communications biology

A Nature Portfolio journal

Article

https://doi.org/10.1038/s42003-025-07964-6

Time-dependent consolidation mechanisms of durable memory in spaced learning

Yifeixue Yang[1], Ziyi Huang[1], Yun Yang[1], Mingxia Fan[2] & Dazhi Yin[1,3]

Emerging studies suggest that time-dependent consolidation enables memory stabilization by promoting memory integration and hippocampal-cortical transfer. Compared to massed learning, how time-dependent consolidation contributes to forming durable memory and what neural signatures predict durable memory in spaced learning remain unclear. We recruited 48 participants who underwent either 3-day spaced learning or 1-day massed learning, and both resting-state and task-based fMRI data were collected in multiple delayed tests (i.e., immediate, 1-week, and 1-month). We use representational similarity analysis to assess neural integration and replay in the hippocampus and default mode network (DMN) subsystems. In contrast with massed learning, spaced learning induces higher neural pattern similarity during immediate retrieval only in DMN subsystems. Particularly, the neural pattern similarity in the dorsal-medial DMN (DMN_{dm}) and medial-temporal DMN subsystems predicts the durable memory defined by 1-month delay. Moreover, we find increased neural replay of durable memory in the DMN_{dm} for spaced learning and in the hippocampus for both spaced and massed learning. Our findings suggest that time-dependent consolidation promotes neural integration and replay in the cortex rather than in the hippocampus, which may underlie the formation of durable memory after spaced learning.

Based on the spacing effect, a well-known phenomenon in learning and memory, spaced learning can enhance memory performance, especially for memory durability, compared to massed learning[1–5]. Previous studies suggest that spaced learning leads to robust memory by eliciting retrieval effort and the reactivation of a preceding memory trace[6], manifests as stronger brain activation[2,7] and enhanced item-specific spatiotemporal pattern similarity[8] during the encoding phase. However, these studies used trial-based spaced learning and mainly focused on the neural mechanisms of encoding, which might ignore the effect of time-dependent consolidation that is more profound in day-based spaced learning. According to a behavioral study, superior memory can last months after day-based spaced learning compared to massed learning[9]. Nevertheless, neuroimaging studies with day-based spaced learning typically assess memory retention following relatively short delays, such as several days[10,11], and they rarely explored the time-dependent consolidation mechanisms underlying the formation of durable memory in spaced learning. Thus, questions are raised about which characteristic neural representations contribute to time-dependent consolidation in spaced learning and whether they can predict durable memory.

The consolidation theory posits that consolidation fosters the formation of long-term memory traces through effective stabilization and reinforcement[12–16]. On the one hand, time-dependent consolidation may lead to a transition from detailed memory to gist-like memory, which integrates memories to promote efficient storage[17–21]. Accordingly, previous functional magnetic resonance imaging (fMRI) studies have demonstrated neural integration as a higher retrieval-related neural pattern similarity between different items after a day[22] or a week[23] than immediately after learning. Although time-dependent consolidation may facilitate the integration of memories, whether spaced learning benefits from this mechanism remains elusive. On the other hand, convergent studies have suggested that memories (especially episodic memories) are initially encoded in the hippocampus and then transferred to distributed cortical networks for stabilization through hippocampal-cortical interactions during consolidation[19,24–28]. Nevertheless, whether spaced learning is conducive to memory stabilization by promoting hippocampal-cortical transfer is largely unknown.

During post-encoding consolidation, spontaneous activities within hippocampal and across hippocampal-cortical networks are the leading

[1]Shanghai Key Laboratory of Brain Functional Genomics (Ministry of Education), Affiliated Mental Health Center (ECNU), School of Psychology and Cognitive Science, East China Normal University, Shanghai, China. [2]Shanghai Key Laboratory of Magnetic Resonance, School of Physics and Electronic Science, East China Normal University, Shanghai, China. [3]Shanghai Changning Mental Health Center, Shanghai, China. e-mail: dzyin@psy.ecnu.edu.cn

mechanism underlying memory stabilization and transformation[20,29–33]. Consolidation relies on the spontaneous memory-related replay after learning[31,32,34,35]. Several fMRI studies evaluated the appearance of task-evoked memory pattern, as well as increased hippocampal-cortical functional connectivity (FC) in post-encoding rest, and revealed a positive relationship between spontaneous replay and subsequent memory[36–39]. For instance, the prior study indicated that after a 1-day delay, the items replayed more during post-encoding consolidation showed better subsequent memory[36]. Besides the memory replay within the hippocampus, previous fMRI studies have suggested that changes in hippocampus-cortex FC after learning reflect the transfer of memory storage[10,40,41]. However, whether lags in spaced learning enable more spontaneous replay and thus contribute to the formation of durable memory remains poorly understood.

Moreover, the emerging model emphasizes that the default mode network (DMN), which has frequently been linked to episodic memory, may serve as a hub for igniting replay cascades and support the reactivation of older memories[42–47]. Due to its complex cognitive engagement, the DMN has been divided into distinct subsystems: the dorsal-medial DMN (DMN_{dm}), core DMN (DMN_{core}), and medial-temporal DMN (DMN_{mt})[48]. Specifically, the DMN_{dm} is closely linked to the integration and storage of memory[49–52]. The DMN_{core}, involved in self-referential and long-term memory processing[53,54], serves as a functional hub and mediates information transfer between DMN_{dm} and the DMN_{mt} subsystems[55–58]. The DMN_{mt}, most close to the hippocampus spatially and functionally, is associated with the encoding and retrieval of episodic memory, particularly in gathering cognitive, emotional, environmental, and other information[54,59,60]. Accordingly, memories might be initially encoded and processed in the hippocampus and the DMN_{mt} and then transferred and integrated into the DMN_{core} and DMN_{dm} for long-term storage. However, the exact destinations of the DMN subsystems that support the hippocampal-cortical transfer of memory traces after spaced learning are unclear.

In this study, we adopted a between-subject and day-based design and totally recruited 69 participants who underwent either 3-day spaced learning or 1-day massed learning. Behavioral analysis involved all 69 participants, while fMRI analysis included 48 participants. A total of 60 picture-word pairs was repeatedly used in six blocks for both groups, ensuring that all participants learned each pair six times. Both resting-state and task-based fMRI data were collected in multiple delayed tests (i.e., immediate, 1-week, and 1-month). Baseline resting-state fMRI data were collected before learning (Fig. 1). Considering that the substantial difference between the two learning groups is the consolidation time before the first test, the primary analysis of this study was to examine the between-group differences in neural pattern similarity and spontaneous replay in the immediate test using representational similarity analysis (RSA), and to test their predictive effects on long-term memory retention. We hypothesized that extended consolidation time in spaced learning group could enhance neural pattern similarity and spontaneous memory replay as well as facilitate the transfer of memory storage, although the immediate memory performance might be comparable for both learning groups. Moreover, these characteristic neural representations may predict durable memory, especially in the DMN_{dm}.

Results

Memory performance differed between spaced and massed learning in the delayed but not in the immediate tests

To estimate memory performance in a single test, the *d-prime*, denoting *hit* rate corrected by *false alarms (FA)* rate, was calculated. To confirm that no between-group difference in memory performance in immediate memory test but do in delayed tests revealed by previous studies[10,61], we conducted independent sample *t*-test for each test. We found that *d-prime* was significantly higher in the spaced learning group than in the massed learning group in the 1-week ($t_{(67)} = 2.38$, $p = 0.020$) and 1-month ($t_{(67)} = 2.95$, $p = 0.004$) delayed tests. No significant difference in *d-prime* was observed in the test immediately after learning ($t_{(67)} = 0.11$, $p = 0.915$) (Fig. 2a).

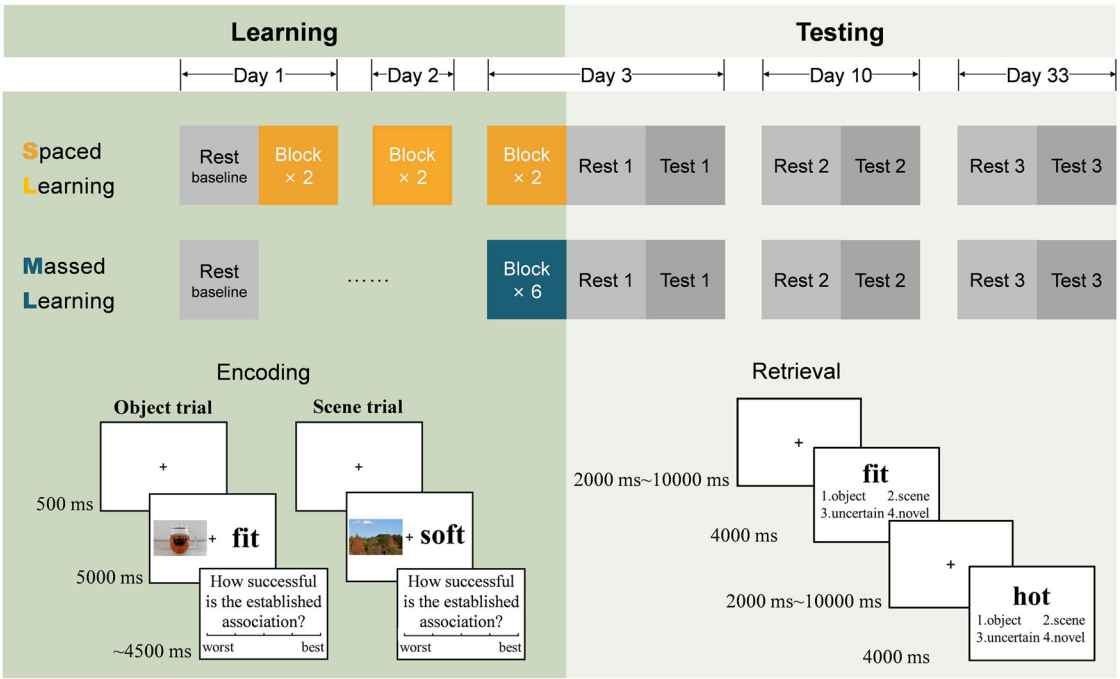

Fig. 1 | Experimental design. The experiment consisted of a memory learning phase and three test sessions (immediate, 1-week, and 1-month) after the learning phase. Participants were randomly assigned to either a 3-day spaced learning group (two blocks of learning each day) or a 1-day massed learning group (six blocks of learning in a single day) to learn a set of 60 picture-word pairs six times. The encoding phases were all conducted out of the scanner. Memory tests were conducted during fMRI scanning with 60 learned and 30 novel words in each test session. During tests, participants were required to indicate whether the presented words were learned and the type of the associated picture. Following only resting-state functional magnetic resonance imaging (fMRI) before learning at baseline, both resting-state and task-based fMRI were collected in the three test sessions after the learning phase.

