the second-order pattern similarity (23, 25), reflecting the similarity of representational space when participants were processing the same set of stimuli. Specifically, we first calculated the representational similarity matrix (RSM, i.e., representational space) across 60 trials (30 items  $\times$  2 repetitions) for each ROI and each participant. These RSMs in each ROI were then averaged across all participants and used as the canonical representational pattern (Fig. 3*A*, *Left*).

To confirm that our averaged similarity matrix is a reliable measure of canonical representational pattern, we randomly split the participants into two equal-sized groups ( $n = 207$  for each group) and computed the correlation of the mean RSMs from the two groups. We performed this procedure 1,000 times and found that the group-averaged RSMs were highly reliable ( $R$ s ranged from 0.905 to 0.949) in all ROIs (SI Appendix, Fig. S1*A*). To examine the effect of sample size, we systematically increased the size of the selected sample from 20 to 200 in each half (step = 10; randomly sampled 1,000 times for each condition). The stability of the canonical representational pattern increased significantly with the sample size (SI Appendix, Fig. S1*B*). Taking the left and right



**Fig. 3.** Similarity to group-averaged representations and activities predicted individuals' memory ability. (A) Schematic diagram depicting the calculation of individual-to-group neural similarity, including RSM and trial-level time series (TTS). For RSM (Left), we first performed representational similarity analysis (RSA) across all trials to generate an RSM for each participant (Step 1). The RSMs were then averaged across all participants to obtain the group-averaged RSM (Step 2). Finally, the individual-to-group representational similarity was obtained by correlating the individual's RSM to group-averaged RSM (Step 3). For TTS (Right), we first averaged the activity across all voxels within an ROI to generate the trial-level time series (TTS) for each participant (Step 4). We then averaged the TTS across all participants to generate the group-averaged TTS (Step 5). Finally, individual-to-group similarity was obtained by correlating individuals' TTS to the group-averaged TTS (Step 6). Corr, correlation. (B) The individual-to-group similarity of representation in bilateral ANG and vmPFC was significantly correlated with individuals' associative memory performance. (C) Higher individual-to-group similarity of the TTS in bilateral ANG, left IFG, left HIP, and PMC was associated with better associative memory performance. (D) Individual-to-group similarities of the RSM and TTS were significantly correlated in most ROIs, except for bilateral HIP and right IFG. \* $P_{Holm} < 0.05$ , \*\* $P_{Holm} < 0.01$ , \*\*\* $P_{Holm} < 0.001$ . (E) The unique contribution of individual-to-group similarity of the RSM and TTS to associative memory performance; \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ . IVVC, left ventral visual cortex; rVVC, right ventral visual cortex; lANG, left angular gyrus; rANG, right angular gyrus; lIFG, left inferior frontal gyrus; rIFG, right inferior frontal gyrus; HIP, hippocampus; PMC, posterior medial cortex; vmPFC, ventromedial prefrontal cortex.



ANG as examples, we achieved a reliability higher than 0.8 with 200 participants in each group (SI Appendix, Fig. S1B).

We calculated the correlation between each individual's RSMs and the group-averaged RSMs. Across ROIs, we found that the averaged correlations in the VVC was significantly higher than those in the PMC (IVVC:  $t = 22.997$ ,  $P_{\text{Holm}} < 0.001$ ; rVVC:  $t = 22.088$ ,  $P_{\text{Holm}} < 0.001$ ), HIP (left:  $t = 14.920$ ,  $P_{\text{Holm}} < 0.001$ ; right:  $t = 17.457$ ,  $P_{\text{Holm}} < 0.001$ ), ANG (left:  $t = 26.298$ ,  $P_{\text{Holm}} < 0.001$ ; right:  $t = 25.013$ ,  $P_{\text{Holm}} < 0.001$ ), IFG (left:  $t = 21.919$ ,  $P_{\text{Holm}} < 0.001$ ; right:  $t = 21.686$ ,  $P_{\text{Holm}} < 0.001$ ), and vmPFC (IVVC:  $t = 14.718$ ,  $P_{\text{Holm}} < 0.001$ ; rVVC:  $t = 13.381$ ,  $P_{\text{Holm}} < 0.001$ ) (SI Appendix, Fig. S2A, Left), suggesting that neural representations were more variable across participants in the higher order regions than in the visual cortex.

Supporting our hypothesis, the correlational analysis showed that the more similar an individual's RSM was to the group-averaged RSM in the left ( $R = 0.179$ ,  $P_{\text{Holm}} = 0.002$ ) and right ANG ( $R = 0.206$ ,  $P_{\text{Holm}} < 0.001$ ) and vmPFC ( $R = 0.161$ ,  $P_{\text{Holm}} = 0.007$ ), the higher the individual's accuracy on the associative recognition task (Fig. 3B). A similar trend was found in the PMC ( $R = 0.112$ ,  $P = 0.023$ , uncorrected). Moreover, the individual-to-group representational similarity also predicted individuals' item memory ( $d'_{\text{item}}$ ) in the bilateral ANG (left:  $R = 0.205$ ,  $P_{\text{Holm}} < 0.001$ ; right:  $R = 0.229$ ,  $P_{\text{Holm}} < 0.001$ ), PMC ( $R = 0.171$ ,  $P_{\text{Holm}} = 0.003$ ), and vmPFC ( $R = 0.233$ ,  $P_{\text{Holm}} < 0.001$ ) (SI Appendix, Fig. S3, Left). Together, the above results show that individual-to-group similarity in neural representations predicts individuals' memory ability.

**Individual-to-group synchronization of neural activity predicted memory ability.** In a further analysis, we examined whether individual-to-group synchronization of neural activity also contributed to individuals' memory ability (30). We extracted the trial-level time series (TTS) in each ROI for each individual and then averaged them to form the group-averaged activity profile (Fig. 3A, Right). Similar to the group-averaged representational pattern, we found that the group-averaged activity profile was highly reliable ( $R_s$  ranged from 0.692 to 0.865) in all ROIs (SI Appendix, Fig. S1C). With 200 participants, we achieved a reliability of 0.657 and 0.756 in the right and left ANG, respectively (SI Appendix, Fig. S1D).

Also similar to the representational pattern, we found that individual-to-group synchronization of neural activity in the VVC was significantly higher than that in the PMC (IVVC:  $t = 6.229$ ,  $P_{\text{Holm}} < 0.001$ ; rVVC:  $t = 7.207$ ,  $P_{\text{Holm}} < 0.001$ ), bilateral HIP (left:  $t = 4.480$ ,  $P_{\text{Holm}} < 0.001$ ; right:  $t = 4.779$ ,  $P_{\text{Holm}} < 0.001$ ), and bilateral IFG (left:  $t = 3.016$ ,  $P_{\text{Holm}} = 0.008$ ; right:  $t = 7.953$ ,  $P_{\text{Holm}} < 0.001$ ), but was lower than that in the right ANG ( $t = 4.212$ ,  $P_{\text{Holm}} < 0.001$ ), and was comparable to that in the left ANG ( $t = -1.101$ ,  $P_{\text{Holm}} = 0.542$ ) and the vmPFC (IVVC:  $t = -0.142$ ,  $P_{\text{Holm}} = 0.887$ ; rVVC:  $t = 0.999$ ,  $P_{\text{Holm}} = 0.318$ ) (SI Appendix, Fig. S2A, Right).

