

Parallel Engagement of Regions Associated with Encoding and Later Retrieval Forms Durable Memories

Isabella C. Wagner, Mariët van Buuren, Leonore Bovy, and Guillén Fernández

Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, 6525 EZ, Nijmegen, The Netherlands

The fate of a memory is partly determined at initial encoding. However, the behavioral consequences of memory formation are often tested only once and shortly after learning, which leaves the neuronal predictors for the formation of durable memories largely unknown. Here, we hypothesized that durable memory formation (as opposed to weak or no memory formation) is reflected through increased activation in the medial temporal lobes and prefrontal cortex, and more consistent processing (i.e., stronger pattern similarity) across encoding material. Thirty-four human subjects studied unique picture–location associations while undergoing fMRI and performed a cued recall test immediately after study as well as 48 h later. Associative memories were defined as “weak” if they were retrieved during the immediate test only. Conversely, “durable” memories persisted also after 48 h. The posterior cingulate cortex showed increased pattern similarity during successful memory formation, independent of the eventual durability. For durable memory encoding, we found increased activation in medial and inferior temporal, prefrontal, and parietal regions. This was accompanied by stronger pattern similarity in lateral prefrontal and parietal regions, as well as in anterior and posterior midline structures that were also engaged during later memory retrieval. Thus, we show that pattern similarity, or consistent processing, in the posterior cingulate cortex predicts associative memory formation at encoding. If this is paralleled by additional activation increases in regions typically related to encoding, and by consistent processing in regions involved in later retrieval, formed memories appear durable for at least 48 h.

Key words: fMRI; memory durability; posterior cingulate cortex; representational similarity analysis; subsequent memory

Significance Statement

Successful memory formation is typically associated with increased neuronal activation in medial temporal and prefrontal regions at encoding, but memory is often assessed only once and shortly after study. Here, we addressed memory durability, and investigated the neuronal underpinnings of encoding for associations remembered over a longer period of time, less long, or immediately forgotten. We showed that durable memory formation is dependent on increased activation in the hippocampus and neocortical regions related to encoding, and on consistent processing of associative memory traces in midline structures that are involved in later memory retrieval. These findings highlight how durable memories are formed.

Introduction

What we remember is partly determined by processes occurring as we initially encounter information. Such processes are experimentally probed by the so-called “subsequent memory effect,” for which stimulus activity acquired at encoding is sorted as later remembered or forgotten (Brewer, 1998; Wagner et al., 1998; Fernández et al., 1999). Memory, however, is frequently deter-

mined by applying a single retrieval test shortly after encoding, which makes it impossible to draw conclusions about longer-lasting memories. The neuronal predictors for the ecologically and biologically more relevant formation of durable memories thus remain largely unknown.

Naturally, durable memory is delineated by testing stimulus material at least at two time points. Previous studies showed that the encoding of durable memories (remembered at a later test) compared with weak memories (remembered only immediately) was associated with increased activation in the hippocampus, surrounding medial temporal lobes (MTLs), and lateral prefrontal cortex (LPFC; Uncapher and Rugg, 2005; Carr et al., 2010). Additionally, durable memory formation yielded enhanced functional coupling between the hippocampus and neocortical representational regions (Sneve et al., 2015), as well as between the LPFC and posterior cingulate cortex (PCC; Liu et al., 2013). In contrast to the MTL and LPFC (Spaniol et al., 2009; Kim, 2011),

Received March 14, 2016; revised June 13, 2016; accepted June 15, 2016.

Author contributions: I.C.W., M.v.B., and G.F. designed research; I.C.W. and L.B. performed research; I.C.W. contributed unpublished reagents/analytic tools; I.C.W. analyzed data; I.C.W., M.v.B., L.B., and G.F. wrote the paper.

This project was supported by the European Research Council (Grant ERC R0001075) to G.F. We thank Eelco van Dongen for initial help with the paradigm.

The authors declare no competing financial interests.

Correspondence should be addressed to Isabella C. Wagner, Donders Institute for Brain, Cognition and Behaviour, Centre for Cognitive Neuroimaging, Kapittelweg 29, 6525 EN, Nijmegen, The Netherlands. E-mail: i.wagner@donders.ru.nl.