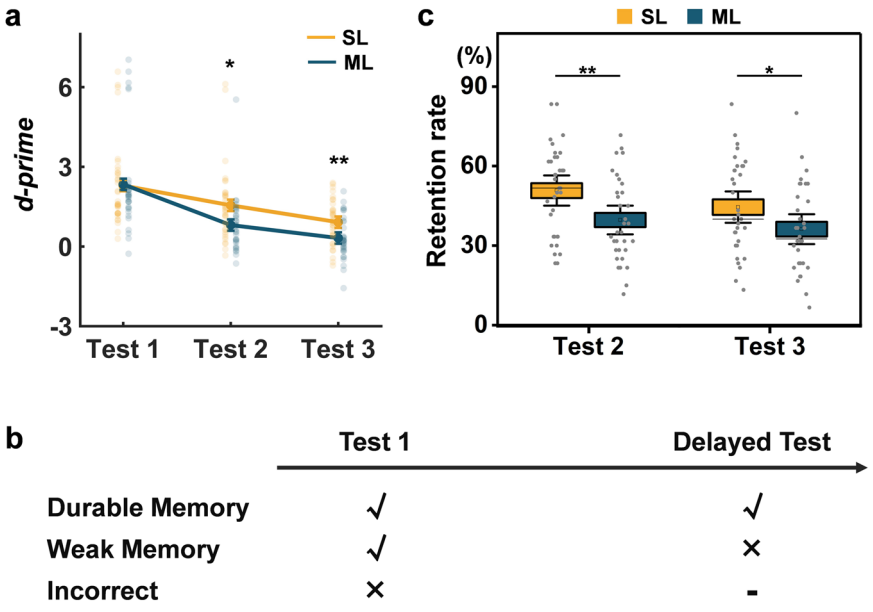

Fig. 2 | Behavioral performance and the definitions of durable and weak memories. a Changes in memory performance in each test in the spaced and massed learning groups. The memory performance differed between spaced and massed learning in the delayed but not in the immediate tests. **b** Durable memories refer to those successfully recalled in immediate and 1-month delayed tests. Weak memories refer to those only recalled in the immediate test. Stimuli recalled that immediate tests were considered incorrect. **c** The spaced learning group showed a higher retention rate in 1-week and 1-month delayed tests than massed learning group. The error bars denote standard errors; center line, median; box limits, standard errors; whiskers, 95% confidence interval; square, mean; SL spaced learning, ML massed learning; Test 1 immediate test; Test 2, 1-week delayed test; Test 3, 1-month delayed test; $^{**}p < 0.01$; and $^{*}p < 0.05$.

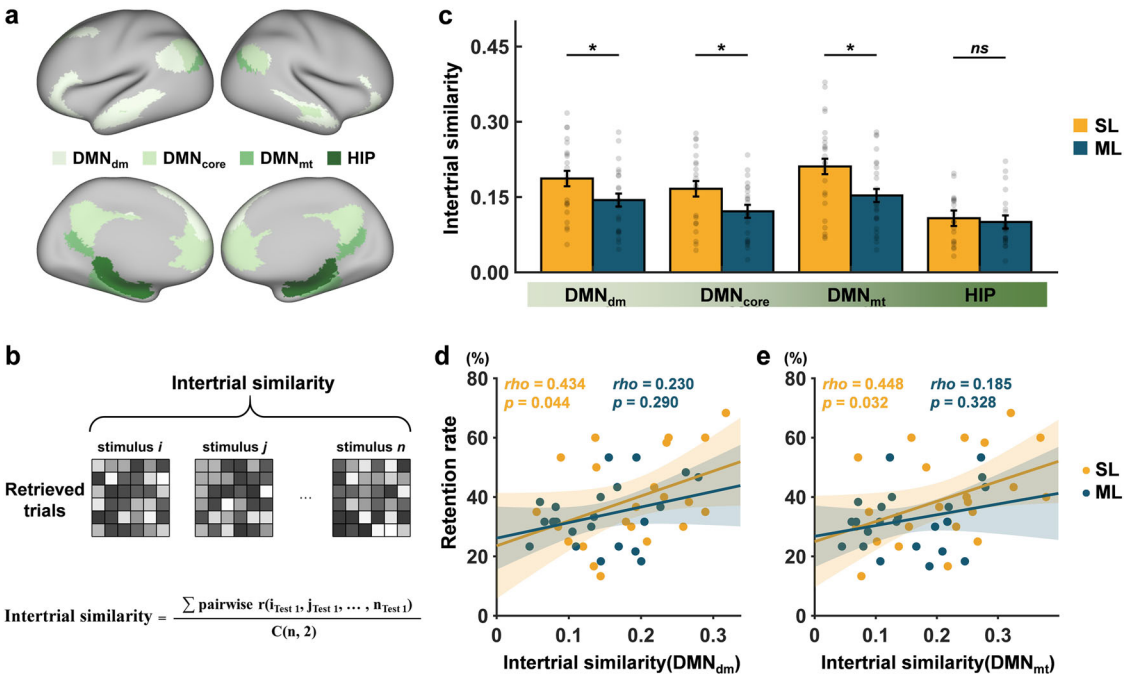

Fig. 3 | Intertrial similarity analysis. a The regions of interest (ROIs) for the hippocampus and the three DMN subsystems are shown in different colors. **b** Intertrial similarity analysis was performed by calculating the mean Pearson's correlation across the neural activity patterns of successfully retrieved memories in the immediate test. *i* and *j* denote the specific stimuli during retrieval; *n* indicates the number of successful memories; *C*(*n*,2) indicates the number of combinations when choosing 2 of *n* stimuli at a time. **c** The average intertrial similarity of each ROI for each learning group. **d** The prediction of the intertrial similarity of the DMN_{dm} to the 1-month retention rate for each learning group. **e** The prediction of the intertrial similarity of the DMN_{mt} to the 1-month retention rate for each learning group. The error bars denote standard errors; SL, spaced learning; ML, massed learning; DMN, default mode network; DMN_{dm}, dorsal-medial DMN; DMN_{core}, core DMN; DMN_{mt}, medial-temporal DMN; HIP, hippocampus; $^{*}p < 0.05$; and *ns*, non-significant.

Spaced learning leads to better memory retention for different delays

To evaluate memory retention, at least two tests after learning are required. In our main analysis, we defined durable memories as those successfully retrieved in both immediate and 1-month delayed tests. The retention rate was quantified as the percentage of durable memories among all successfully retrieved memories in the immediate test (Fig. 2b). To explore the delay effect, we also defined durable memories based on the 1-week delays. Through independent sample *t* test, we found that the retention rate after spaced learning was significantly higher than that after massed learning for both 1-week ($t_{(67)} = 2.87$, $p = 0.006$) and 1-month ($t_{(67)} = 2.06$, $p = 0.043$) delays (Fig. 2c). This result indicates that spaced learning leads to higher memory retention for both 1-week and 1-month delays, compared to massed learning.

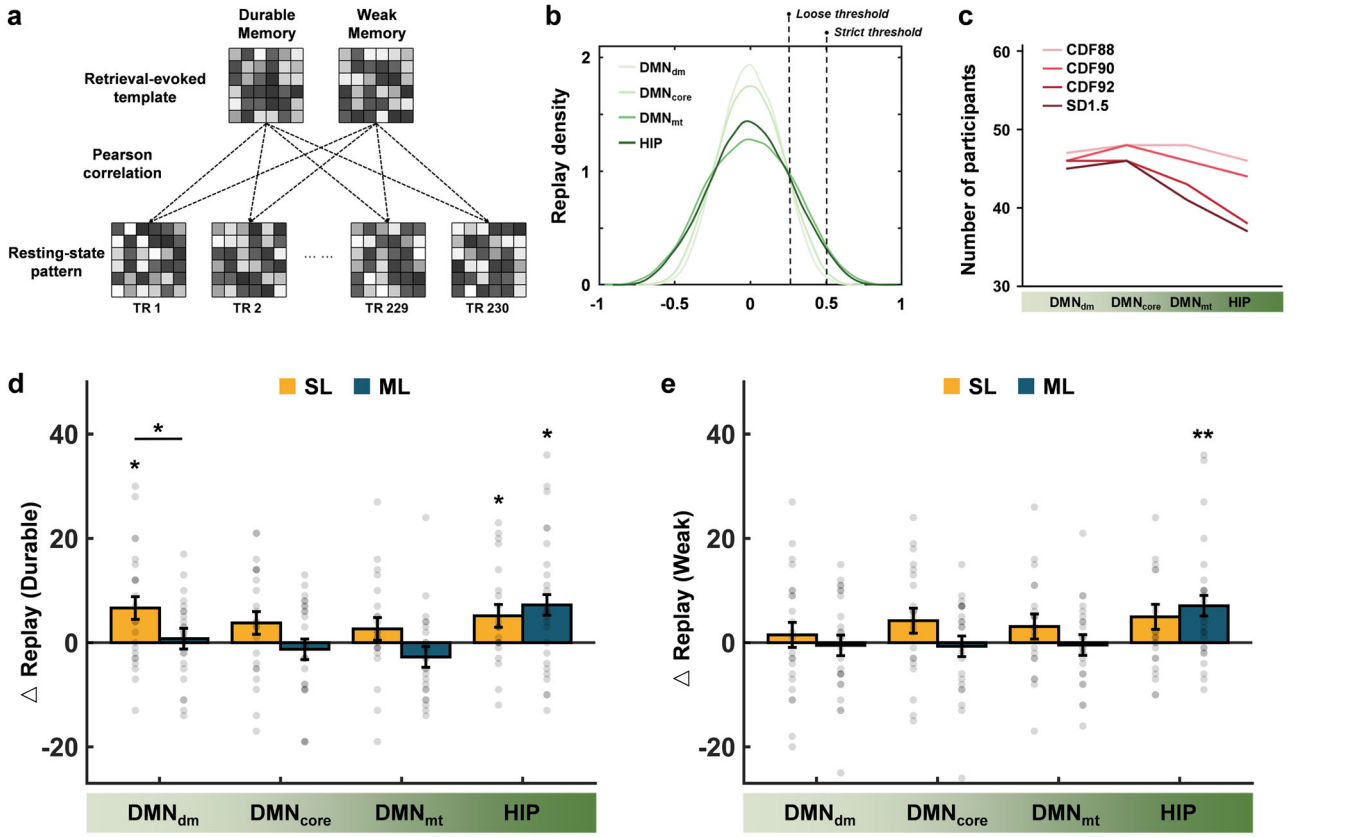

Fig. 4 | Changes of replay for durable and weak memories after learning. a For each participant, we created multivoxel neural activity templates for durable and weak memories based on the retrieval-related activity in the hippocampus and the three DMN subsystems. Correlation analyses were then performed between the retrieval-evoked neural activity templates and the framewise resting activity patterns. Replay was defined as a resting activity pattern that showed high similarity (above the 90th percentile) with the retrieval-evoked neural activity templates. **b** The distribution of correlations between retrieval-evoked neural activity templates and resting activity patterns in the hippocampus and the DMN subsystems. The 90th percentile, not an absolute correlation coefficient, was selected as the threshold to ensure consistency for different brain systems. For example, an absolute threshold of 0.5 might be appropriate for the DMN_{mt} and hippocampus, but it is too strict for the DMN_{dm} and DMN_{core}. **c** The number of effective participants with different thresholds. **d** Changes of replay for durable memories in the three DMN subsystems and hippocampus for each learning group. **e** Changes of replay for weak memories in the three DMN subsystems and hippocampus for each learning group. The error bars denote standard errors; DMN, default mode network; DMN_{dm}, dorsal-medial DMN; DMN_{core}, core DMN; DMN_{mt}, medial-temporal DMN; HIP, hippocampus; CDF88, the 88th percentile thresholds; CDF90, the 90th percentile thresholds; CDF92, the 92nd percentile thresholds; SD1.5, threshold of 1.5 standard deviations; △ replay, change of replay; SL, spaced learning; ML, massed learning; $^{**}p < 0.01$; and $^{*}p < 0.05$.