By correlating the degree of individual-to-group synchronization of neural activity with individuals' memory ability, we found that in the left ( $R = 0.242$ ,  $P_{\text{Holm}} < 0.001$ ) and right ANG ( $R = 0.214$ ,  $P_{\text{Holm}} < 0.001$ ), left IFG ( $R = 0.132$ ,  $P_{\text{Holm}} = 0.043$ ), left HIP ( $R = 0.134$ ,  $P_{\text{Holm}} = 0.043$ ), and PMC ( $R = 0.167$ ,  $P_{\text{Holm}} = 0.005$ ), participants whose activity profiles were more synchronized to the group-averaged activity profile showed better associative memory (Fig. 3C). The individual-to-group neural synchronization was associated with better item memory performance in all ROIs (SI Appendix, Fig. S3, Right;  $R_s = 0.122 \sim 0.317$ ,  $P_{s_{\text{Holm}}} \leq 0.013$ ).

**Using both individual-to-group similarity measures to predict memory ability.** Interestingly, we found that the degree of

individual-to-group similarities in representation and the degree of the individual-to-group synchronization in neural activity were highly correlated in most ROIs ( $R_s = 0.145 \sim 0.391$ ,  $P_{s_{\text{Holm}}} \leq 0.012$ ), except for the right IFG ( $R = 0.029$ ,  $P_{\text{Holm}} = 1$ ) and the bilateral HIP (left:  $R = 0.062$ ,  $P_{\text{Holm}} = 0.617$ ; right:  $R = 0.021$ ,  $P_{\text{Holm}} = 1$ ) (Fig. 3D), suggesting that the fluctuations of attention could affect the fidelity of stimulus representation and contribute to memory strength (15). To dissociate the effects of brain activity and stimulus representations on memory ability, we constructed regression models that included both measures to assess their unique contributions. Overall, individual-to-group synchronization in neural activity contributed more to memory ability than did the individual-to-group representational similarity, except for the vmPFC (Fig. 3E). In the left and right ANG, PMC, and vmPFC, both measures uniquely contributed to memory ability (Fig. 3E), indicating that activity synchronization and representational similarity might capture different aspects of stimulus processing, all of which contribute to successful memory encoding (15).

#### **Individual-to-group similarities did not predict memory content.**

The above analyses supported our hypothesis that individual-to-group similarities in neural representations and activities predicted memory ability. Would they also predict one's memory content, such that individuals showing overall higher similarity to the group-averaged response also showed higher similarity to the group-averaged memory profile (i.e., the memorability of items)? We found that individual-to-group similarity of neural representations was not related to individuals' similarity to group-averaged memory pattern ( $R_s = -0.009 \sim 0.047$  for different ROIs,  $P_s > 0.415$ , uncorrected). Similarly, the individual-to-group synchronization of neural activities did not predict individuals' similarity to group-averaged memory pattern either ( $R_s = 0.051 \sim 0.130$  for different ROIs,  $P_{s_{\text{Holm}}} > 0.080$ ). These results suggest that individual-to-group similarity in neural response and representational pattern did not predict individuals' memory content.

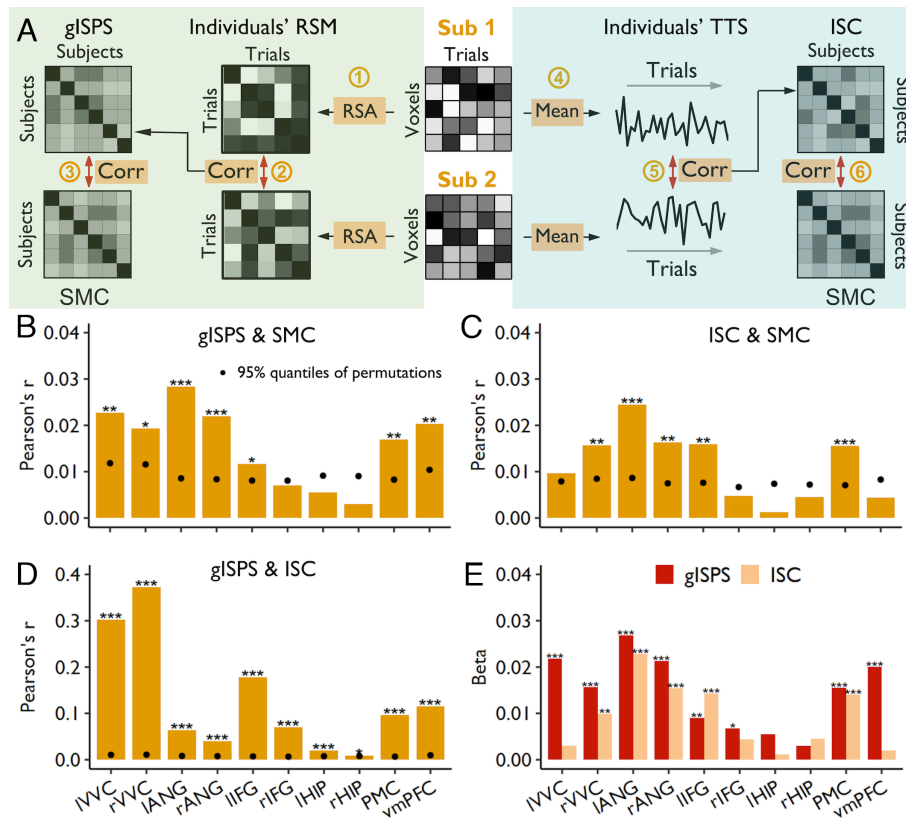
#### **Intersubject Similarities in Representations and Activities Predicted Memory Content.**

Thus far, we have shown that the individual-to-group similarities in representational patterns and neural activities predicted individuals' memory ability but not their memory content. In the following analysis, we tested the hypothesis that the intersubject similarity in neural representations and brain activities would predict memory content.

#### **Intersubject similarity in neural representations predicted SMC.**

As described earlier, the SMC is quantified as the degree of overlapping memory contents between two participants. The ISPS was quantified using the global ISPS (gISPS), which is the Pearson correlation of the representational pattern between each pair of participants (Methods and Fig. 4A, Left). This measure of ISPS is less affected by their differences in anatomical structure and anatomy-function correspondences for different participants.

Similar to the SMC reported in the previous section, we found that the gISPS varied significantly across participant pairs. Taking the left ANG as an example, we found that the gISPS varied from  $-0.260$  to  $0.396$  across participant pairs, with a mean of  $0.045 \pm 0.061$ . Further, we found that, similar to the results of individual-to-group similarity, the averaged gISPS in the VVC was significantly higher than that in the PMC (IVVC:  $\Delta = 0.043$ ,  $P_{\text{Holm}} < 0.001$ ; rVVC:  $\Delta = 0.039$ ,  $P_{\text{Holm}} < 0.001$ ), bilateral HIP (left:  $\Delta = 0.029$ ,  $P_{\text{Holm}} < 0.001$ ; right:  $\Delta = 0.032$ ,  $P_{\text{Holm}} < 0.001$ ), bilateral ANG (left:  $\Delta = 0.049$ ,  $P_{\text{Holm}} < 0.001$ ; right:  $\Delta = 0.044$ ,  $P_{\text{Holm}} < 0.001$ ), bilateral IFG (left:  $\Delta = 0.042$ ,  $P_{\text{Holm}} < 0.001$ ; right:  $\Delta = 0.039$ ,  $P_{\text{Holm}} < 0.001$ ), and vmPFC (IVVC:  $\Delta = 0.029$ ,  $P_{\text{Holm}} < 0.001$ ; rVVC:  $\Delta = 0.025$ ,  $P_{\text{Holm}} < 0.001$ ) (SI Appendix, Fig. S2B,