DOI:10.1523/JNEUROSCI.0830-16.2016

Copyright © 2016 the authors 0270-6474/16/367985-11\$15.00/0

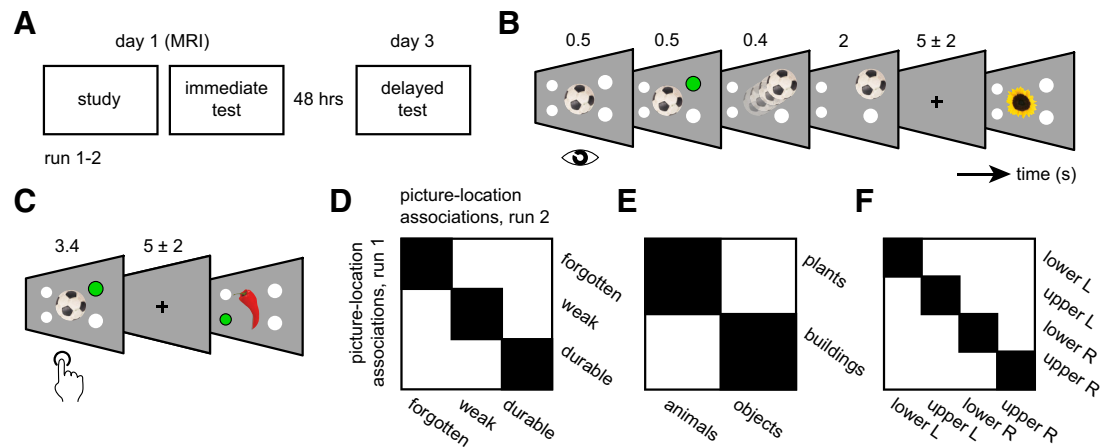


Figure 1. Study timeline, associative memory task, and representational similarity analysis. **A**, Subjects intentionally encoded unique picture–location associations inside the MR scanner across two runs. After each study phase, subjects performed a cued recall test (immediate test, day 1), as well as a delayed test after 48 h (day 3, behavioral laboratory; see also Materials and Methods, Task and procedure). **B**, During study trials, pictures were randomly associated with one of four locations. Each trial started with the presentation of the picture in the center of the screen (0.5 s), after which the target location was indicated (0.5 s). The picture then moved to the target location (0.4 s; movement trajectory is schematically indicated) and remained there for 2 s. The intertrial interval varied randomly between 3 and 7 s (mean, 5 s), and the next trial started thereafter. **C**, During immediate and delayed tests, the picture was presented centrally, and subjects were required to indicate the correct location by pressing one of four buttons (3.4 s). The intertrial interval varied randomly between 3 and 7 s (mean, 5 s), and the next trial started thereafter. **D–F**, RSA was performed across the unique picture–location associations: trials were sorted based on their memory durability (**D**), their picture category (**E**), and their location (**F**; L, left; R, right). Pattern similarity values were extracted from respective quadrants (marked in black; see Materials and Methods, Representational similarity analysis).

the PCC is commonly deactivated during successful memory encoding (Daselaar et al., 2004, 2009), possibly reflecting internal orientation and self-referential processing (Huijbers et al., 2012). Together, these studies provide first evidence of how levels of activation and functional connectivity might foster durable memory formation.

Complementary to levels of activation and connectivity, durable memory might rely on the spatial organization of distributed activation patterns at encoding, which can be assessed using representational similarity analysis (RSA; Kriegeskorte et al., 2008). RSA quantifies the spatial correspondence of activation patterns, whereby similar patterns are thought to share information content. For example, the pattern similarity of representations across item repetitions (i.e., “self-similarity”; Xue et al., 2010, 2013) or different stimuli (i.e., “global similarity”; Visser et al., 2013) at encoding was associated with better memory. These representations could further be stabilized through attention (Aly and Turk-Browne, 2016a,b) and reward (Wolosin et al., 2013). Moreover, pattern similarity in the PCC was advocated to promote memory encoding (Xue et al., 2013) and consolidation (Bird et al., 2015). These findings suggest that associative memory formation (i.e., the linking of multiple stimulus features) is facilitated by means of increased similarity, or “consistency”, across encoding patterns. How consistent encoding representations modulate memory durability is, however, unknown.