Spaced learning induced higher intertrial similarity of the DMN subsystems during immediate retrieval

Although no significant between-group differences in immediate memory performance were observed, we used RSA to examine the intertrial similarity of the hippocampus and the three DMN subsystems (Fig. 3a) during the immediate retrieval (Fig. 3b). Using the trial-by-trial general linear model (GLM), retrieval-evoked brain activity can be obtained for individual trials. The intertrial similarity was calculated by averaging the correlations between the brain activity patterns of the successfully retrieved trials. The independent sample *t*-test showed that the intertrial similarity of the DMN_{dm} ($t_{(43)} = 2.05$, $p = 0.046$), the DMN_{core} ($t_{(43)} = 2.30$, $p = 0.027$), and the DMN_{mt} ($t_{(44)} = 2.33$, $p = 0.024$) was significantly higher in the spaced learning group than in the massed learning group. No significant between-group differences in intertrial similarity were observed in the hippocampus ($t_{(42)} = 0.50$, $p = 0.617$) (Fig. 3c).

Intertrial similarity of DMN subsystems during immediate retrieval predicted memory retention after spaced learning

Through correlation analysis, we further found that the intertrial similarity of DMN_{dm} ($rho = 0.43$, $p = 0.044$) (Fig. 3d) and DMN_{mt} ($rho = 0.45$, $p = 0.032$) (Fig. 3e) was significantly correlated with the retention rate only for spaced learning. However, there was no significant difference in the correlations between the two groups for DMN_{dm} ($z = 0.75$, $p = 0.228$) and DMN_{mt} ($z = 0.96$, $p = 0.170$). The intertrial similarity was not correlated with *d-prime* in the immediate test in any learning group (Supplementary Fig. 1a, b). These findings indicate that spaced learning leads to a more similar neural activity pattern in cortical networks during immediate retrieval, which was associated with durable memory.

As an exploratory analysis, we have performed mixed analysis of variance (ANOVA) with a 2 (group: SL, ML) × 3 (test: immediate, 1-week, 1-month) for each region of interest (ROI) to examine how between-group differences in intertrial similarity change over time. We found significant main effect of group in the DMN_{core} ($F_{(1, 43)} = 4.54$, $p = 0.039$, $\eta^2_p = 0.10$) and DMN_{mt} ($F_{(1, 44)} = 5.45$, $p = 0.024$, $\eta^2_p = 0.11$) and main effect of test in the DMN_{dm} ($F_{(1, 43)} = 8.00$, $p = 0.001$, $\eta^2_p = 0.16$), the DMN_{core} ($F_{(1, 43)} = 3.81$, $p = 0.026$, $\eta^2_p = 0.08$), the DMN_{mt} ($F_{(1, 44)} = 7.81$, $p = 0.001$, $\eta^2_p = 0.15$) and the hippocampus ($F_{(1, 42)} = 13.68$, $p < 0.001$, $\eta^2_p = 0.25$). However, no significant interaction of group × test was observed in any ROI. Post-hoc analyses showed that intertrial similarity of the DMN_{core} and DMN_{mt} in the spaced learning group was higher than that in the massed learning group. The intertrial similarity in the immediate test was significantly higher than that in the 1-week test in the DMN_{core} ($p = 0.040$), the DMN_{mt} ($p = 0.006$), and the hippocampus ($p < 0.001$). The intertrial similarity in the immediate test was significantly higher than that in the 1-month test in the DMN_{dm}

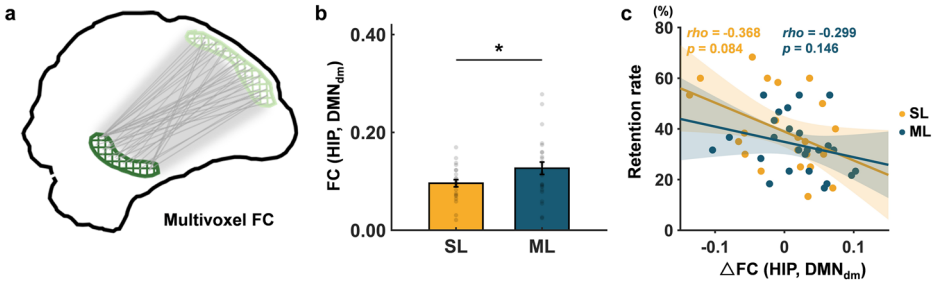

Fig. 5 | Multivoxel resting-state functional connectivity between the hippocampus and default mode network subsystems after spaced and massed learning. a Multivoxel FC was obtained by mean Pearson's correlations between the time course of each voxel in the hippocampus and that of each voxel in the DMN subsystems. **b** Differences in multivoxel FC between the spaced and massed learning groups in immediate rest. **c** The prediction of the change in FC of hippocampus and DMN_{dm} to the retention rate for each learning group. The error bars denote standard errors; SL spaced learning, ML massed learning, DMN default mode network, DMN_{dm} dorsal-medial DMN, HIP hippocampus, FC functional connectivity, △FC changes in FC between the immediate test after learning and baseline, and $^{*}p < 0.05$.

($p = 0.004$), the DMN_{mt} ($p = 0.008$,) and the hippocampus ($p = 0.001$) (Supplementary Fig. 2). These findings indicate that neural pattern similarity in the DMN subsystems is higher in the spaced learning group across the three delayed tests. The neural pattern similarity in the DMN subsystems and hippocampus reduces over time in both learning groups.

Temporal similarity in the DMN_{dm} was associated with memory retention after spaced learning

Besides the neural integration, we used RSA to examine the stimulus-specific temporal similarity of neural activity patterns between delayed tests for durable memory. Temporal similarity was calculated by averaging the correlations between the brain activity patterns for the same stimulus across different tests. No significant between-group differences were observed in mean temporal similarity across all delayed tests (Supplementary Fig. 3a). However, the temporal similarity between the immediate and 1-week delayed tests was significantly higher in the spaced learning group than in the massed learning group in DMN_{core} ($t_{(46)} = 2.10$, $p = 0.041$) and DMN_{mt} ($t_{(46)} = 2.08$, $p = 0.044$) (Supplementary Fig. 3b). Moreover, both the mean temporal similarity across all delayed tests ($rho = 0.46$, $p = 0.027$) (Supplementary Fig. 3c) and the temporal similarity between the immediate and 1-week delayed tests ($rho = 0.47$, $p = 0.023$) (Supplementary Fig. 3d) only in the DMN_{dm} were significantly correlated with the retention rate after spaced learning. However, there was no significant difference in the correlations between the two groups for both the mean temporal similarity across all delayed tests ($z = 0.79$, $p = 0.214$) and the temporal similarity between the immediate and 1-week delayed tests ($z = 0.66$, $p = 0.256$).

Spaced learning induced more neural replay in the DMN_{dm} for durable memories

To explore whether spaced learning induces more spontaneous replay for durable memories, we quantified the emergence of replay of durable and weak memories at rest. For each participant, we created a multivoxel neural activity template for durable memories based on the retrieval-related activity in the hippocampus and the three DMN subsystems separately, as well as for weak memories as a contrast. Correlation analyses were then performed between each retrieval-evoked neural activity template and the framewise resting activity patterns. Replay was defined as a resting activity pattern that showed high similarity (above the 90th percentile) with the retrieval-evoked neural activity template (Fig. 4a). Thus, the number of replay during resting state can be calculated both after learning and at baseline.

First, we examined whether the increment of replay was significantly higher than 0 with one-sample *t*-tests for each group within each ROI. Then, independent sample *t*-tests were conducted to examine the between-group difference in replay. The increment of replay for durable memory was significantly higher than 0 within the DMN_{dm} ($t_{(22)} = 2.81$, $p = 0.010$) and hippocampus ($t_{(21)} = 2.54$, $p = 0.019$) after spaced learning, while only within the hippocampus after massed learning ($t_{(24)} = 2.69$, $p = 0.013$). Through independent sample *t*-tests, we found that the increment of replay for durable memory within the DMN_{dm} after spaced learning was significantly higher than that after massed learning ($t_{(46)} = 2.10$, $p = 0.041$) (Fig. 4d). These findings indicate an increased replay in the DMN_{dm} and hippocampus after spaced learning, while the increase only occurs in the hippocampus after massed learning.

We also conducted similar analyses based on the neural activity templates of weak memories. The increment of replay was significantly higher than 0 only within the hippocampus after massed learning ($t_{(24)} = 2.95$, $p = 0.007$) (Fig. 4e). This result suggests that the increment of replay for weak memories only occurs in the hippocampus after massed learning.

To consider the separation of durable and weak memory templates, we have performed a permutation test to compare the similarity of the two templates to templates generated from shuffled trials. We found that the similarity of the two templates was not significantly greater or less than that generated from shuffled trials (Supplementary Fig. 4a, b). This result suggests that the two templates are not meaningfully different or similar. Through the multivariate pattern analysis (MVPA), we calculated the classification accuracy in distinguishing neural patterns for durable and weak memories in each ROI. The classification accuracy was significantly higher than chance in the DMN subsystems and hippocampus ($ps < 0.001$) (Supplementary Fig. 4c). These results indicate that the DMN subsystems and hippocampus are capable of differentiating durable and weak memories.

Weaker hippocampus-DMN_{dm} functional connectivity after spaced learning than after massed learning

Besides the increased replay of local neural activity, previous studies considered that the changes in hippocampus-cortex FC after learning reflect the transfer of memory storage[10,40,41]. Considering that prior analyses such as intertrial similarity and replay were multivariate, we, therefore, performed multivoxel FC rather than the correlation between the average time courses of the two regions. Specifically, we conducted multivoxel resting-state FC between the hippocampus and each DMN subsystem to assess hippocampal-cortical interactions (Fig. 5a). Through independent sample *t*-test, we found that the FC between the hippocampus and the DMN_{dm} in the spaced learning group was significantly lower than that in the massed learning group ($t_{(46)} = 2.05$, $p = 0.046$) in the immediate test (Fig. 5b). Moreover, the change in resting-state FC of the hippocampus with DMN_{dm} from the baseline to the immediate test after learning showed a marginally significant correlation with the retention rate in the spaced learning group ($rho = -0.37$, $p = 0.084$) (Fig. 5c). These findings suggest functional decoupling of the hippocampus and DMN_{dm} after spaced learning.