**Fig. 4.** Intersubject similarity in representations and activities predicted SMC. (A) Schematic diagram depicting the calculation of gISPS and intersubject correlation (ISC) in TTS and their correlations with the SMC. To examine the relationship between gISPS and SMC (Left), we first performed representational similarity analysis (RSA) across all trials to generate an RSM for each participant (Step 1). The resultant individuals' RSMs were then used to calculate cross-subject correlations to obtain gISPS (Step 2). Finally, the relationship between gISPS and SMC was examined by using the Mantel test (Step 3). To examine the relationship between ISC and SMC (Right), we first averaged the activity across all voxels within an ROI to generate TTS for each participant (Step 4). We then calculated cross-subject correlations of neural fluctuation by using the resultant individuals' TTSs to obtain ISC (Step 5). Finally, the correlation between ISC and SMC was calculated by using the Mantel test (Step 6). Corr, correlation. (B) The gISPS in most ROIs, including the bilateral VVC and ANG, left IFG, PMC, and vmPFC, was significantly correlated with SMC. (C) The ISC in the right VVC, bilateral ANG, left IFG, and PMC was significantly correlated with SMC. (D) The gISPS was significantly correlated with ISC.  $*P_{\text{Holm}} < 0.05$ ,  $**P_{\text{Holm}} < 0.01$ ,  $***P_{\text{Holm}} < 0.001$ . The black dots represent 95% quantiles of permutations; thus, the black dot within a bar means the significant effect for the Mantel test. (E) The unique contribution of gISPS and ISC to SMCs in bilateral ANG, right VVC, left IFG, and PMC;  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$ . lVVC, left ventral visual cortex; rVVC, right ventral visual cortex; lANG, left angular gyrus; rANG, right angular gyrus; lIFG, left inferior frontal gyrus; rIFG, right inferior frontal gyrus; lHIP, left hippocampus; rHIP, right hippocampus; PMC, posterior medial cortex; vmPFC, ventromedial prefrontal cortex.

Left), again indicating that the representations in the higher order brain regions are more uniquely transformed than the representations in the perceptual areas.

Since similarities in both memory content and representational pattern were assessed using distance measures, we used the Mantel test with permutation to examine their associations and determined the significance level (Methods). Specifically, we used Pearson correlation to measure the relationship between gISPS and SMC and used the permutation test to determine the statistical significance. We found a significant correlation between the degree of the SMC and shared representational patterns in most ROIs, including the left ( $R = 0.028$ ,  $P_{\text{Holm}} < 0.001$ ) and right ANG ( $R = 0.022$ ,  $P_{\text{Holm}} < 0.001$ ), left ( $R = 0.023$ ,  $P_{\text{Holm}} = 0.005$ ) and right VVC ( $R = 0.019$ ,  $P_{\text{Holm}} = 0.016$ ), vmPFC ( $R = 0.020$ ,  $P_{\text{Holm}} = 0.004$ ), PMC ( $R = 0.017$ ,  $P_{\text{Holm}} = 0.003$ ), and left IFG ( $R = 0.012$ ,  $P_{\text{Holm}} = 0.036$ ) (Fig. 4B).

We performed several control analyses to validate the above results. First, individual differences in memory ability could contribute to the measure of SMC ( $R = 0.40$ ,  $P < 0.001$ ). To control for individual differences in memory ability, we performed a partial Mantel test by including the cross-participant differences in memory ability (p\_association, associative recognition rate) as a covariate. The results again showed a significant positive correlation between the SMC and shared neural representations in the

left ( $R = 0.027$ ,  $P_{\text{Holm}} < 0.001$ ) and right ANG ( $R = 0.027$ ,  $P_{\text{Holm}} < 0.001$ ), left ( $R = 0.028$ ,  $P_{\text{Holm}} = 0.002$ ) and right VVC ( $R = 0.024$ ,  $P_{\text{Holm}} = 0.003$ ), vmPFC ( $R = 0.022$ ,  $P_{\text{Holm}} = 0.003$ ), PMC ( $R = 0.017$ ,  $P_{\text{Holm}} = 0.003$ ), and left IFG ( $R = 0.016$ ,  $P_{\text{Holm}} = 0.003$ ) (SI Appendix, Fig. S4, Left).

Second, to demonstrate that our results were robust to different ways of quantifying memory performance, we reanalyzed the data using a fine-grained measure of memory performance. The old faces that were recognized with the correct name (i.e., association correct) were scored as 2, those recognized as old but with an incorrect name (i.e., item correct) were scored as 1, and those recognized as new (i.e., forgotten) were scored as 0. Results again showed a significant correlation between the SMC and shared neural representations in the left ( $R = 0.057$ ,  $P_{\text{Holm}} < 0.001$ ) and right ANG ( $R = 0.049$ ,  $P_{\text{Holm}} < 0.001$ ), left ( $R = 0.051$ ,  $P_{\text{Holm}} = 0.005$ ) and right VVC ( $R = 0.049$ ,  $P_{\text{Holm}} = 0.005$ ), vmPFC ( $R = 0.058$ ,  $P_{\text{Holm}} < 0.001$ ), and PMC ( $R = 0.051$ ,  $P_{\text{Holm}} < 0.001$ ) (SI Appendix, Fig. S5, Left).

Third, to ensure that this shared memory was not associated with general processing strength due to attention and/or familiarity, we examined individual differences in the mean neural activation level in these ROIs. The differences in activation level across participants were not correlated with the degree of SMC ( $P_{\text{Holm}} > 0.9$ , corrected) in any regions.

Fourth, in the above analyses, the two repetitions of a given item were separately coded in the RSM to better model the context and learning effects. In an additional analysis, we first averaged the activation patterns of the two repetitions for each item and then constructed an RSM based on the activation patterns of 30 items. Using this RSM to calculate the gISPS, we found that the gISPS in the vmPFC ( $R = 0.017$ ,  $P_{\text{Holm}} = 0.018$ ) was significantly correlated with SMC. In addition, the gISPS in the left ( $R = 0.012$ ,  $P = 0.018$ , uncorrected) and right ANG ( $R = 0.01$ ,  $P = 0.030$ , uncorrected), left IFG ( $R = 0.009$ ,  $P = 0.035$ , uncorrected), and right VVC ( $R = 0.013$ ,  $P = 0.025$ , uncorrected) were marginally related to SMC (SI Appendix, Fig. S6). The weaker effect suggests that the context and repetition effects also contribute to individual differences in memory content.

Finally, to ensure that the shared representations between participants were not due to their similarity to the canonical representational pattern, we included the group-averaged RSM as a covariate when calculating the gISPS. Using this partial gISPS, which should reflect the unique neural transformations shared by each pair of participants, we still found that the gISPS was significantly related to SMC in most ROIs ( $R_s = 0.013 \sim 0.024$ ,  $P_{\text{Holm}} \leq 0.001$ ; SI Appendix, Fig. S7, Left), except the right IFG ( $R = 0.005$ ,  $P = 0.055$ , uncorrected) and bilateral HIP (left:  $R = 0.003$ ,  $P = 0.165$ , uncorrected; right:  $R = 0.0005$ ,  $P = 0.437$ , uncorrected). These results provide strong support for our hypothesis that the shared transformations of neural representations contribute to SMC.