Here, we asked not only how durable memory formation is related to the level of activation, but to the consistency of processing across encoding trials. We chose a design that allowed us to delineate durable memory formation by testing all stimulus material twice. Subjects studied unique picture–location associations and performed a cued recall test immediately after study (both inside the MR scanner), as well as 48 h later (in the behavioral laboratory; Fig. 1A). We defined associative memories as “weak” if they were preserved during immediate retrieval only. “Durable” memories also persisted 48 h later.

First, we hypothesized that the encoding of durable compared to weak (and forgotten) associations would produce enhanced

activation within the hippocampus, surrounding MTL, and LPFC. Second, and most critically, we proposed that consistent processing across different stimulus content might mirror the buildup of associative traces. We expected that durable memory formation would particularly benefit from consistent processing. This consistency was quantified by deriving the multivoxel pattern similarity of unique picture–location associations, embedded within a whole-brain, single-trial RSA framework. We predicted increased pattern similarity for durable associations relative to weak and forgotten associations within the hippocampus and surrounding MTL, LPFC, and PCC.

Materials and Methods

Subjects. Thirty-five subjects participated in this experiment (23 female; age range, 18–29 years; mean age, 23 years). All subjects were right handed, healthy, had normal or corrected-to-normal vision, and gave written informed consent before participation. One subject was excluded due to technical problems with the MR gradient coil. Final analyses were thus completed including 34 subjects (23 female; age range, 18–29 years; mean age, 23 years). The study was approved by the institutional review board (Commissie Mensgebonden Onderzoek, region Arnhem-Nijmegen, The Netherlands).

Task and procedure. Subjects were instructed to memorize 192 picture–location associations that were distributed over two experimental runs. Each run consisted of a study (15 min) and a test phase (15 min; Fig. 1A). In between, subjects remained in the scanner for a short rest period (6 min) while a white fixation cross was presented on the computer screen, and subjects were instructed to remain awake with their eyes open (not depicted).

During each study phase, subjects memorized 96 color pictures (i.e., 192 across two runs; see Materials and Methods, Stimulus material and randomization) that were randomly associated with one of four locations presented on the computer screen (lower left, upper left, lower right, upper right; similar to Takashima et al., 2009; van Dongen et al., 2011, 2012). We chose this paradigm since the binding of different pictures with a location robustly engages the hippocampus, leading to a hippocampal-dependent spatial associative memory trace (Brown and Aggleton, 2001; Mayes et al., 2007). Additionally, we incorporated different stimulus categories and locations to induce activation (patterns) in