As an exploratory analysis, we calculated the increment of replay and FC during the 1-week and 1-month delayed tests to examine how spontaneous consolidation changes over time. For the replay, the increment of replay for durable memory was significantly higher than 0 within the

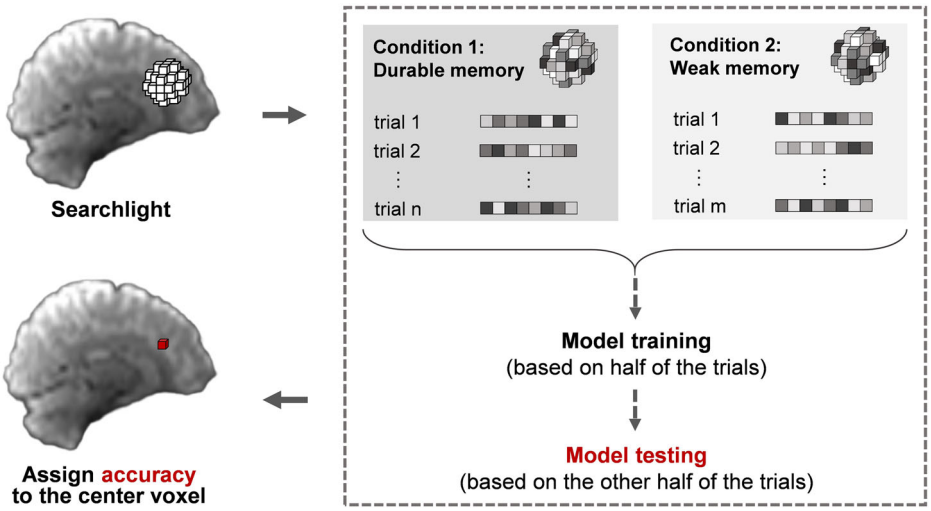

Fig. 6 | Framework of whole-brain searchlight multivariate pattern analysis. For each participant, the brain activity images of durable and weak memories were selected and evenly divided into training and testing datasets for machine learning. A spherical region of interest (ROI, radius = 3 voxels) was moved voxel by voxel through the brain. For each spherical ROI, a support vector machine with a specified kernel was utilized to train a classifier based on the training dataset, and the testing dataset was used to assess the classification accuracy, which was assigned to the center voxel of that ROI.

hippocampus ($t_{(22)} = 3.14$, $p = 0.005$) in the 1-week delayed test, and was significantly higher than 0 within the DMN_{core} ($t_{(22)} = 2.80$, $p = 0.033$) in the 1-month delayed test after spaced learning (Supplementary Fig. 5). However, through mixed ANOVA with a 2 (group: SL, ML) × 3 (test: immediate, 1-week, 1-month) design, no significant interaction or main effect was observed in any ROI for both the increment of replay and FC. These results indicate that cortical replay persists after delays in spaced learning, while replay in the hippocampus decays after 1 month. However, the massed learning group showed no significant replay within the hippocampus or cortical regions, even with more exposure after delays. Additionally, the between-group differences in FC between the hippocampus and DMN subsystems decay after delays.

Brain regions for distinguishing durable and weak memories differed in spaced and massed learning groups

In order to reveal the specific neural representation related to durable memory, a data-driven analysis framework was also conducted. First, we used whole-brain searchlight MVPA to identify brain regions that can distinguish durable and weak memories in the immediate test for both learning groups. Subsequently, we investigated the neural integration and spontaneous replay in these specific regions. Specifically, for each participant, brain activity images of durable and weak memories were divided equally into training and testing datasets. A spherical ROI (radius = 3 voxels) was moved voxel-by-voxel through the brain[62]. For each spherical ROI, a support vector machine with a specified kernel was utilized to train the classifier based on the training dataset. Classification accuracy was assessed based on the testing dataset and assigned to the center voxel of the spherical ROI (Fig. 6).

The brain regions capable of distinguishing between durable and weak memories were obtained through a group-level one-sample *t* test and were finally overlaid on the masks of the hippocampus and the three DMN subsystems (Fig. 7a). We observed that brain regions distinguishing durable and weak memories were mainly located in DMN_{dm} and anterior DMN_{core} for the spaced learning group. In contrast, the identified brain regions for massed learning were mainly located in the hippocampus, DMN_{mt}, and posterior DMN_{core}. These findings suggest a dissociation of the hippocampus and cortex in distinguishing durable and weak memories after spaced and massed learning.

Furthermore, we calculated the intertrial similarity, temporal similarity, and replay number based on the masks resulting from MVPA for each learning group. Intertrial similarity ($rho = 0.46$, $p = 0.027$) (Fig. 7b), temporal similarity ($rho = 0.44$, $p = 0.036$) (Fig. 7c), and the replay number for durable memories ($rho = 0.44$, $p = 0.038$) (Fig. 7d) were significantly correlated with the retention rate only in the spaced learning group. Moreover, the differences in the correlations between the two groups were significant for the intertrial similarity ($z = 1.95$, $p = 0.025$) and replay ($z = 1.65$, $p = 0.050$) but were not for the temporal similarity ($z = 1.53$, $p = 0.062$). These findings suggest that neural activity patterns in the brain regions that enable the distinction between durable and weak memories can also predict memory retention for spaced learning.

Control analysis

To estimate whether greater intertrial similarity in the spaced learning group over the massed learning group is memory-specific, we have conducted the same analyses for all incorrect trials, as well as separately for object and scene trials. For incorrect trials, we conducted independent sample *t*-tests and found that the intertrial similarity was significantly higher in the spaced learning group than in the massed learning group in the DMN_{dm} ($t_{(45)} = 2.56$, $p = 0.014$), the DMN_{core} ($t_{(45)} = 2.99$, $p = 0.005$) and the DMN_{mt} ($t_{(45)} = 2.36$, $p = 0.023$) (Supplementary Fig. 6). For object and scene trials, we performed a mixed ANOVA with a 2 (group: SL, ML) × 2 (picture type: object, scene) design for each ROI. We found a significant main effect of picture type in the DMN_{dm} ($F_{(1, 45)} = 6.53$, $p = 0.014$, $\eta^2_p = 0.13$). Post-hoc analysis showed higher intertrial similarity for the objects. In addition, we found a significant main effect of group in the DMN_{dm} ($F_{(1, 45)} = 4.69$, $p = 0.036$, $\eta^2_p = 0.09$), the DMN_{core} ($F_{(1, 45)} = 6.63$, $p = 0.013$, $\eta^2_p = 0.13$) and the DMN_{mt} ($F_{(1, 45)} = 5.53$, $p = 0.023$, $\eta^2_p = 0.11$). Post-hoc analyses showed higher intertrial similarity for the spaced learning group. However, no significant interaction of group × picture type was observed in any ROI (Supplementary Fig. 7). These findings suggest that greater intertrial similarity in the spaced learning group over the massed learning group is not memory-specific.

Validation analysis

To estimate the robustness of the replay analysis to different thresholds, we also used the 88th (CDF88) and 92nd (CDF92) percentile thresholds, as well as a threshold of 1.5 standard deviations (SD1.5) above the mean. For durable memory, we consistently found that the increment of replay within the DMN_{dm} in the spaced learning group was significantly higher than 0 with the CDF88 ($t_{(22)} = 2.54$, $p = 0.019$) and CDF92 ($t_{(22)} = 2.21$, $p = 0.038$) thresholds and were marginally significant with the SD1.5 threshold ($t_{(22)} = 1.89$, $p = 0.072$). The increment of replay within the hippocampus in spaced learning group was significantly higher than 0 with the CDF88 threshold ($t_{(22)} = 2.97$, $p = 0.007$) and was marginally significant with the SD1.5 threshold ($t_{(21)} = 1.84$, $p = 0.079$). The increment of replay within the hippocampus in the massed learning group was significantly higher than 0 with the CDF88 ($t_{(22)} = 2.59$, $p = 0.016$), CDF92 ($t_{(22)} = 3.13$, $p = 0.005$) and SD1.5 ($t_{(21)} = 2.89$, $p = 0.008$) thresholds. For weak memory, we found that

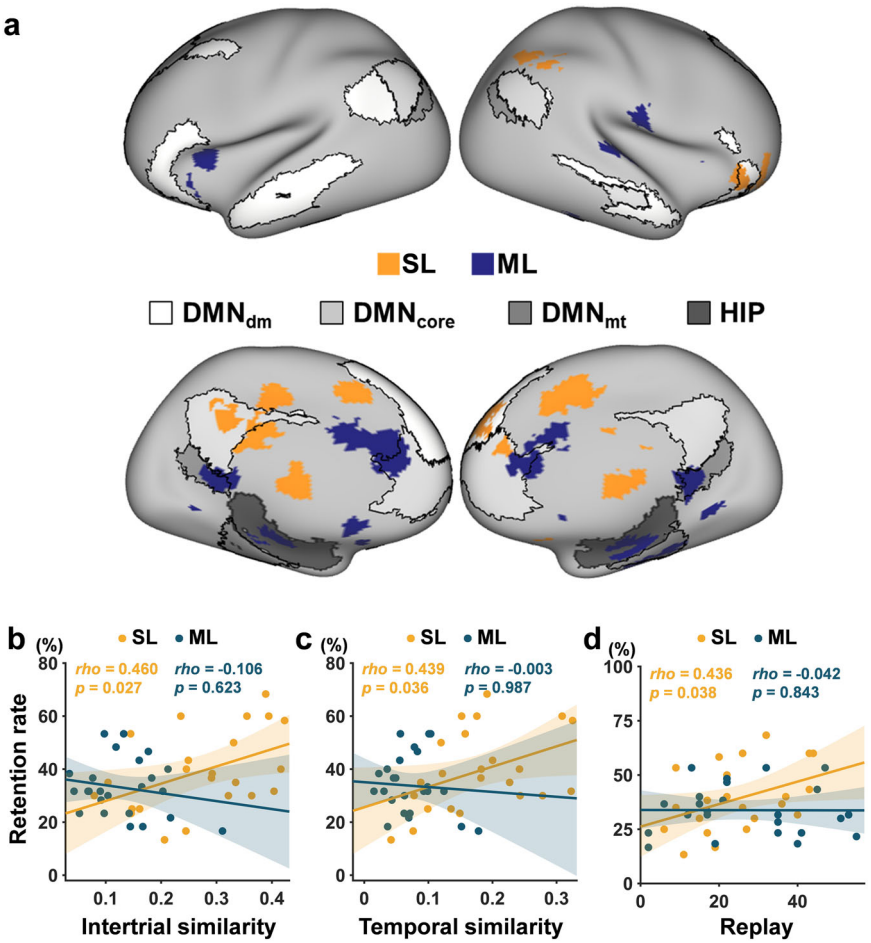

Fig. 7 | Brain regions capable of distinguishing durable and weak memories as identified by whole-brain searchlight multivariate pattern analysis. a The brain regions identified by multivariate pattern analysis (MVPA) for both spaced (yellow) and massed (blue) learning groups were overlaid on the masks of the hippocampus and the DMN subsystems. **b** The prediction of the intertrial similarity in the masks resulting from MVPA to the retention rate for each learning group. **c** The correlation between the temporal similarity in the masks resulting from MVPA and the retention rate for each learning group. **d** The prediction of the number of replays in the masks resulting from MVPA to the retention rate for each learning group. SL spaced learning, ML massed learning, DMN default mode network, DMN_{dm} dorsal-medial DMN, DMN_{core} core DMN, DMN_{mt} medial-temporal DMN, and HIP hippocampus.

the increment of replay within the DMN_{dm} in the spaced learning group was not significant with any thresholds. The increment of replay within the hippocampus was significantly higher than 0 with the CDF88 ($t_{(22)} = 2.85$, $p = 0.009$), CDF92 ($t_{(22)} = 2.83$, $p = 0.009$), and SD1.5 ($t_{(21)} = 2.59$, $p = 0.016$) thresholds only in massed learning group (Supplementary Fig. 8). These results suggest that our findings are robust to different thresholds. However, the effect is weaker with the threshold of SD1.5.