**Intersubject neural synchronization had a limited predictive power of SMC.** To test whether two participants who showed synchronized fluctuation of attention would remember similar content, we also examined the intersubject correlation (ISC) of TTS (Methods and Fig. 4A, Right). Comparisons of ISC across different ROIs showed that the ISC in the VVC was significantly higher than that in the PMC (IVVC:  $\Delta = 0.017$ ,  $P_{\text{Holm}} < 0.001$ ; rVVC:  $\Delta = 0.020$ ,  $P_{\text{Holm}} < 0.001$ ), bilateral HIP (left:  $\Delta = 0.012$ ,  $P_{\text{Holm}} < 0.001$ ; right:  $\Delta = 0.014$ ,  $P_{\text{Holm}} < 0.001$ ), and bilateral IFG (left:  $\Delta = 0.009$ ,  $P_{\text{Holm}} < 0.001$ ; right:  $\Delta = 0.022$ ,  $P_{\text{Holm}} < 0.001$ ) (SI Appendix, Fig. S2B, Right). The right VVC showed higher ISC than the right ANG ( $\Delta = 0.013$ ,  $P_{\text{Holm}} < 0.001$ ) and vmPFC ( $\Delta = 0.004$ ,  $P_{\text{Holm}} < 0.001$ ), but the left VVC did not (lANG:  $\Delta = -0.003$ ,  $P_{\text{Holm}} < 0.001$ ; vmPFC:  $\Delta = 0.00001$ ,  $P_{\text{Holm}} = 0.995$ ). Again, these findings indicated that the brain activities in the higher order brain regions were more individualized than those related to the perceptual processing.

Using the Mantel test, we examined the correlations between synchronization of brain activity and SMC. This analysis revealed that the synchronization of brain activity in the left ( $R = 0.0245$ ,  $P_{\text{Holm}} < 0.001$ ) and right ANG ( $R = 0.0163$ ,  $P_{\text{Holm}} = 0.005$ ), PMC ( $R = 0.0155$ ,  $P_{\text{Holm}} < 0.001$ ), right VVC ( $R = 0.0157$ ,  $P_{\text{Holm}} = 0.008$ ), and left IFG ( $R = 0.016$ ,  $P_{\text{Holm}} = 0.003$ ) was significantly correlated with SMC (Fig. 4C). Validation analyses revealed that the associations remained significant in several ROIs after controlling for memory ability, including the left ( $R = 0.0224$ ,  $P_{\text{Holm}} = 0.002$ ) and right ANG ( $R = 0.0152$ ,  $P_{\text{Holm}} = 0.017$ ), left IFG ( $R = 0.0155$ ,  $P_{\text{Holm}} = 0.021$ ), and right VVC ( $R = 0.0177$ ,  $P_{\text{Holm}} = 0.017$ ) (SI Appendix, Fig. S3, Right), and remained significant in most ROIs when using fine-grained memory performance score, including the left ( $R = 0.057$ ,  $P_{\text{Holm}} = 0.002$ ) and right ANG ( $R = 0.032$ ,  $P_{\text{Holm}} = 0.003$ ), left ( $R = 0.027$ ,  $P_{\text{Holm}} = 0.015$ ) and right IFG ( $R = 0.016$ ,  $P_{\text{Holm}} = 0.034$ ), left ( $R = 0.030$ ,  $P_{\text{Holm}} = 0.006$ ) and right VVC ( $R = 0.044$ ,  $P_{\text{Holm}} = 0.002$ ), PMC ( $R = 0.030$ ,  $P_{\text{Holm}} = 0.002$ ), and vmPFC ( $R = 0.028$ ,  $P_{\text{Holm}} = 0.014$ ) (SI Appendix, Fig. S5, Right). However, when using partial correlation to control individuals' similarity to group-averaged activity profile, the resultant partial ISC barely predicted SMC, except a marginal effect in

the PMC ( $R = 0.009$ ,  $P_{\text{Holm}} = 0.060$ ) and another marginal effect in the left IFG ( $R = 0.008$ ,  $P_{\text{Holm}} = 0.078$ ) (SI Appendix, Fig. S7, Right). This result suggests that the effect of intersubject synchronization in neural activation on SMC was mainly due to their synchronization to the group-averaged neural activities.

**Using both intersubject measures to predict SMC.** We further found that the two intersubject neural measures (intersubject synchronization of brain activity and shared neural representations) were significantly correlated with each other in all ROIs ( $R_s = 0.009 \sim 0.372$ ,  $P_{\text{Holm}} \leq 0.025$ ; Fig. 4D). The highest correlation was found in the VVC, suggesting that synchronization in neural activity has the strongest effect in shaping neural representation in the visual cortex as compared to other regions. We further constructed regression models to examine whether intersubject shared representations and synchronized activities made unique contributions to SMC. The results showed that the shared representations in most ROIs (except the bilateral HIP) still contributed to SMC, and that their contributions were higher than those of synchronized activities, particularly in the bilateral ANG and VVC, right IFG, PMC, and vmPFC (Fig. 4E). These results together suggest that the shared representations in these regions mainly reflect the common transformation of the stimulus representations rather than synchronized attention, and the former contributes to SMC.

## Discussion

The current study aimed to uncover the neural mechanisms of individualized memory using a cross-participant approach. At the group level, we found that the ISPS for remembered items was greater than that for forgotten items in the right ANG, left HIP, and vmPFC, which is consistent with a recent study (24). Interestingly, unlike the previous study, we also found a subsequent memory effect of ISPS in the visual cortex (i.e., VVC), suggesting that perceptual representations of human faces are also important for subsequent memory. These findings extend previous studies that found greater intersubject synchronization of BOLD response for remembered items than for forgotten items (29, 31). We also found a reversed effect, i.e., greater ISPS for forgotten than remembered items, in the left IFG. Similarly, Koch and colleagues found a trend of reversed effect in the parahippocampal areas (24). This reversed effect could either suggest that these shared representations might be disruptive to memory formation (24) and/or that more unique representations are beneficial to memory. Future studies should further replicate these findings and investigate the underlying mechanisms.

Given the shared experiences (i.e., stimuli and context) for all participants during learning, several studies have suggested that the synchronized activities and the shared representations might reflect higher degree of attention and greater fidelity of stimulus representations (24, 28, 29). Alternatively, according to the transformative perspective of memory (17–19), intersubject representational similarity could also reflect transformations shared only within each pair of participants, which differ from those of other participants. To test these alternative perspectives, the current study first generated canonical representational patterns (i.e., RSMs) and TTS based on a large sample of 415 participants and clearly demonstrated these indices' reliability using random splits. Interestingly, we found that the group-averaged pattern was highly variable with a sample size like those used in typical memory imaging studies (e.g., around 30), suggesting that the group-averaged response patterns from small samples are either noisy or unique to this particular group of participants.