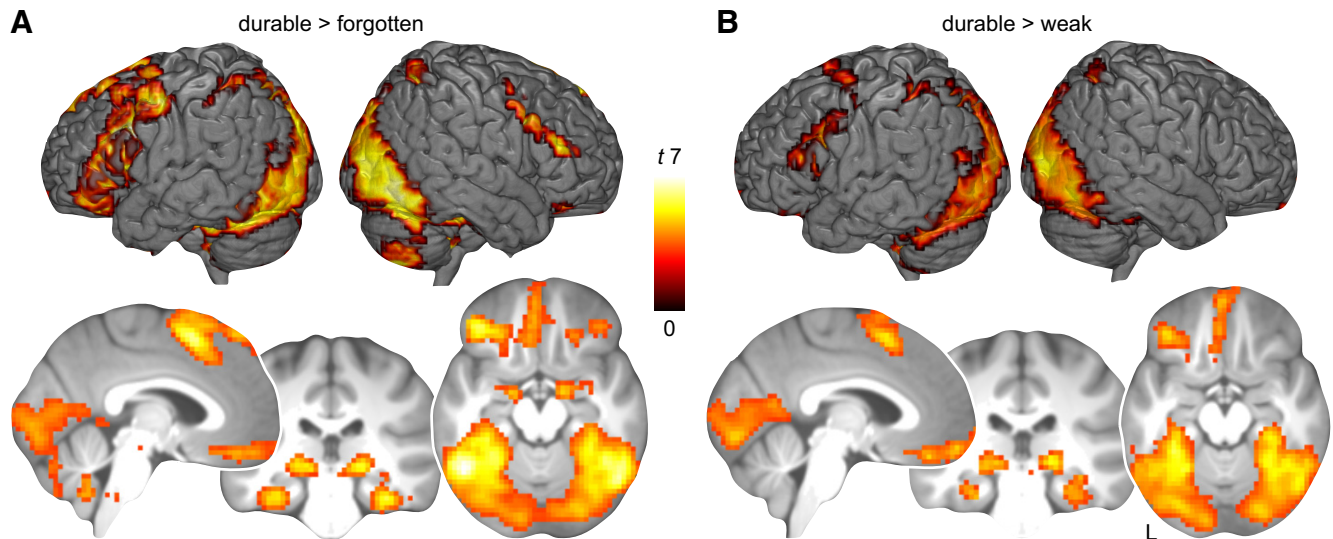


Figure 2. Activation during encoding of durable memories. **A, B**, Increased BOLD activation during encoding of durable compared to forgotten associations (durable > forgotten; **A**) and during encoding of durable compared to weak associations (durable > weak; **B**). Results are shown at $p < 0.001$ ($p < 0.05$, FWE-corrected at cluster level; Table 1), and at MNI coordinates: $x = -3, y = -27, z = -17$. Slices for all figures are based on the average structural scan of the subjects. L, Left.

Table 1. Activation during encoding

| Contrast and brain region | MNI coordinates | | | z-value | Cluster size |
|-------------------------------|-----------------|-----|-----|---------|--------------|
| | x | y | z | | |
| Main effect* | | | | | |
| L fusiform gyrus | -27 | -51 | -15 | | 22,880 |
| R angular gyrus | 63 | -51 | 36 | | 447 |
| R middle frontal gyrus | 45 | 12 | 30 | 7.82 | 1021 |
| L supramarginal gyrus | -60 | -48 | 45 | 6.84 | 174 |
| R cingulate gyrus | 3 | -21 | 39 | 6.78 | 137 |
| L cingulate gyrus | -6 | 36 | 0 | 6.77 | 1187 |
| L cingulate gyrus | -3 | 3 | 30 | 6.77 | 132 |
| L middle frontal gyrus | -30 | 33 | 33 | 5.71 | 71 |
| Cerebellum | -36 | -87 | -33 | 5.67 | 204 |
| R superior temporal gyrus | 63 | -18 | -3 | 4.44 | 91 |
| Durable > forgotten | | | | | |
| L inferior temporal gyrus | -48 | -51 | -12 | 6.86 | 12,747 |
| Cerebellum | 21 | -39 | -45 | 5.94 | 230 |
| R middle frontal gyrus | 45 | 9 | 30 | 5.12 | 261 |
| R lateral orbitofrontal gyrus | 36 | 36 | -15 | 4.33 | 92 |
| Durable > weak | | | | | |
| R lingual gyrus | 33 | -42 | -9 | 5.67 | 8508 |
| L lateral orbitofrontal gyrus | -39 | 30 | -15 | 5.30 | 192 |
| Cerebellum | -24 | -36 | -45 | 5.20 | 153 |
| L precentral gyrus | -42 | 3 | 27 | 4.90 | 525 |
| L caudate | -15 | 6 | 12 | 4.88 | 102 |
| L superior frontal gyrus | -3 | 12 | 57 | 4.75 | 290 |
| L gyrus rectus | -3 | 39 | -21 | 4.57 | 193 |

MNI coordinates represent the location of peak voxels. We report the first local maximum within each cluster. Effects were tested for significance using cluster inference with a cluster-defining threshold of $p < 0.001$ and a cluster probability of $p < 0.05$, FWE-corrected for multiple comparisons (critical cluster size: 65 voxels). L, Left; R, right. Anatomical nomenclature for all tables was obtained from the Laboratory for Neuro Imaging (LONI) Brain Atlas (LPBA40; <http://www.loni.usc.edu/atlas/>).