It should be noted that the threshold of SD1.5 obtained by the pooled correlation coefficients from pre-encoding and post-encoding rest sessions of all participants is likely stricter than that obtained for a single participant in previous studies[36,63]. The stricter thresholds will lead to a reduction in the number of effective participants, because the number of replay might be 0 for some participants (Fig. 4c).

Additionally, we calculated univariate FC using a correlation between the average time courses between the hippocampus and the DMN subsystems. Through independent sample *t*-test, we consistently found that FC between the hippocampus and the DMN_{dm} in the spaced learning group was significantly lower than that in the massed learning group ($t_{(46)} = 2.11$, $p = 0.040$) in the immediate test after learning (Supplementary Fig. 9a). Moreover, the changes in FC before and after encoding showed a significant correlation with the retention rate in the massed learning group ($rho = -0.45$, $p = 0.023$) (Supplementary Fig. 9b). These results suggest that our findings are robust to different estimates of FC.

To examine the delay effect, we also defined durable and weak memories based on a 1-week delayed test, and the main analysis was redone. For intertrial similarity, we performed a correlation analysis between intertrial similarity and retention rate defined by the 1-week delay. We found that the intertrial similarity in DMN_{dm} ($rho = 0.38$, $p = 0.085$) and DMN_{mt} ($rho = 0.36$, $p = 0.096$) showed marginally significant correlations with the retention rate for spaced learning, whereas no significant correlations were observed for massed learning (Supplementary Fig. 1c, d). This result indicates that associations between the intertrial similarity in cortical networks and retention rate were weak for a relatively short delay.

For replay analyses, we consistently found that, for durable memories, the increment of replay was significantly higher than 0 within the DMN_{dm} ($t_{(22)} = 2.81$, $p = 0.010$) and hippocampus ($t_{(21)} = 2.54$, $p = 0.019$) after spaced learning, while only within the hippocampus after massed learning ($t_{(24)} = 2.69$, $p = 0.013$). However, no between-group differences were observed (Supplementary Fig. 10a). For weak memories, the increment of replay was significantly higher than 0 only within the hippocampus after massed learning ($t_{(24)} = 2.24$, $p = 0.035$) (Supplementary Fig. 10b). This result indicates that the increased replay in the DMN_{dm} and hippocampus for durable memory after spaced learning also occurs with a short delay, although it was more pronounced for a long delay.

We also conducted a whole-brain searchlight MVPA for durable and weak memories defined by the 1-week delay. We found that the brain regions capable of distinguishing between durable and weak memories for massed learning largely overlapped with those identified based on the 1-month delay. However, no regions were noticed in the spaced learning group based on a relatively short delay (Supplementary Fig. 11). This result suggests that the use of a relatively short delay may classify some weak memories into durable memories, which is not conducive to specific neural representations for durable memory.

Discussion

Despite much progress has been made in understanding the spacing effect, little is known about the time-dependent consolidation mechanisms in spaced learning as well as the predictions of durable memory, at least

partially due to trial-based spaced learning often shown in previous studies. This study investigated the spacing effect using a multi-day learning and a between-subject design. Moreover, memory retention was evaluated using multiple delayed tests at different time scales, and both retrieval-evoked and spontaneous brain activities were adopted to reveal the neural representations related to durable memory.

Behaviorally, we found significant between-group differences in memory performance at delayed tests and retention rate during delay, whereas no significant differences were observed in the immediate test. Previous studies on multi-day learning have consistently reported no significant differences in memory performance between two learning groups at immediate test[10,61,64,65]. By contrast, studies using trial-based intervals have shown a substantial spacing effect minutes after learning[2,8], but the effect on long-term memory retention, such as lasting for months, was rarely reported. This is because the persistence of the spacing effect may rely on learning intervals, especially sleep[66,67]. Previous studies with multi-day learning designs demonstrated that the spacing effect can persist for one year[9,68]. Therefore, adopting multi-day learning helps to study the long-term memory retention of spaced learning as well as its neural mechanisms.

Although no significant between-group differences in the number of successfully retrieved memories were observed in the immediate test, we found that spaced learning formed more similar neural activity patterns across memories in the DMN subsystems than massed learning. This suggests that time-dependent consolidation in spaced learning facilitates the neural integration of memories[17,22,23,69]. Consistently, previous studies have indicated the higher retrieval-related neural pattern similarity after consolidation, which is related to the fast spindle density that is thought to play an important role in systems memory consolidation[21]. In contrast, a recent study observed lower neural pattern similarity in the medial prefrontal cortex but not in the hippocampus during 1-week delayed retrieval in the spaced learning group[11]. This disparity of differences in neural activity similarity between spaced and massed learning is likely due to different lags between learning and different brain regions. Despite this disparity, we consistently observed dissociable differences in neural activities of the cortex and hippocampus between spaced and massed learning. More importantly, our findings further indicated that the neural pattern similarity of cortical networks predicted durable memory for spaced learning. This suggests that neural integration in the immediate test plays a crucial role in supporting durable memory after spaced learning.

Beyond retrieval-related neural activity, we observed an increment of replay for durable memories in the hippocampus in both learning groups. More interestingly, we observed an increased replay of the DMN_{dm} exclusively after spaced learning, which may have contributed to the formation of durable memories. In addition, we observed significant replay for weak memories only in the hippocampus after massed learning. Consistently, a previous study indicated no significant difference between the replay for remembered and forgotten stimuli in the hippocampus, while significantly higher replay for remembered stimuli than that for forgotten stimuli in the entorhinal cortex was observed[39]. It appears that the hippocampus equally replays the experienced information, whether it is subsequently remembered or not. However, when evaluating the proportion of correct features of remembered items, the hippocampus replayed the items more with a lower proportion of correct features immediately after learning, while replayed the items more with a higher proportion of correct features a day after learning[36]. Thus, although the hippocampus replays both high and low-precise memories, there is a prioritization for replaying lower-precise memories. Furthermore, a recent study revealed a higher co-reactivation of the hippocampus and cortex immediately after learning about forgotten stimuli than for correct stimuli, which also indicated the preference for weaker memories[63]. As memory information is initially processed in the hippocampus and then transferred to the cortex, the hippocampus plays a fundamental role in processing experienced information equally. Memories that cannot be well transferred to the cortex seem to be strongly processed in the hippocampus, however, remaining weak and subsequently forgotten. Thus, the cortical replay that only occurs after spaced learning might be attributed to the system consolidation and underlie the mechanisms of spacing effect.

According to the consolidation theory, the formation of long-term memory traces is attributed to effective stabilization and reinforcement[14]. Previous studies have indicated a transfer from the hippocampus to a distributed cortical network through memory consolidation to achieve long-term memory storage[19,24–28]. Besides the intertrial similarity and replay in the DMN_{dm} during immediate retrieval, we also observed that the temporal similarity of durable memories in this subnetwork is associated with memory retention in the spaced learning group. Our findings indicate that the cortex, especially the DMN_{dm}, supports durable memory after spaced learning. Moreover, we found weaker resting-state FC between the hippocampus and the DMN_{dm} after spaced learning than after massed learning. This functional decoupling further suggests less dependence on the hippocampal-cortical transfer after spaced learning, and probably, the memories have already been stably stored in the cortical networks.

Through whole-brain searchlight MVPA, we found that brain regions that distinguished durable memories from weak memories mainly included the DMN_{dm} and the anterior DMN_{core} for spaced learning, whereas the regions largely overlapped with the hippocampus, DMN_{mt}, and posterior DMN_{core} for massed learning. This suggests that the dissociation of the hippocampus and cortex distinguishes between durable and weak memories in the two learning groups. For the successfully retrieved memories in the immediate test, a previous study showed that brain activations differ between the durable and weak memories during encoding[70]. Consistently, our findings provide further evidence that the differences in neural representations underlying durable and weak memories at retrieval depend on learning strategies. Moreover, the dissociation of the hippocampus and cortex aligns with the distinctions of neural activity patterns between the two learning groups, as revealed by hypothesis-driven analysis. Therefore, these results suggest that the neural representation formed in the cortex rather than in the hippocampus contributes to long-term memory retention in spaced learning.

More time between the initial learning and the immediate test in a spaced learning group provides more time for consolidation. This extended period may underlie the higher intertrial similarity and more replay in the cortical networks for spaced learning. Memories undergo a neural integration with time-dependent consolidation, leading memories to change from detailed to gist-like[16,26]. Prior work has shown that cortical neural representations of memories become more integrated with time[23,71]. Moreover, the retrieval-related neural pattern similarity was related to the fast spindle density, which is thought to play an important role in systems memory consolidation[21]. With time, memories are strengthened and integrated across hippocampal-cortical networks through spontaneous replay, which is considered to be the leading mechanism of system consolidation[32,33]. We focused on retrieval patterns rather than encoding patterns, and the memories may undergo different neural reorganization according to their encoding condition. While encoding patterns allow capturing details of memories, retrieval patterns may reflect integrated and gist-like memories. As expected, we found higher intertrial similarity and more replay in the cortical network after spaced learning compared to massed learning. Extended time for consolidation in spaced learning groups enables a greater possibility for spontaneous memory replay, as well as the transfer of memory storage, thus leading to the strengthening and integration of memories in the cortex.

We examined the delay effect by defining durable and weak memories according to different delayed tests. Behaviorally, the differences in retention rates between the two learning groups were consistent according to the 1-week and 1-month delayed tests. However, the correlations between the intertrial similarity of the DMN_{dm} and DMN_{mt} and the 1-weak retention rate were weaker compared to that with the 1-month retention rate. This finding suggests that intertrial similarity after spaced learning preferably predicts durable memory in a relatively long delay. In addition, the findings regarding changes in replay and the identified brain regions capable of distinguishing between durable and weak memories were more pronounced

when using the definition with a relatively long delay. Thus, a relatively long delay might be more suitable for defining durable memory. However, it should be noted that the retention rate might be extremely low when using a very long delay, such as lasting for years.