Using the indices of RSMs and TTS, we revealed distinct neural mechanisms for memory ability and content. First, we found that the individual-to-group synchronization of brain activity predicted individuals' memory ability. Previous studies found that the group-averaged time course might reflect the fluctuation of attention and the strength of stimulus processing (28, 30, 37–40). For example, greater and more widespread brain synchronizations across subjects were found when watching attention-grabbing Hollywood-style commercial movies as compared to less exciting real-life, unedited videos (38). Similarly, video clips with higher emotional arousal also elicit stronger inter-subject synchronization than the less emotional clips (40). Such fluctuations in attention can affect not only our ongoing perceptual experience but also our later memory performance (12, 28). A recent study recorded spontaneous fluctuations in participants' attention during memory encoding and retrieval tasks and found that loss of attention was significantly associated with reductions in target encoding and memory signal (13). Our finding is consistent with these observations and further suggests that individual-to-group synchronization might also index subjects' overall task engagement and predict memory performance. Supporting this interpretation, we also found higher individual-to-group synchronization was associated with greater fidelity of stimulus processing as reflected by individual-to-group similarity of neural representation. Future studies could further examine the relationship between individual-to-group synchronization and attention by either manipulating or measuring task engagement (26–28) or include concurrent measurements of heart rate and pupil diameter (41).

Second, we found that the individual-to-group similarity of representational patterns also contributed to memory ability, although the effect was not as strong as that of synchronized neural activities. Two specific findings are worth emphasizing. The first one is that the similarity to canonical response was correlated with the similarity to canonical representations, particularly in the VVC and vmPFC, suggesting that attention could contribute to the fidelity of representations (14–16). The second finding is that both the similarity to canonical activities and the similarity to canonical representations made unique contributions to memory ability, extending the previous findings from within-subject studies (15). Together these two findings based on the intersubject analyses emphasize the interdependent and complementary roles of activation level and fidelity of representation in successful memory encoding.

Third, we found that the intersubject similarity of neural representations might be a better indicator of individualized memory content than is the individual-to-group similarity. In other words, each pair of participants' SMC is not due to the two individuals' similarity to the canonical representations, but rather due to the pair's unique transformation of representations. Although there was also some evidence that the intersubject synchronization of neural activity contributed to the SMC, the effect was smaller and mainly due to their similarity to group-averaged activity profile (probably reflecting shared attention to the stimuli) rather than due to each pair's unique attention fluctuation. Despite a significant correlation between the synchronization of neural activities and the shared representations, we still found that shared neural representations made unique contributions to SMC. It is also notable that the similarity of representations in ANG, IFG, HIP, PMC, and vmPFC was less affected by the synchronized brain activity than that in VVC, suggesting that the representations in these higher order regions are less affected by the brain activities and processing strength.

More broadly, the findings of this study contribute to the literature on the interactive and constructive nature of human episodic memory, namely, memory encoding involves the constructive neural transformation of the learning materials via the interaction with

existing long-term knowledge (6, 19–21, 42–44). For example, one study found that participants with the same interpretation of ambiguous stories showed similar brain activity in the posterior cingulate cortex and the other regions of the default mode network (DMN) (45). Furthermore, participants with more similar interpretations of the storyline and narrative also showed more similar representational similarities in the PMC (43, 44). Even using simple words (2) and pictures (25) as learning materials, significant individual differences in representational patterns have been found, and these differences are associated with individuals' unique memory (2).

Consistently, our findings suggest that the shared representations reflect the individualized transformation of memory representations that contribute to personalized memory content. The regions where the representations showed strong and reliable effects of SMC included the ANG and vmPFC, whose representations showed more robust individual differences and were less affected by neural activities as compared to those in the VVC. Existing studies have shown that the representation in the ANG is identity-specific and invariant to viewpoints (46), as well as sensitive to semantic similarity (47). In addition, its representation is significantly transformed from encoding to retrieval (25, 48) and is strongly modulated by the task goals (49). Similarly, the vmPFC is consistently involved in the representation of structured knowledge, such as cognitive maps (50) and schemas (51) [also see Xue (19) for a recent review]. Together, our results suggest that personalized memory content arises from the individualized transformation of memory representations from the shared sensory inputs.

The current study also has two methodological implications. First, previous studies examining intersubject neural similarity mainly used naturalistic stimuli, such as stories and movies. Although the naturalistic stimuli have high ecological validity and are able to capture highly similar behavioral and neural responses, they may not be good choices when examining the subsequent memory effect and its individual differences. For example, many stimulus features (i.e., the presence of cuts and changes in scenes and characters; the length of each event; film features, including shot scale, music, and dialog; and low-level visual features, such as luminance and contrast) could potentially influence neural activity and memory. The complexity of these effects also makes it challenging to examine the neural mechanisms of subsequent memory. In addition, these strong stimulus effects would have reduced individual differences in both memory performance and neural representations. To mitigate these effects, the current study used simple stimuli, such as human faces or words, revealing significant individual differences in memory ability and content, as well as in neural activation patterns.

Second, intersubject analysis has been increasingly used to examine the pattern of stimulus processing (39, 52). In particular, we can use one participant's brain activity or neural representation as a model for a second participant's brain activity or neural representation, representing a powerful data-driven technique for detecting neural responses (52). The idea is that as long as two participants receive the same input at the same order, any shared variance must be due to their stimulus processing (52). This approach has been shown to be more sensitive than deconvolution/GLM-based analyses in some circumstances (53). Unlike GLM-based approaches, this approach abolishes the need to explicitly model the task parameters and modulation functions, which is challenging for complex designs (54). When examining the fidelity of representation, the intersubject analysis no longer requires within-subject repetition of stimuli (55), which is ideal for experimental designs when stimuli cannot be repeated, or when repetition could affect the underlying cognitive and neural mechanisms. Despite these advantages, our results suggest that the intersubject similarity could reflect both

canonical representations shared by the population and the unique representations that are only shared by specific participant pairs. This issue could be mitigated by calculating the group-averaged response patterns with a large sample.

Taken together, the current study uncovered intersubject neural representational mechanisms for individualized memory content. These findings emphasize the constructive and interactive nature of human episodic memory (19, 21). Future studies should further investigate the interaction between preexisting knowledge and state-dependent factors in shaping the neural representational patterns during learning, which should lead to a better understanding and more accurate prediction of individualized memory.

## Methods

**Participants.** Five hundred healthy Han Chinese college students were recruited for this study as a part of a large cohort study in China, i.e., the Cognitive Neurogenetic Study of Chinese Young Adults. All participants provided informed consent, and the study was approved by the Institutional Review Board of the State Key Laboratory of Cognitive Neuroscience and Learning at Beijing Normal University. All participants self-reported as having normal or corrected-to-normal vision and no history of neurological or psychiatric problems. Four hundred seventy-eight volunteers participated in the face-name associative memory task.

Seventy-three participants were excluded due to incomplete imaging data ( $n = 1$ ), large head motion (mean framewise displacement  $> 0.3$ ,  $n = 9$ ), or inconsistent learning sequences ( $n = 53$ ). As a result, 415 participants (214 females, mean age =  $21.33 \pm 2.14$  y) were included in the analyses.

### Experimental Design: Face-Name Associative Memory Task.

**Materials.** The experimental stimuli consisted of 30 unfamiliar face photographs (15 men and 15 women, chosen from the internet). Fictional first names (e.g., "Aiguo") and common surnames (e.g., "Li") were assigned to the faces and were used for encoding. An additional 20 unfamiliar faces (10 men and 10 women) were used for the memory test. All face pictures were converted into grayscale images of the same size ( $256 \times 256$  pixels) on a gray background. All of the faces had neutral facial expressions in the pictures. Four additional face-name pairs were used in the practice session.