*Results for the main effect were thresholded at $p < 0.05$ FWE-corrected for multiple comparisons.

clearly separable neuronal regions. A trial started with the presentation of the picture in the center of the screen (1 s) together with the four surrounding screen locations as filled white circles. After 500 ms, one of the filled circles turned green indicating the target location of the respective picture. The picture then moved to that target location (400 ms) and remained there for 2 s. Intertrial intervals varied randomly between 3 and 7 s (mean, 5 s), during which time a fixation cross was presented (Fig. 1B). Subjects were provided with a break of 25 s every 32 trials, indicated by asterisks on the computer screen.

During the immediate test, subjects were prompted for their memory of all picture–location associations that were shown during the preceding study phase (i.e., 96 trials/run). Again, pictures were presented in the center of the screen surrounded by the four filled circles indicating the four alternative screen locations (3.4 s; Fig. 1C). Subjects were required to press one of four buttons (each assigned to a specific location) using the middle and index fingers of both hands. Trials were separated by a fixation period ranging between 3 and 7 s (mean, 5 s), and a break of 25 s was given every 32 trials.

The delayed test was performed in front of a computer screen in a behavioral laboratory on day 3 (day 1–day 3 difference: mean, 47 h; range, 45–50 h). Timing and structure were identical to the immediate test (day 1), but with a new pseudorandom order of all 192 items (see Materials and methods, Stimulus material and randomization). The experiment was programmed and presented with Presentation (version 16.4, Neurobehavioral Systems, www.neurobs.com).

Stimulus material and randomization. Stimulus material consisted of 192 color photographs (e.g., animals, plants, objects, buildings; 48 pictures each), derived from the Hemera Photo-Object database (Hemera Technologies) and the Internet. All were unique and easy to name. Pictures were resized to 400×300 pixels and presented on a gray background. For each run, the pictures of two categories were presented (always one animate with one inanimate picture category). This resulted in four presentation sequences that were randomized in groups of four subjects (run1, run2; sequence 1: animals and objects, plants and buildings; sequence 2: animals and buildings, plants and objects; sequence 3: plants and buildings, animals and objects; sequence 4: plants and objects, animals and buildings).

During the immediate test on day 1, all picture–location associations from the preceding study phase were tested. The presentation order from the study phase was split into quarters (24 trials each) to control for temporal distance between study and test presentations of a specific picture. Within these quarters, trial presentation was shuffled randomly. Hence, trials presented in the first quarter during the study appeared in the first quarter during the immediate test, but in a different order.

For the delayed test (day 3), all associations that were learned on day 1 were tested again. Pictures were presented in a pseudorandom order without controlling for temporal distance to the respective study presentation. During all study and test phases of the experiment (day 1 and 3), presentation orders were restricted such that no more than three pictures of the same category (e.g., animals, plants, objects, buildings) were presented in succession. Additionally, no more than three successive pictures were associated with the same location. The pairing of picture–