This study has limitations. First, although sleep is thought to play a crucial role in memory consolidation[72,73], we did not directly explore the neural mechanism of consolidation during sleep. In the future, using electroencephalogram (EEG) or stereo-EEG to examine neural activity during sleep may provide direct evidence of memory consolidation in the spacing effect. Second, the neural activity during encoding, which was usually examined in trial-based design[2,8,74,75], was not recorded during our multi-day learning. In order to investigate long-term memory retention, we conducted multi-day learning and mainly focused on neural activity during retrieval and post-encoding resting state. Between-group differences in intertrial similarity for incorrect trials indicated that the effect was not memory-specific. Although incorrect trials were not successfully retrieved, these trials may be encoded differently in two learning groups. Therefore, we speculate that the between-group differences in the intertrial similarity over all incorrect trials likely reflect differences in encoding for spaced and massed learning. Future research using fMRI scans during the encoding process could help test this possibility. Third, in order to estimate the memory retention for each object-word pair, we repeatedly tested the same pairs in different tests, which may affect the memory durability[76] and the systems consolidation[77,78]. Future studies should consider increasing the number of memory items and testing them independently to minimize potential influences. Fourth, the information load for the two learning groups seems to be different on the day of the immediate test. This might influence the neural activity despite no significant between-group differences in memory performance at the immediate test. An appropriate control for the differences in information load should be considered in future studies. In addition, DMN is particularly associated with boredom or fatigue[79,80]. Although the learning task was relatively simple, all participants were told they would be tested later. We found no significant between-group differences in memory performance at the immediate test, despite the fact that the information load may be different for the two learning groups. Therefore, we speculate that our findings may not be confounded by boredom or fatigue. Finally, although the 1-month definition for durable memories precedes that based on the 1-week delayed test, we cannot confirm whether the 1-month definition is the most suitable. Future studies should include much longer delays for further tests to address this issue.

In conclusion, our findings revealed that spaced learning-induced higher neural pattern similarity in the DMN subsystems during immediate retrieval. Particularly, the neural pattern similarity in the DMN_{dm} and DMN_{mt} subsystems predicted the durable memory. Moreover, we found increased replay of durable memory in the DMN_{dm} for spaced learning and in the hippocampus for both spaced and massed learning. Our findings suggest that time-dependent consolidation promotes neural integration and replay in the cortex, which may underlie the formation of durable memory after spaced learning. This study also sheds light on the importance of delay time scales in predicting durable memory.

Methods

Participants

Sixty-nine participants (38 females; age = 22.49 ± 1.96 years; 35 for spaced learning and 34 for massed learning) were recruited from East China Normal University and were all included in the subsequent behavioral analyses. All participants were right-handed with normal or corrected-to-normal vision and had no history of neurological or psychiatric diseases or psychotropic medication use. For functional imaging analyses, nine participants were excluded due to excessive head motion. Twelve participants were excluded due to the number of trials being less than five in any condition (e.g., durable memory condition and weak memory condition), to ensure a high signal to noise ratio for functional imaging analysis. Thus, 48 (23 for spaced learning and 25 for massed learning) participants were included in the functional imaging analyses. The study protocol was approved by the Ethics Committee of East China Normal University, and each participant signed written consent prior to the experiment. All ethical regulations relevant to human research participants were followed.

Materials

Seventy pairs of picture-word associations were used in this study. The same number of neutral scene and object pictures were chosen from the Geneva Affective Picture Database (GAPED) (https://www.unige.ch/cisa/research/materials-and-online-research/research-material/)[81]. The word stimuli (adjectives) were randomly selected from the MRC psycholinguistics database (https://websites.psychology.uwa.edu.au/school/mrcdatabase/uwa_mrc.htm) and translated into three-character Chinese words. Pictures and words were randomly paired to create 10 and 60 picture-word associations for practice and formal experiments, respectively. An additional 90 words (30 each) were selected for the three memory tests (immediate, 1-week, and 1-month).

Experimental design

The experiment consisted of a learning phase and three memory test sessions after learning (i.e., immediate, 1-week, and 1-month). The participants were randomly assigned to either a 3-day spaced learning group (with two blocks of learning each day) or a 1-day massed learning group (with six blocks of learning per day). The encoding phase was conducted out of the scanner. A total of 60 picture-word pairs were repeatedly used in six blocks for both groups, ensuring that all participants learned each pair six times. For each trial, participants were asked to associate the picture with the word during the 5-s presentation and then rate the success of the established association from 1 (worst) to 5 (best) within the subsequent 4.5 s.

Memory tests were conducted using an MRI scanner and consisted of 60 learned and 30 novel words in each test session. Participants were required to indicate within 4 s whether the presented words were learned and what type of the associated picture is. The response options were as follows: 1 = learned and paired with an object, 2 = learned and paired with a scene, 3 = learned but with an uncertain picture type, and 4 = novel. The intertrial interval was randomly set between 2 s and 10 s accompanied by a fixation cross positioned at the center of the screen.

Both resting-state and task-based fMRI data were collected during the three memory test sessions after learning, with only resting-state fMRI scanning before learning used as the baseline.

Memory performance analysis

The *d*-prime was calculated for each test to estimate each participant's memory performance. The *hit* trials were strictly defined as words that were successfully recalled with the correct picture type to minimize the impact of repeated retrieval, whereas uncertain trials were disregarded. *FA* trials refer to instances in which novel words are incorrectly reported as learned. The *z*-scores of *FA* and *hit* rates were calculated using the inverse of the cumulative Gaussian distribution. Thus, *d-prime* was obtained by subtracting the z-score of the *FA* rate from the z-score of the *hit* rate.

At least two tests are required after learning to evaluate memory retention. We defined durable memories as those successfully retrieved in both the immediate and 1-month delayed tests. The retention rate was quantified based on the percentage of durable memories. Durable and weak memories were defined according to the 1-week delayed test to examine the delay effect, and the retention rate was recalculated.

MRI data acquisition

Imaging data were acquired using a 3.0 Tesla MRI scanner (Siemens, Erlangen, Germany) equipped with a 20-channel head coil at the Shanghai Key Laboratory of Magnetic Resonance, East China Normal University. High-resolution T1-weighted images were obtained using a magnetization-prepared rapid gradient-echo sequence with 192 sagittal slices, slice thickness of 1 mm, no gap, repetition time (TR) of 2,530 ms, echo time (TE) of 2.98 ms, flip angle (FA) of 7°, field of view (FOV) of 256 × 256 mm^2, and voxel size of 1.0 × 1.0 × 1.0 mm^3. Functional MRI scans were collected using

a $T2^*$-weighted echo-planar imaging (EPI) sequence with 33 axial slices, thickness of 3.5 mm, gap of 0.7 mm, TR of 2000 ms, TE of 30 ms, FA of 90°, FOV of 224 × 224 mm^2, and voxel size of 3.5 × 3.5 × 3.5 mm^3. Resting-state and task fMRI data acquisitions lasted for 8 min (240 volumes) and 15 min (450 volumes) in each scanning session.

Preprocessing of functional MRI data

Task fMRI data were preprocessed using Statistical Parametric Mapping 12 (SPM12; http://www.fil.ion.ucl.ac.uk/spm). After converting the raw DICOM data to the NIFTI format, the following preprocessing steps were conducted[82]. The images underwent the correction for slice acquisition delay and rigid-body head movement. Excessive head motion was defined as translation > 2 mm or rotation > 2° in any direction. Subsequently, the corrected images were spatially normalized to the MNI (Montreal Neurological Institute) space and resampled to 3 mm isotropic voxels. According to the requirements of the multivoxel analysis[83], we skipped spatial smoothing to preserve the spatial characteristics of the original data.

For the preprocessing of resting-state fMRI data, the toolbox for Data Processing & Analysis for Brain Imaging (DPABI; https://rfmri.org/DPABI) was used. The main steps were similar to those for task fMRI data except for the steps below. (1) Removing the first 10 volume for signal equilibrium and allowing participants to adapt to the scanning environment; (2) Performing nuisance regression to remove confounding factors, including six head motion parameters, cerebrospinal signals, and white matter signals[84,85]; and (3) A linear drift and bandpass filter (0.01 ~ 0.08 Hz) was taken to reduce low-frequency drift and high-frequency physiological noises[86]. Although differences exist, all the preprocessing steps and parameters for both resting-state and task fMRI data were standard and were used following previous studies[36,87]. For task fMRI data, head-motion parameters were included in the general linear model, and high-pass filtering (1/128 Hz) was also conducted.

ROI definition

Previous studies demonstrated the crucial role of the hippocampus in the spacing effect, and the DMN is increasingly recognized to be associated with human cognition, especially episodic memory[42,43]. In particular, the DMN plays a vital role in the hippocampal-cortical transfer, which supports memory formation[45]. Considering the complex cognitive engagement, the DMN has recently been divided into three subsystems: DMN_{dm}, DMN_{core}, and DMN_{mt}. Therefore, we selected the hippocampus and these three DMN subsystems as ROIs in this study. The bilateral hippocampus was extracted based on the Automated Anatomical Labeling Atlas[62]. The three DMN subsystems were defined based on Yeo's 17-network parcellation[88], which was generated by using a clustering approach and resting-state FC of 1000 participants.

Intertrial similarity analysis

We used RSA[83] to examine the intertrial similarity of the hippocampus and the three DMN subsystems during the retrieval of each test. Specifically, we first modeled retrieval-evoked brain activity using a trial-by-trial GLM[89]. Each trial was modeled using a separate regressor to estimate the neural response. Thus, 90 beta images of interest were obtained from each participant. We constructed a representation similarity matrix for each ROI using pairwise Pearson's correlation based on the beta images of successfully retrieved memories. This step was performed using the RSA toolbox (http://www.mrc-cbu.cam.ac.uk/methods-and-resources/toolboxes/license/). The resultant $N_{trials} \times N_{trials}$ representational similarity matrix was then subjected Fisher's z-transformation. Finally, intertrial similarity was computed by averaging the representational similarity matrix, and the diagonal elements that reflected the within-trial similarity were excluded.

Temporal similarity analysis for durable memories

A temporal similarity analysis was performed to evaluate the persistence of neural representations of durable memories after delay. Specifically, all beta images of durable memories were obtained from trial-by-trial GLMs for each test. For each ROI, pairwise Pearson's correlation analysis was performed for the neural activity patterns of each stimulus among the three tests. The correlation coefficients were subsequently subjected to Fisher's z-transformation. The mean correlation coefficients across the three tests and all durable memories were treated as the temporal similarity for each participant.

Replay analysis

To examine the replay of memories immediately after learning, we first used a GLM to identify the retrieval-evoked neural activity templates. As we hypothesized that durable memories were integrated and owned shared neural representation, we obtained a multivoxel neural activity template based on all durable memories for each participant. Briefly, the durable and weak conditions were considered as variables of interest to obtain the retrieval-evoked neural activity templates of durable and weak memories for each ROI, with incorrect, and novel (unlearned words) conditions as covariates. Beta images of interest obtained from the GLMs were treated as neural activity templates. Subsequently, correlation analyses were performed between each retrieval-evoked neural activity template and the framewise resting activity patterns for each ROI. For each participant, a vector consisting of 230 (TRs) correlation coefficients was obtained for each retrieval-evoked neural activity template and ROI, which was then subjected to Fisher's z-transformation. Subsequently, we aggregated all z-transformed correlation coefficients across all participants for each ROI to obtain a distribution. The 90th percentile[90], not an absolute correlation coefficient, was selected as the threshold to ensure consistency for different brain systems. For example, an absolute threshold of 0.5 might be appropriate for the DMN_{mt} and hippocampus, but it is too strict for the DMN_{dm} and DMN_{core} (Fig. 4b). To diminish baseline individual differences, we calculated the changes in replay number from pre-encoding rest to post-encoding rest.