**Procedures.** During fMRI data acquisition, participants were asked to remember 30 unfamiliar face-name pairs. For each face-name pair, participants were instructed to remember the name associated with the face for a later memory test by pressing a button to indicate whether each name "fits" the face (right index finger = the name fits the face; right middle finger = the name does not fit the face). Participants were informed that it was a purely subjective judgment designed to help them memorize the association between faces and names. Each face-name pair was presented twice with an inter-repetition interval ranging from 8 to 15 trials. The stimuli were presented in the same order to all participants. A slow event-related design (12 s for each trial) was used in this study to obtain better estimates of the single-trial BOLD response associated with each trial (Fig. 1A). Each trial started with a 0.5-s fixation; the picture was presented for 2.5 s. Then, the frame of the picture turned red, indicating that the participants should press the button to indicate their response within 1.5 s. To prevent further encoding of the pictures, participants were asked to perform a perceptual orientation judgment task for 7.5 s. In this task, a Gabor image tilting  $45^\circ$  to the left or the right was presented on the screen, and participants were asked to identify the orientation of the Gabor image as quickly as possible by pressing a corresponding button. The next Gabor image appeared 0.2 s after the response. A self-paced procedure was used for this task to be engaging to the subjects. Participants completed only one run of the encoding task, which lasted 12 min. Before the scan, participants completed a practice session to familiarize themselves with the task and key responses. Participants were informed that a subsequent memory test would occur later, but they were not informed of the specific procedure of the memory test.

Approximately 24 min later, during which the participants performed some other fMRI experiments (an n-back task and a decision-making task),

participants were asked to complete the retrieval test in the fMRI scanner. The stimuli for the retrieval test consisted of the same 30 face pictures from the encoding stage and 20 new face pictures. All the pictures were randomly mixed. For each face picture, three old names from the encoding stage without a surname (the correct name that was paired with the face during encoding) and two other names that were paired with different faces during encoding) and the option of "new [face]" were presented underneath the face (Fig. 1A). The location of the correct name was counterbalanced for equal numbers on the three locations of the names and the new option always appeared at the rightmost position. Participants were asked to judge whether they had seen the face and to indicate the corresponding correct name if the face was old; otherwise, they pressed the "new [face]" button if the face was new. Each trial started with a 0.5-s fixation period followed by a picture presentation for 4 s. Participants were asked to press the button to indicate their response within 4 s. The responses were made using a button box (the left index finger corresponded to the first name, the left middle finger corresponded to the second name, the right index finger corresponded to the third name, and the right middle finger corresponded to "new [face]"). After the presentation of each face, a fixation cross of jittered duration (0 to 8 s) was placed on the center of the screen. This testing run lasted approximately 5 min.

**Quantification of Memory Ability.** We used the associative recognition rate ( $p_{\text{association}}$ ) as an indicator of memory ability. It was quantified as the proportion of recognized faces associated with the correct names (i.e., choosing the correct names for studied faces) minus the proportion of recognized faces associated with incorrect names (i.e., choosing one of the incorrect names for studied faces). In addition, we calculated the  $d'$  of item memory ( $d'_{\text{item}}$ ) for each individual, which is the difference between item hit rate (correct recognition of old faces regardless of the correctness of names) and false-alarm rate (assigning a name to a new face):  $d'_{\text{item}} = Z(\text{hit rate}) - Z(\text{false-alarm rate})$ . The memorability of each face-name pair was computed by averaging the performance of associative recognition for each item across all participants ( $n = 415$ ).

**Quantification of SMC.** We used the MD to characterize the distance of the memory content between two participants. The MD is the sum of the absolute difference of each item's memory score between two participants, calculated using the following formula:  $\text{MD} = |x_1 - y_1| + |x_2 - y_2| + \dots + |x_n - y_n|$ , in which  $x, y$  represent two different participants and  $n$  represents the number of items. The *vegdist* function of the *vegan* toolkit (56) in *R* was used to calculate the MD of the items' memory scores for all participant pairs. Since greater MD across two participants indicates less overlapped memory content, we then quantified SMC by subtracting MD from the maximum possible distance, which is the number of items multiplied by the maximum possible distance of each item. A larger value of SMC suggests that the given two participants tended to remember a larger number of the same items.

We used two indices (i.e., associative memory and fine-grained memory precision) to measure SMC across participants. For associative memory, studied faces recognized with the correct name were defined as remembered items and scored as 1, whereas those recognized with an incorrect name or judged as new were defined as forgotten items and scored as 0. Then, the MD matrix across all participants was obtained to index the cross-participant memory content dissimilarity. The degree of associative SMC was quantified as the maximum possible distance (i.e., 30) minus MD.

For fine-grained memory performance, we defined different memory strengths for each studied face. Specifically, studied faces recognized with the correct name were defined as associative memory with a memory strength of 2. In contrast, those recognized with an incorrect name were defined as item memory with a memory strength of 1, and those judged as new were defined as forgotten with a memory strength of 0. Subsequently, the MD for each participant pair was calculated as the measurement of fine-grained memory content distance. Similarly, the degree of SMC was quantified as the maximum possible distance (i.e., 60) minus MD.

### Neuroimage Data Collection and Processing.

**MRI data acquisition.** Image data were acquired using a 3.0 T Siemens MRI scanner in the Brain Imaging Center at Beijing Normal University. A single-shot T2\*-weighted gradient-echo, EPI sequence was used for the functional

scan with the following parameters: TR/TE = 2,000 ms/25 ms, FA = 90°, FOV = 192 × 192 mm<sup>2</sup>, and 64 × 64 matrix size with a resolution of 3 × 3 mm<sup>2</sup>. Forty-one 3-mm transversal slices parallel to the AC-PC line were obtained to cover the cerebrum and partial cerebellum. The anatomical scan was acquired using the T1-weighted magnetization prepared rapid acquisition gradient-echo sequence with the following parameters: inversion time (TI) = 1,100 ms; TR/TE/FA = 2,530 ms/3.39 ms/7°, FOV = 256 × 256 mm<sup>2</sup>, matrix = 192 × 256, slice thickness = 1.33 mm, 144 sagittal slices, and voxel size = 1.3 × 1.0 × 1.3 mm<sup>3</sup>.

**fMRI data preprocessing.** The task-based fMRI data were preprocessed by using fMRIPrep 1.4.1, including slice-timing correction, head motion correction, and spatial normalization. The first 3 volumes for the task were discarded before preprocessing to allow for T1 equilibrium. For more details, please see Sheng et al. (4).

**Single-item response estimation.** Before estimating the single trial activation pattern, fMRI data were filtered in the temporal domain using a nonlinear high-pass filter with a 100-s cutoff. A generalized linear model (GLM) was performed to estimate the activation pattern for each repetition of the face-name pairs during encoding. A least square separate method was used in this single-trial model. Here, the target trial served as the model of one explanatory variable (EV), and all other trials were modeled as another EV (57). Each trial was modeled at its presentation time and convolved with a canonical hemodynamic response function (double gamma). This voxelwise GLM was used to compute the activation associated with each of the 60 trials in the task. The single-item response estimation was conducted in the native space, which was then transformed to the MNI152 space using the *antsApplyTransforms* tool from the Advanced Normalization Tools (58). The t-map in standard space was used for subsequent representational similarity analysis (59).