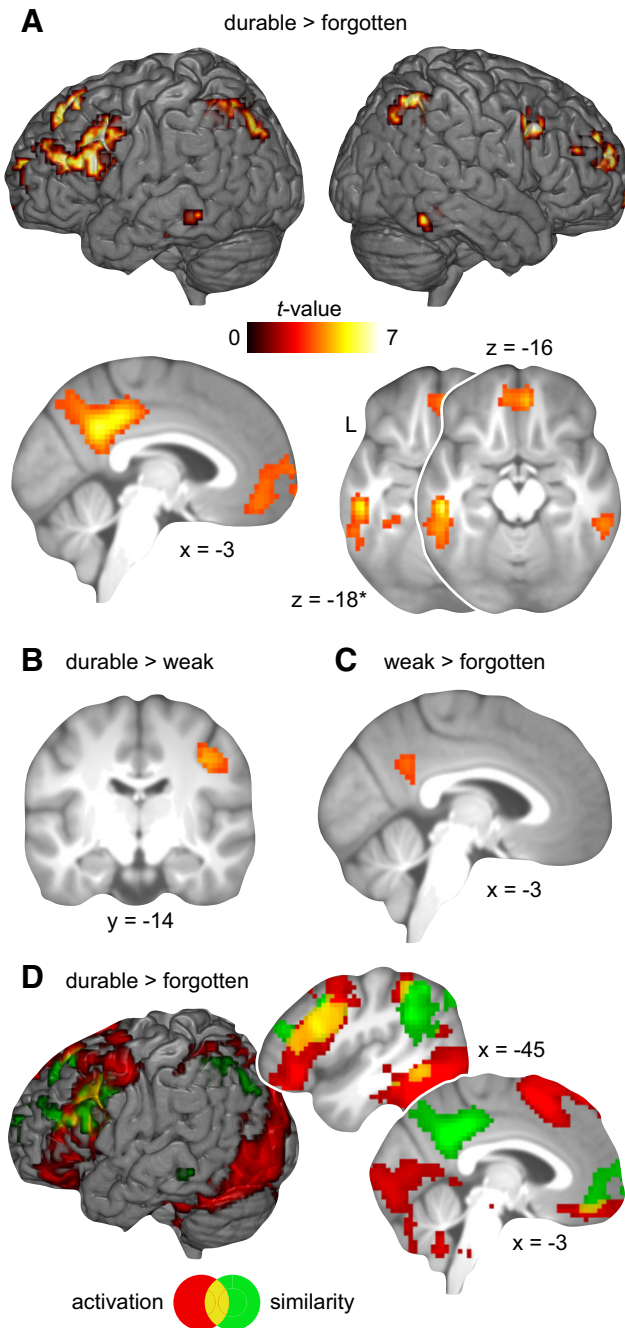


Figure 3. Pattern similarity during encoding, and overlap with univariate levels of activation. **A–C**, Increased pattern similarity during encoding of durable > forgotten (**A**), durable > weak (**B**), and weak > forgotten (**C**) associations. Results are shown at $p < 0.001$ ($p < 0.05$, FWE-corrected at cluster level; see also Table 2). *Slice is shown at $p < 0.001$, uncorrected. The parahippocampal result survived small volume correction with an a priori anatomical ROI. **D**, Overlay of univariate activation (durable > forgotten; Fig. 2A) and pattern similarity results (same contrast; **A**). L, Left.

location associations was randomized across subjects, and picture categories were associated with the four different locations in equal amounts.

Behavioral data analysis. We derived a measure of memory durability by sorting trials based on the subjects' performance at the immediate test (day 1) and the delayed test (day 3). This resulted in three types of responses: picture–location associations that were (1) already forgotten on day 1 (“forgotten”); (2) remembered on day 1 but forgotten on day 3 (“weak”); or (3) remembered at both tests (“durable”). Picture–location associations that were forgotten at the immediate test (day 1) but were recalled correctly at the delayed test (day 3; mean \pm SEM: 16.1 \pm 1.3

Table 2. Pattern similarity during encoding

| Contrast and brain region | MNI coordinates | | | z-value | Cluster size |
|-------------------------------|-----------------|-----|-----|---------|--------------|
| | x | y | z | | |
| Main effect* | | | | | |
| R middle occipital gyrus | 27 | −90 | 6 | | 33,905 |
| Durable > forgotten | | | | | |
| R posterior cingulate cortex† | 6 | −36 | 39 | 5.55 | 1684 |
| L middle temporal gyrus† | −51 | −27 | −15 | 5.46 | 302 |
| R middle frontal gyrus† | 21 | 45 | 18 | 4.84 | 749 |
| L middle frontal gyrus† | −27 | 30 | 51 | 4.71 | 144 |
| L inferior frontal gyrus† | −48 | 18 | 24 | 4.66 | 929 |
| R angular gyrus† | 48 | −45 | 36 | 4.54 | 324 |
| R middle frontal gyrus† | 54 | 15 | 39 | 4.40 | 120 |
| R inferior temporal gyrus† | 57 | −39 | −15 | 3.80 | 105 |
| Durable > weak | | | | | |
| R precentral gyrus | 39 | −15 | 39 | 4.43 | 109 |
| Weak > forgotten | | | | | |
| L posterior cingulate cortex | −6 | −51 | 30 | 3.59 | 85 |

MNI coordinates represent the location of peak voxels. We report the first local maximum within each cluster. Effects were tested for significance using cluster inference with a cluster-defining threshold of $p < 0.001$ and a cluster probability of $p < 0.05$ FWE-corrected for multiple comparisons (critical cluster size, 72 voxels). L, Left; R, right.