For validation analyses, the 88th percentile and the 92nd percentile were used. We also used a threshold of 1.5 SD above the mean for each participant following previous studies[36,63], The threshold of 1.5 SD above the mean obtained by pooling the correlation coefficients from the pre-encoding and post-encoding rest sessions of all participants.

Regarding the separation of durable and weak memory templates, we first conducted permutation analysis for each participant. To obtain the shuffled templates for durable and weak memories, we randomly exchanged the trials for durable and weak memories when conducting the GLMs. The shuffled templates for durable and weak memories contained the same number of trials as the corresponding durable and weak memory templates. We conducted the shuffled GLMs 1000 times. Then, we calculated Pearson's correlation between the durable and weak memory templates to obtain the real r value. The 1000 shuffled r values were obtained by calculating Pearson's correlation between shuffled templates for durable and weak memories. For the distribution of the 1000 shuffled r value, we can evaluate the 95% confidence interval. Next, we estimated whether the real r value falls within the 95% confidence interval. For better visualization, the real r value and the 1000 shuffled r values were standardized to z-scores for each participant. Thus, the distribution for each participant can be overlapped and compared, as the positions of the z-scores at ± 1.96 reflect the 95% confidence interval.

Besides the permutation, we also evaluated the separation of durable and weak memory templates through MVPA. A support vector machine with a specified kernel was used to train the classifier for beta images of durable and weak memories obtained from the trial-by-trial GLM. The classification accuracy in each ROI was obtained and compared to chance.

Multivoxel resting-state functional connectivity analysis

To assess the interaction between the hippocampus and the DMN subsystems, the multivoxel FC was calculated. The time courses of all voxels in each ROI were extracted for each participant. Given that ROI_i and ROI_j contain m and n voxels, respectively, Pearson's correlations were calculated between the time courses of each pairwise voxel from the

two different ROIs, resulting in an $m \times n$ connectivity matrix. The connectivity matrix was then subjected to Fisher's z-transformation. The multivoxel FC between ROI_i and ROI_j was calculated by averaging all elements in the connectivity matrix. We calculated multivoxel FC at both baseline and immediately after learning. To reduce the effect of individual differences, we subtracted the baseline FC from the FC in the immediate test for subsequent correlation analysis.

Whole-brain searchlight MVPA analysis

We used a whole-brain searchlight multivariate decoding approach[91] to determine brain areas that distinguish durable from weak memories. The Decoding Toolbox (TDT)[92] was used for this searchlight analysis based on beta images obtained from the trial-by-trial GLM. For each participant, all beta images for durable and weak memories were selected and then randomly and evenly divided into training and testing datasets for machine learning. A spherical ROI (radius = 3 voxels) was moved voxel-by-voxel throughout the entire brain. For each spherical ROI, a support vector machine with a specified kernel was used to train the classifier. The classification accuracy was then assessed and assigned to the center voxel of the ROI. Consequently, we obtained a map of the classification accuracy for each participant. Finally, we used a one-sample t-test to generate group-level classification accuracy maps for the spaced and massed learning groups. We considered AlphaSim corrected $p < 0.05$ as statistically significant[93].

Statistics and reproducibility

We used SPSS 26 and MATLAB R2023a for statistical analyses of the behavioral and fMRI data. Sixty-nine participants were included in behavioral analyses, and 48 participants were included in the functional imaging analyses. Extreme values beyond 2.5 SD from the mean were excluded. The values are presented as the mean ± standard error of mean. Mixed ANOVA was used to assess the interaction of session × group as well as the main effect with partial η^2 (indicated as η^2_p) as the effect size. Post-hoc analyses were conducted using the least significant difference (LSD) test with Bonferroni correction. Independent t tests were used to assess the differences between groups. Spearman's rank correlation coefficient (*rho*) was used to assess the relationships between variables.

Reporting summary

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

Data availability

Source data underlying each graph presented in the main figures are deposited as Supplementary Data. Behavioral data and fMRI metadata are available in https://osf.io/wc3mh/. Raw fMRI data are available from the corresponding author upon reasonable request and the need for a formal data-sharing agreement.

Code availability

The main analysis scripts are available in https://osf.io/wc3mh/.

Received: 7 September 2024; Accepted: 19 March 2025;
Published online: 01 April 2025

References

1. Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T. & Rohrer, D. Distributed practice in verbal recall tasks: a review and quantitative synthesis. *Psychol. Bull.* **132**, 354–380 (2006).
2. Xue, G. et al. Spaced learning enhances subsequent recognition memory by reducing neural repetition suppression. *J. Cogn. Neurosci.* **23**, 1624–1633 (2011).
3. Donovan, J. J. & Radosevich, D. J. A meta-analytic review of the distribution of practice effect: now you see it, now you don't. *J. Appl. Psychol.* **84**, 795–805 (1999).
4. Shea, C. H., Lai, Q., Black, C. & Park, J. H. Spacing practice sessions across days benefits the learning of motor skills. *Hum. Mov. Sci.* **19**, 737–760 (2000).
5. Philips, G. T., Kopec, A. M. & Carew, T. J. Pattern and predictability in memory formation: from molecular mechanisms to clinical relevance. *Neurobiol. Learn Mem.* **105**, 117–124 (2013).
6. Braun, K. & Rubin, D. C. The spacing effect depends on an encoding deficit, retrieval, and time in working memory: evidence from once-presented words. *Memory* **6**, 37–65 (1998).
7. Bradley, M. M. et al. Imaging distributed and massed repetitions of natural scenes: spontaneous retrieval and maintenance. *Hum. Brain Mapp.* **36**, 1381–1392 (2015).
8. Feng, K. et al. Spaced learning enhances episodic memory by increasing neural pattern similarity across repetitions. *J. Neurosci.* **39**, 5351–5360 (2019).
9. Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T. & Pashler, H. Spacing effects in learning: a temporal ridgeline of optimal retention. *Psychol. Sci.* **19**, 1095–1102 (2008).
10. Vilberg, K. L. & Davachi, L. Perirhinal-hippocampal connectivity during reactivation is a marker for object-based memory consolidation. *Neuron* **79**, 1232–1242 (2013).
11. Ezzyat, Y., Inhoff, M. C. & Davachi, L. Differentiation of human medial prefrontal cortex activity underlies long-term resistance to forgetting in memory. *J. Neurosci.* **38**, 10244–10254 (2018).
12. Squire, L. R. & Alvarez, P. Retrograde-amnesia and memory consolidation - a neurobiological perspective. *Curr. Opin. Neurobiol.* **5**, 169–177 (1995).
13. Moscovitch, M., Cabeza, R., Winocur, G. & Nadel, L. Episodic memory and beyond: the hippocampus and neocortex in transformation. *Annu. Rev. Psychol.* **67**, 105–134 (2016).
14. Lechner, H. A., Squire, L. R. & Byrne, J. H. 100 years of consolidation - remembering muller and pilzecker. *Learn. Mem.* **6**, 77–87 (1999).
15. Dudai, Y. The neurobiology of consolidations, or, how stable is the engram?. *Annu. Rev. Psychol.* **55**, 51–86 (2004).
16. Winocur, G., Moscovitch, M. & Bontempi, B. Memory formation and long-term retention in humans and animals: convergence towards a transformation account of hippocampal-neocortical interactions. *Neuropsychologia* **48**, 2339–2356 (2010).
17. Dandolo, L. C. & Schwabe, L. Time-dependent memory transformation along the hippocampal anterior-posterior axis. *Nat. Commun.* **9**, 1205 (2018).
18. Robin, J. & Moscovitch, M. Details, gist and schema: hippocampal-neocortical interactions underlying recent and remote episodic and spatial memory. *Curr. Opin. Behav. Sci.* **17**, 114–123 (2017).
19. Mcclelland, J. L., Mcnaughton, B. L. & Oreilly, R. C. Why there are complementary learning-systems in the hippocampus and neocortex - insights from the successes and failures of connectionist models of learning and memory. *Psychol. Rev.* **102**, 419–457 (1995).
20. Alvarez, P. & Squire, L. R. Memory consolidation and the medial temporal-lobe - a simple network model. *PNAS* **91**, 7041–7045 (1994).
21. Cowan, E. et al. Sleep spindles promote the restructuring of memory representations in ventromedial prefrontal cortex through enhanced hippocampal-cortical functional connectivity. *J. Neurosci.* **40**, 1909–1919 (2020).
22. Liu, Y., et al. Memory consolidation reconfigures neural pathways involved in the suppression of emotional memories. *Nat. Commun.* **7**, 13375 (2016).
23. Tompary, A. & Davachi, L. Consolidation promotes the emergence of representational overlap in the hippocampus and medial prefrontal cortex. *Neuron* **96**, 228–241 (2017).
24. Frankland, P. W. & Bontempi, B. The organization of recent and remote memories. *Nat. Rev. Neurosci.* **6**, 119–130 (2005).
25. Takashima, A. et al. Shift from hippocampal to neocortical centered retrieval network with consolidation. *J. Neurosci.* **29**, 10087–10093 (2009).