**Definition of the ROI.** We focused our analyses on the VVC, PMC, HIP, ANG, IFG, and vmPFC (Fig. 2B). These regions have been consistently shown to be involved in human episodic memory (11, 25, 34, 35, 48). For example, several studies based on univariate and multivariate analyses have found significant subsequent memory effects in the VVC, HIP, ANG, IFG, and vmPFC (16, 32, 33, 35). Based on the intersubject analysis, PMC was consistently shown to be associated with episodic memory (23, 25, 29). The VVC, HIP, and IFG were defined using the Harvard-Oxford probabilistic atlas with a threshold of 25% probability (25). The VVC consists of the ventral lateral occipital cortex, occipital fusiform, occipital temporal fusiform, and paraHIP (25). Following previous studies of ISC (23, 24, 43), we defined the PMC based on functional parcellations using resting-state functional connectivity (60). Specifically, it contains the posterior medial cluster of the dorsal DMN ([http://findlab.stanford.edu/functional\\_ROIs.html](http://findlab.stanford.edu/functional_ROIs.html)). The vmPFC was defined by NeuroSynth (<https://identifiers.org/neurovault.image:460116>), a large-scale database for meta-analysis (61). The meta-analysis of nearly 10,000 fMRI studies identified several functionally separable subregions in the medial frontal cortex and showed that vmPFC is involved in episodic memory and coactivated with the amygdala and HIP (61). The ANG was defined using the Schaefer2018 atlas (400 parcels) (62) and contained all parietal nodes within the DMN (Network

15 to 17) (8, 35, 49). Except the PMC and vmPFC, all the other ROIs were separately defined for the left and right hemispheres.

**Construction of the gISPS Matrix.** First, we constructed an RSM for each participant by calculating cross-trial pattern similarity using Pearson correlation coefficients. Second, we computed the between-subject correlation of Fisher's Z-transformed RSM in each ROI (i.e., second-order similarity) to obtain the gISPS matrix (25) (Fig. 4A).

**Construction of ISC Matrix.** We first averaged activation values for each trial and each ROI to obtain TTS and then constructed an ISC matrix by computing the between-subject correlation of the TTS in each ROI (Fig. 4B). We used TTS to calculate ISC because the participants needed to perform an orientation judgment task in a self-paced way after encoding a face-name pair in our experimental paradigm. Therefore, the original time points were not strictly time-locked for different participants. Given that all the participants learned face-name pairs with the same learning sequence, we computed ISC based on TTS.

**Mantel Test.** We used the Mantel test (63) implemented in the *vegan* package (56) (<https://CRAN.R-project.org/package=vegan>) to examine the relationship between cross-participant neural similarity and their SMC. The Mantel statistic is simply a Pearson correlation between entries of two correlation or dissimilarity matrices. However, the significance cannot be directly assessed because there are  $N(N-1)/2$  entries for just  $N$  observations. Therefore, the  $P$ -value is based on the permutation test. Specifically, the first distance or similarity matrix (*xmat*) was permuted, and the second (*y**mat*) was kept constant. Then, the permutation  $P$ -value of the Mantel test was computed as  $(N+1)/(n+1)$ , where  $N$  is the number of randomized correlation coefficients greater than the actual observation, and  $n$  is the number of permutations, which was set to 9,999 in the current study. Similarly, to exclude the effects of potentially confounding variables (i.e., memory performance), the *mantel.partial* function was also used to examine the significance of partial correlations.

**Data, Materials, and Software Availability.** All data needed to evaluate the conclusions in the paper are presented in the paper and/or *SI Appendix*. Processed anonymized data and custom code have been deposited in GitHub (<https://github.com/Cynthia1229/Individualized-memory>) (64).

**ACKNOWLEDGMENTS.** G.X. received grants from the National Natural Science Foundation of China (31730038), the Sino-German Collaborative Research Project "Crossmodal Learning" (NSFC 62061136001/DFG TRR169), the China-Israel collaborative research grant (NSFC 31861143040).

Author affiliations: <sup>a</sup>State Key Laboratory of Cognitive Neuroscience and Learning and International Data Group/McGovern Institute for Brain Research, Beijing Normal University, Beijing 100875, China; <sup>b</sup>Department of Psychological Science, University of California, Irvine, CA 92697; and <sup>c</sup>Chinese Institute for Brain Research, Beijing 102206, China

1. D. A. Bors, C. M. MacLeod, "Individual differences in memory" in *Memory* (Elsevier, 1996), pp. 411–441.
2. M. J. Chadwick et al., Semantic representations in the temporal pole predict false memories. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 10180–10185 (2016).
3. B. A. Kirchhoff, R. L. Buckner, Functional-anatomic correlates of individual differences in memory. *Neuron* **51**, 263–274 (2006).
4. J. Sheng et al., Higher-dimensional neural representations predict better episodic memory. *Sci. Adv.* **8**, eabm3829 (2022).
5. B. Zhu et al., Individual differences in false memory from misinformation: Cognitive factors. *Memory* **18**, 543–555 (2010).
6. B. Zhu et al., Multiple interactive memory representations underlie the induction of false memory. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 3466–3475 (2019).
7. K. H. Alm et al., Medial temporal lobe white matter pathway variability is associated with individual differences in episodic memory in cognitively normal older adults. *Neurobiol. Aging* **87**, 78–88 (2020).
8. A. N. Trelle et al., Hippocampal and cortical mechanisms at retrieval explain variability in episodic remembering in older adults. *ELife* **9**, e55335 (2020).
9. G. S. Wig et al., Medial temporal lobe BOLD activity at rest predicts individual differences in memory ability in healthy young adults. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 18555–18560 (2008).
10. A. D. Wagner et al., Building memories: Remembering and forgetting of verbal experiences as predicted by brain activity. *Science* **281**, 1188–1191 (1998).
11. G. Xue et al., Greater neural pattern similarity across repetitions is associated with better memory. *Science* **330**, 97–101 (2010).
12. M. T. deBettencourt, K. A. Norman, N. B. Turk-Browne, Forgetting from lapses of sustained attention. *Psychon. Bull. Rev.* **25**, 605–611 (2018).
13. K. P. Madore et al., Memory failure predicted by attention lapsing and media multitasking. *Nature* **587**, 87–91 (2020).
14. Y. Lu, C. Wang, C. Chen, G. Xue, Spatiotemporal neural pattern similarity supports episodic memory. *Curr. Biol.* **25**, 780–785 (2015).
15. G. Xue et al., Complementary role of frontoparietal activity and cortical pattern similarity in successful episodic memory encoding. *Cereb. Cortex* **23**, 1562–1571 (2013).
16. L. Zheng et al., Reduced fidelity of neural representation underlies episodic memory decline in normal aging. *Cereb. Cortex* **28**, 2283–2296 (2018).
17. S. E. Favila, H. Lee, B. A. Kuhl, Transforming the concept of memory reactivation. *Trends Neurosci.* **43**, 939–950 (2020).
18. M. Moscovitch, A. Gilboa, Systems consolidation, transformation and reorganization: Multiple trace theory, trace transformation theory and their competitors. *PsyArXiv [Preprint]* (2021). <https://doi.org/10.31234/osf.io/xybrs>. (8 November 2022).
19. G. Xue, From remembering to reconstruction: The transformative neural representation of episodic memory. *Prog. Neurobiol.* **219**, 102351 (2022).
20. J. Liu et al., Transformative neural representations support long-term episodic memory. *Sci. Adv.* **7**, eabg9715 (2021).
21. G. Xue, The neural representations underlying human episodic memory. *Trends Cogn. Sci.* **22**, 544–561 (2018).