*Results for the main effect were thresholded at $p < 0.05$ FWE-corrected for multiple comparisons.

†Significant clusters after exclusively masking pattern similarity (durable > forgotten) with combined effects from picture category and location RSA analyses (see Results, Pattern similarity of picture categories and locations; Fig. 4D–F, Table 3).

trials) reflected correct guesses (forgotten \cap forgotten-remembered: 61.3 \pm 5.5 trials; 61/4 locations, chance level, 15; $p = 0.423$) and were grouped together with associations that were forgotten at both tests (day 1 and 3). Subjects only displayed very few trials with no responses (“misses”; 3 \pm 1 trials across both days), which were excluded from behavioral and multivariate fMRI data analyses. For univariate fMRI analysis, missed trials were collapsed together with forgotten associations. To test whether memory performance (weak, durable) was significantly above chance level that could be reached by guessing, we applied one-sample t tests. Chance level was calculated based on the average number of remembered associations (weak \cap durable: 127.6 \pm 5.7 trials; 128/4 locations, chance level, 32). α was set to 0.05 throughout.

fMRI data acquisition. Imaging data were acquired using a 3.0 tesla MRI scanner (Skyra, Siemens) equipped with a 32-channel head coil. We obtained 405 T_2^* -weighted blood oxygenation level-dependent (BOLD) images for each study and immediate test phase, using a gradient multi-echo EPI sequence. The application of multiple echo times (TEs) was shown to increase the signal-to-noise ratio because it allows region-specific TEs (Poser et al., 2006). For instance, signal from the MTL and the ventromedial prefrontal cortex benefits from shorter TEs, given the neighboring air-filled cavities. Signals from other brain regions, like areas at the convexity, yields an optimal BOLD contrast at longer TEs. Parameters were as follows: repetition time (TR), 2180 ms; TEs, 7.5, 18.3, 29, 40 ms; flip angle, 90°; field of view (FOV), 224 \times 224 mm; matrix, 74 \times 74; 34 ascending axial slices; 21% slice gap; voxel size, 3 mm. Structural scans were acquired using a magnetization-prepared rapid acquisition gradient echo sequence with the following parameters: TR, 2300 ms; TE, 3.03 ms; flip angle, 8°; FOV, 256 \times 256 mm; voxel size, 1 mm isotropic.

fMRI data preprocessing. All imaging data were analyzed using SPM8 (<http://www.fil.ion.ucl.ac.uk/spm/>) in combination with Matlab (MathWorks). As a first step, echoes from the four different echo times were combined into single volumes. We used 56 scans that were acquired during a short resting-state scan (2 min) before the start of the first study phase to determine the optimal weighting of echo times for each voxel. This was performed by calculating the contrast-to-noise ratio for each echo per scan. Images from multiple echo times were then combined by performing motion correction on the first echo, estimating iterative rigid-body realignment to minimize the residual sum of squares between the first echo of the first scan and all remaining scans. The estimated parameters were then applied to all other echoes, realigning all echoes to the first echo of the first scan. Finally, the calculated optimal echo time weightings were used to combine the four echo images into a single

image. These combined images were used for all further preprocessing and analyses.

The first six volumes were discarded to allow for T1 equilibration. The combined EPI volumes were then slice time corrected to the middle slice and realigned to the mean image of both runs. The structural scan was coregistered to the mean functional scan and segmented into gray matter, white matter, and cerebrospinal fluid using the "New Segmentation" algorithm. Multivariate RSA was performed in each subjects' native space. For univariate analyses, all images (functional and structural) were spatially normalized to the Montreal Neurological Institute (MNI) EPI template using Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL; Ashburner, 2007), and functional images were further smoothed with a 3D Gaussian kernel (8 mm full-width at half maximum, FWHM).