26. Dudai, Y., Karni, A. & Born, J. The consolidation and transformation of memory. *Neuron* **88**, 20–32 (2015).
27. Krenz, V., Alink, A., Sommer, T., Roozendaal, B. & Schwabe, L. Time-dependent memory transformation in hippocampus and neocortex is semantic in nature. *Nat. Commun.* **14**, 6037 (2023).
28. Bontempi, B., Laurent-Demir, C., Destrade, C. & Jaffard, R. Time-dependent reorganization of brain circuitry underlying long-term memory storage. *Nature* **400**, 671–675 (1999).
29. Rasch, B. & Born, J. Maintaining memories by reactivation. *Curr. Opin. Neurobiol.* **17**, 698–703 (2007).
30. Foster, D. J. Replay comes of age. *Annu. Rev. Neurosci.* **40**, 581–602 (2017).
31. Deuker, L. et al. Memory consolidation by replay of stimulus-specific neural activity. *J. Neurosci.* **33**, 19373–19383 (2013).
32. Carr, M. F., Jadhav, S. P. & Frank, L. M. Hippocampal replay in the awake state: a potential substrate for memory consolidation and retrieval. *Nat. Neurosci.* **14**, 147–153 (2011).
33. Tambini, A. & Davachi, L. Awake reactivation of prior experiences consolidates memories and biases cognition. *Trends Cogn. Sci.* **23**, 876–890 (2019).
34. van de Ven, G. M., Trouche, S., McNamara, C. G., Allen, K. & Dupret, D. Hippocampal offline reactivation consolidates recently formed cell assembly patterns during sharp wave-ripples. *Neuron* **92**, 968–974 (2016).
35. Girardeau, G. & Zugaro, M. Hippocampal ripples and memory consolidation. *Curr. Opin. Neurobiol.* **21**, 452–459 (2011).
36. Schapiro, A. C., McDevitt, E. A., Rogers, T. T., Mednick, S. C. & Norman, K. A. Human hippocampal replay during rest prioritizes weakly learned information and predicts memory performance. *Nat. Commun.* **9**, 3920 (2018).
37. Schuck, N. W. & Niv, Y. Sequential replay of nonspatial task states in the human hippocampus. *Science* **364**, 1254–125 (2019).
38. Tambini, A. & Davachi, L. Persistence of hippocampal multivoxel patterns into postencoding rest is related to memory. *PNAS* **110**, 19591–19596 (2013).
39. Staresina, B. P., Alink, A., Kriegeskorte, N. & Henson, R. N. Awake reactivation predicts memory in humans. *PNAS* **110**, 21159–21164 (2013).
40. Tambini, A., Ketz, N. & Davachi, L. Enhanced brain correlations during rest are related to memory for recent experiences. *Neuron* **65**, 280–290 (2010).
41. Wagner, I. C., van Buuren, M. & Fernandez, G. Thalamo-cortical coupling during encoding and consolidation is linked to durable memory formation. *NeuroImage* **197**, 80–92 (2019).
42. Menon, V. 20 years of the default mode network: a review and synthesis. *Neuron* **111**, 2469–2487 (2023).
43. Kaefer, K., Stella, F., McNaughton, B. L. & Battaglia, F. P. Replay, the default mode network and the cascaded memory systems model. *Nat. Rev. Neurosci.* **23**, 628–640 (2022).
44. Kaplan, R. et al. Hippocampal sharp-wave ripples influence selective activation of the default mode network. *Curr. Biol.* **26**, 686–691 (2016).
45. Higgins, C. et al. Replay bursts in humans coincide with activation of the default mode and parietal alpha networks. *Neuron* **109**, 882–893 (2021).
46. Huang, Q., et al. Replay-triggered brain-wide activation in humans. *Nat. Commun.* **15**, 7185 (2024).
47. Reagh, Z. M. & Ranganath, C. Flexible reuse of cortico-hippocampal representations during encoding and recall of naturalistic events. *Nat. Commun.* **14**, 1279 (2023).
48. Smallwood, J. et al. The default mode network in cognition: a topographical perspective. *Nat. Rev. Neurosci.* **22**, 503–513 (2021).
49. Kim, H. Default network activation during episodic and semantic memory retrieval: a selective meta-analytic comparison. *Neuropsychologia* **80**, 35–46 (2016).
50. Eichenbaum, H. Prefrontal-hippocampal interactions in episodic memory. *Nat. Rev. Neurosci.* **18**, 547–558 (2017).
51. Sandrini, M., Censor, N., Mishoe, J. & Cohen, L. G. Causal role of prefrontal cortex in strengthening of episodic memories through reconsolidation. *Curr. Biol.* **23**, 2181–2184 (2013).
52. Nolde, S. F., Johnson, M. K. & D'Esposito, M. Left prefrontal activation during episodic remembering: an event-related fMRI study. *Neuroreport* **9**, 3509–3514 (1998).
53. Xue, C. et al. Distinct disruptive patterns of default mode subnetwork connectivity across the spectrum of preclinical Alzheimer's disease. *Front. Aging Neurosci.* **11**, 307 (2019).
54. Andrews-Hanna, J. R., Smallwood, J. & Spreng, R. N. The default network and self-generated thought: component processes, dynamic control, and clinical relevance. *Ann. N. Y. Acad. Sci.* **1316**, 29–52 (2014).
55. Yao, W. et al. Core-centered connection abnormalities associated with pathological features mediate the progress of cognitive impairments in Alzheimer's disease spectrum patients. *J. Alzheimer's. Dis.* **82**, 1499–1511 (2021).
56. Hou, Y. et al. Primary disruption of the default mode network subsystems in drug-naive Parkinson's disease with mild cognitive impairments. *Neuroradiology* **62**, 685–692 (2020).
57. Kaboodvand, N., Bäckman, L., Nyberg, L. & Salami, A. The retrosplenial cortex: a memory gateway between the cortical default mode network and the medial temporal lobe. *Hum. Brain Mapp.* **39**, 2020–2034 (2018).
58. Chen, Q., Turnbull, A., Baran, T. M. & Lin, F. V. Longitudinal stability of medial temporal lobe connectivity is associated with tau-related memory decline. *Elife* **9**, e62114 (2020).
59. Terry, D. P., Sabatinelli, D., Puente, A. N., Lazar, N. A. & Miller, L. S. A meta-analysis of fMRI activation differences during episodic memory in Alzheimer's disease and mild cognitive impairment. *J. Neuroimaging***25**, 849–860 (2015).
60. Rizzolo, L. et al. Relationship between brain AD biomarkers and episodic memory performance in healthy aging. *Brain. Cogn.* **148**, 105680 (2021).
61. Litman, L. & Davachi, L. Distributed learning enhances relational memory consolidation. *Learn. Mem.* **15**, 711–716 (2008).
62. Tzourio-Mazoyer, N. et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage* **15**, 273–289 (2002).
63. Tanriverdi, B. et al. Awake hippocampal-cortical co-reactivation is associated with forgetting. *J. Cogn. Neurosci.* **35**, 1446–1462 (2023).
64. Li, C. & Yang, J. Role of the hippocampus in the spacing effect during memory retrieval. *Hippocampus* **30**, 703–714 (2020).
65. Takashima, A. et al. Memory trace stabilization leads to large-scale changes in the retrieval network: a functional MRI study on associative memory. *Learn. Mem.* **14**, 472–479 (2007).
66. Ellenbogen, J. M., Hulbert, J. C., Stickgold, R., Dinges, D. F. & Thompson-Schill, S. L. Interfering with theories of sleep and memory: sleep, declarative memory, and associative interference. *Curr. Biol.* **16**, 1290–1294 (2006).
67. Lutz, N. D., Martinez-Albert, E., Friedrich, H., Born, J. & Besedovsky, L. Sleep shapes the associative structure underlying pattern completion in multielement event memory. *PNAS* **121**, e2314423121 (2024).
68. Toppino, T. C. & Gerbier, E. About practice: repetition, spacing, and abstraction. *Psychol. Learn. Motiv.* **60**, 113–189 (2014).
69. Tompary, A. & Davachi, L. Integration of overlapping sequences emerges with consolidation through medial prefrontal cortex neural ensembles and hippocampal–cortical connectivity. *eLife* **13**, e84359 (2024).
70. Wagner, I. C., van Buuren, M., Bovy, L. & Fernandez, G. Parallel engagement of regions associated with encoding and later retrieval forms durable memories. *J. Neurosci.* **36**, 7985–7995 (2016).

71. Richards, B. A. et al. Patterns across multiple memories are identified over time. *Nat. Neurosci.* **17**, 981–986 (2014).
72. Bendor, D. & Wilson, M. A. Biasing the content of hippocampal replay during sleep. *Nat. Neurosci.* **15**, 1439–1444 (2012).
73. Gorriz, M. H., Takigawa, M. & Bendor, D. The role of experience in prioritizing hippocampal replay. *Nat. Commun.* **14**, 8157 (2023).
74. Callan, D. E. & Schweighofer, N. Neural correlates of the spacing effect in explicit verbal semantic encoding support the deficient-processing theory. *Hum. Brain Mapp.* **31**, 645–659 (2010).
75. Zou, F., et al. Re-expression of CA1 and entorhinal activity patterns preserves temporal context memory at long timescales. *Nat. Commun.* **14**, 4350 (2023).
76. Rowland, C. A. The effect of testing versus restudy on retention: a meta-analytic review of the testing effect. *Psychol. Bull.* **140**, 1432–1463 (2014).
77. Hebscher, M., Wing, E., Ryan, J. & Gilboa, A. Rapid cortical plasticity supports long-term memory formation. *Trends Cogn. Sci.* **23**, 989–1002 (2019).
78. Antony, J. W., Ferreira, C. S., Norman, K. A. & Wimber, M. Retrieval as a fast route to memory consolidation. *Trends Cogn. Sci.* **21**, 573–576 (2017).
79. Danckert, J. & Merrifield, C. Boredom, sustained attention and the default mode network. *Exp. Brain Res.* **236**, 2507–2518 (2018).
80. Gui, D. et al. Resting spontaneous activity in the default mode network predicts performance decline during prolonged attention workload. *NeuroImage* **120**, 323–330 (2015).
81. Dan-Glauser, E. S. & Scherer, K. R. The Geneva affective picture database (GAPED): a new 730-picture database focusing on valence and normative significance. *Behav. Res. Methods* **43**, 468–477 (2011).
82. Yin, D., et al. Failure in cognitive suppression of negative affect in adolescents with generalized anxiety disorder. *Sci. Rep.* **7**, 6583 (2017).
83. Kriegeskorte, N., Mur, M. & Bandettini, P. A. Representational similarity analysis-connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* **2**, 4 (2008).
84. Liao, X., Cao, M., Xia, M. & He, Y. Individual differences and time-varying features of modular brain architecture. *NeuroImage* **152**, 94–107 (2017).
85. Zhang, J., et al. Neural, electrophysiological and anatomical basis of brain-network variability and its characteristic changes in mental disorders. *Brain* **139**, 2307–2321 (2016).
86. Biswal, B., Yetkin, F. Z., Haughton, V. M. & Hyde, J. S. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn. Reson. Med.* **34**, 537–541 (1995).
87. de Voogd, L. D., Fernandez, G. & Hermans, E. J. Awake reactivation of emotional memory traces through hippocampal-neocortical interactions. *NeuroImage* **134**, 563–572 (2016).
88. Yeo, B. T., et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* **106**, 1125–1165 (2011).
89. Mumford, J. A., Turner, B. O., Ashby, F. G. & Poldrack, R. A. Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *NeuroImage* **59**, 2636–2643 (2012).
90. Livne, T. et al. Spontaneous activity patterns in human motor cortex replay evoked activity patterns for hand movements. *Sci. Rep.* **12**, 16867 (2022).
91. Kriegeskorte, N., Goebel, R. & Bandettini, P. Information-based functional brain mapping. *PNAS* **103**, 3863–3868 (2006).
92. Hebart, M. N., Gorgen, K. & Haynes, J. D. The Decoding Toolbox (TDT): a versatile software package for multivariate analyses of functional imaging data. *Front. Neuroinform.* **8**, 88 (2014).
93. Yan, C., Wang, X., Zuo, X. & Zang, Y. DPABI: data processing & analysis for (resting-state) brain imaging. *Neuroinformatics* **14**, 339–351 (2016).

Acknowledgements

This work was supported by the National Natural Science Foundation of China (32271096 to D.Z.Y.), the STI 2030—Major Projects (2021ZD0200500 to D.Z.Y.), and the Fundamental Research Funds for the Central Universities (D.Z.Y.).

Author contributions

Conceptualization: D.Z.Y. and Y.F.X.Y. Methodology: Y.F.X.Y., Z.Y.H., and Y.Y. Investigation: Y.F.X.Y. and M.X.F. Visualization: Y.F.X.Y. and Z.Y.H. Supervision: D.Z.Y. Writing (original draft): D.Z.Y. and Y.F.X.Y. Writing (review and editing): Y.F.X.Y., Z.Y.H., Y.Y., M.X.F., and D.Z.Y.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-025-07964-6.

Correspondence and requests for materials should be addressed to Dazhi Yin.

Peer review information *Communications Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editor: Benjamin Bessieres. A peer review file is available.