22. U. Hasson, Y. Nir, I. Levy, G. Fuhrmann, R. Malach, Intersubject synchronization of cortical activity during natural vision. *Science* **303**, 1634–1640 (2004).
23. J. Chen *et al.*, Shared memories reveal shared structure in neural activity across individuals. *Nat. Neurosci.* **20**, 115–125 (2017).
24. G. E. Koch, J. P. Paulus, M. N. Coutanche, Neural patterns are more similar across individuals during successful memory encoding than during failed memory encoding. *Cereb. Cortex* **30**, 3872–3883 (2020).
25. X. Xiao *et al.*, Individual-specific and shared representations during episodic memory encoding and retrieval. *NeuroImage* **217**, 116909 (2020).
26. J. P. Dmochowski *et al.*, Audience preferences are predicted by temporal reliability of neural processing. *Nat. Commun.* **5**, 4567 (2014).
27. J. J. Ki, S. P. Kelly, L. C. Parra, Attention strongly modulates reliability of neural responses to naturalistic narrative stimuli. *J. Neurosci.* **36**, 3092–3101 (2016).
28. H. Song, E. S. Finn, M. D. Rosenberg, Neural signatures of attentional engagement during narratives and its consequences for event memory. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2021905118 (2021).
29. U. Hasson, O. Furman, D. Clark, Y. Dudai, L. Davachi, Enhanced intersubject correlations during movie viewing correlate with successful episodic encoding. *Neuron* **57**, 452–462 (2008).
30. S. S. Cohen, L. C. Parra, Memorable audiovisual narratives synchronize sensory and supramodal neural responses. *eNeuro* **3**, ENEURO.0203-16.2016 (2016).
31. E. Simony *et al.*, Dynamic reconfiguration of the default mode network during narrative comprehension. *Nat. Commun.* **7**, 12141 (2016).
32. G. Brod, Y. L. Shing, Specifying the role of the ventromedial prefrontal cortex in memory formation. *Neuropsychologia* **111**, 8–15 (2018).
33. H. Kim, Neural activity that predicts subsequent memory and forgetting: A meta-analysis of 74 fMRI studies. *NeuroImage* **54**, 2446–2461 (2011).
34. B. A. Kuhl, M. M. Chun, Successful remembering elicits event-specific activity patterns in lateral parietal cortex. *J. Neurosci.* **34**, 8051–8060 (2014).
35. Z. Ye, L. Shi, A. Li, C. Chen, G. Xue, Retrieval practice facilitates memory updating by enhancing and differentiating medial prefrontal cortex representations. *ELife* **9**, e57023 (2020).
36. M. Xia, J. Wang, Y. He, BrainNet Viewer: A network visualization tool for human brain connectomics. *PLoS One* **8**, e68910 (2013).
37. L. J. Chang *et al.*, Endogenous variation in ventromedial prefrontal cortex state dynamics during naturalistic viewing reflects affective experience. *Sci. Adv.* **7**, eabf7129 (2021).
38. U. Hasson, R. Malach, D. J. Heeger, Reliability of cortical activity during natural stimulation. *Trend Cogn. Sci.* **14**, 40–48 (2010).
39. S. A. Nastase, V. Gazzola, U. Hasson, C. Keysers, Measuring shared responses across subjects using intersubject correlation. *Soc. Cogn. Affect. Neurosci.* **14**, 667–685 (2019).
40. L. Nummenmaa *et al.*, Emotions promote social interaction by synchronizing brain activity across individuals. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 9599–9604 (2012).
41. J. N. van der Meer, M. Breakspear, L. J. Chang, S. Sonkusare, L. Cocchi, Movie viewing elicits rich and reliable brain state dynamics. *Nat. Commun.* **11**, 5004 (2020).
42. O. Bein, N. Reggev, A. Maril, Prior knowledge promotes hippocampal separation but cortical assimilation in the left inferior frontal gyrus. *Nat. Commun.* **11**, 4590 (2020).
43. M. Nguyen, T. Vanderwal, U. Hasson, Shared understanding of narratives is correlated with shared neural responses. *NeuroImage* **184**, 161–170 (2019).
44. S. Saalasti *et al.*, Inferior parietal lobule and early visual areas support elicitation of individualized meanings during narrative listening. *Brain Behav.* **9**, e01288 (2019).
45. J. M. Lahnakoski *et al.*, Synchronous brain activity across individuals underlies shared psychological perspectives. *NeuroImage* **100**, 316–324 (2014).
46. S. K. Jeong, Y. Xu, Behaviorally relevant abstract object identity representation in the human parietal cortex. *J. Neurosci.* **36**, 1607–1619 (2016).
47. Z. Ye *et al.*, Neural global pattern similarity underlies true and false memories. *J. Neurosci.* **36**, 6792–802 (2016).
48. X. Xiao *et al.*, Transformed neural pattern reinstatement during episodic memory retrieval. *J. Neurosci.* **37**, 2986–2998 (2017).
49. S. E. Favila, R. Samide, S. C. Sweigart, B. A. Kuhl, Parietal representations of stimulus features are amplified during memory retrieval and flexibly aligned with top-down goals. *J. Neurosci.* **38**, 7809–7821 (2018).
50. N. W. Schuck, M. B. Cai, R. C. Wilson, Y. Niv, Human orbitofrontal cortex represents a cognitive map of state space. *Neuron* **91**, 1402–1412 (2016).
51. A. Gilboa, H. Marlatt, Neurobiology of schemas and schema-mediated memory. *Trend Cogn. Sci.* **21**, 618–631 (2017).
52. E. S. Finn *et al.*, Idiosyncrony: From shared responses to individual differences during naturalistic neuroimaging. *NeuroImage* **215**, 116828 (2020).
53. M. P. Hejnar, K. A. Kiehl, V. D. Calhoun, Interparticipant correlations: A model free fMRI analysis technique. *Hum. Brain Mapping* **28**, 860–867 (2007).
54. J. Pajula, J.-P. Kauppi, J. Tohka, Inter-subject correlation in fMRI: Method validation against stimulus-model based analysis. *PLoS One* **7**, e41196 (2012).
55. J. D. Koen, Age-related neural dedifferentiation for individual stimuli: An across-participant pattern similarity analysis. *Aging Neuropsychol. Cogn.* **29**, 552–576 (2022).
56. J. Oksanen *et al.*, "Vegan: Community ecology package" (R package version 2.5-7, Github, 2020).
57. J. A. Mumford, B. O. Turner, F. G. Ashby, R. A. Poldrack, Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *NeuroImage* **59**, 2636–2643 (2012).
58. B. B. Avants *et al.*, The Insight ToolKit image registration framework. *Front. Neuroinform.* **8**, 44 (2014).
59. A. Walther *et al.*, Reliability of dissimilarity measures for multi-voxel pattern analysis. *NeuroImage* **137**, 188–200 (2016).
60. W. R. Shirer, S. Ryali, E. Rykhlevskaia, V. Menon, M. D. Greicius, Decoding subject-driven cognitive states with whole-brain connectivity patterns. *Cereb. Cortex* **22**, 158–165 (2012).
61. A. de la Vega, L. J. Chang, M. T. Banich, T. D. Wager, T. Yarkoni, Large-scale meta-analysis of human medial frontal cortex reveals tripartite functional organization. *J. Neurosci.* **36**, 6553–6562 (2016).
62. A. Schaefer *et al.*, Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cereb. Cortex* **28**, 3095–3114 (2018).
63. P. Legendre, L. Legendre, *Numerical Ecology* (Elsevier, ed. 3 English, 2012).
64. S. Jintao, Individualized memory (Version 1.0.0, Computer software). *GitHub*. <https://github.com/Cynthia1229/Individualized-memory>. Deposited 10 January 2023.