Univariate activation analysis. To investigate subsequent memory effects during encoding, all trials were sorted based on individual memory performance (forgotten, weak, durable; see above). The BOLD response for all trials was modeled with separate task regressors time-locked to the onset of the trials. Trials that were forgotten at the immediate test (day 1) but were correctly recalled at the delayed test (day 3), as well as missed responses were included in the task regressor for forgotten trials (see above). All events were estimated as a boxcar function with the duration of one trial (3.4 s) and were convolved with a canonical hemodynamic response function. In addition, the six realignment parameters, their first derivatives, and the squared first derivatives were included in the design matrix. This resulted in 18 additional regressors that accounted for noise due to head movement. Finally, a high-pass filter with a cutoff at 128 s was applied. Both runs were combined in a first-level model, and task regressors were contrasted against the implicit baseline.

For group analysis, contrast images were entered into a second-level random-effects one-way ANOVA with memory durability (forgotten, weak, durable) as a within-subject factor. Conditions were compared using *post hoc* paired-sample *t* tests. Unless stated otherwise, activation was tested for significance using cluster inference with a cluster-defining threshold of $p < 0.001$ and a cluster probability of $p < 0.05$ familywise error (FWE) corrected for multiple comparisons. For all analyses, the corrected cluster size threshold (i.e., the spatial extent of a cluster that is required to be labeled significant) was calculated using the SPM extension "CorrClusTh.m," together with the Newton–Raphson search method (script provided by Thomas Nichols, University of Warwick, United Kingdom, and Marko Wilke, University of Tübingen, Germany; <http://www2.warwick.ac.uk/fac/sci/statistics/staff/academic-research/nichols/scripts/spm/>).

Representational similarity analysis. RSA (Kriegeskorte et al., 2008) was used to determine the neural pattern similarity across the unique picture–location associations during encoding. Single-trial estimates were obtained by modeling each encoding trial with a separate regressor (Mumford et al., 2012). Remaining trials and nuisance regressors were appended identically to the univariate analysis (see above), but runs were modeled independently. This resulted in approximately 192 beta images per subject (actual numbers varied slightly since missed trials were excluded from the estimation). RSA was performed on unsmoothed data and within the native space of each subject.

For all RSA analyses, we moved a spherical searchlight (Kriegeskorte et al., 2006) with a radius of 8 mm (73 voxels) throughout the brain volume. Only searchlights that contained at least 30 gray matter voxels were considered. Single-trial beta estimates from voxels within a given searchlight were extracted and reshaped into a trial \times voxel matrix, whereby trials were sorted according to their memory durability (Fig. 1D; forgotten, weak, durable). Data were z-scored across trials and runs to remove mean activation differences, and the beta estimate of each trial was correlated with the beta estimates of all other trials, resulting in a trial \times trial similarity matrix. These data were then Fisher's z-transformed, and overall similarity scores were computed by averaging across the respective quadrants of the similarity matrix. Specifically, we were interested in the neural pattern similarities across the unique picture–location associations that shared the same memory durability (forgotten \times forgotten, weak \times weak, durable \times durable). Since nonrandom trial orders (as typically present in subsequent memory designs) can spuriously drive RSA results (i.e., trials that are closer in time tend to be more similar;

Mumford et al., 2014), we only report similarities between trials from different runs.

The overall similarity values were assigned to the center voxel of the respective searchlight, resulting in a 3D whole-brain similarity image for forgotten, weak, and durable associations. Images were normalized using DARTEL and were smoothed with a 3 mm FWHM Gaussian kernel. To test significance on a group level, we submitted the images to a second-level random-effects one-way ANOVA with memory durability (forgotten, weak, durable) as a within-subject factor. Conditions were compared using *post hoc* paired-sample *t* tests. Unless stated otherwise, we applied cluster inference with a cluster-defining threshold of $p < 0.001$ and a cluster probability of $p < 0.05$, FWE-corrected for multiple comparisons. We further tested similarities within a priori defined hippocampal and parahippocampal regions of interest (ROIs; left and right; based on the Automatic Anatomical Labeling atlas, Tzourio-Mazoyer et al., 2002), using small volume correction (SVC; $p < 0.05$, FWE-corrected at cluster level).

Activation and pattern similarity. Levels of activation might drive pattern similarity such that, for example, increased activation could lead to stronger pattern similarity due to a stronger signal. We therefore investigated the relationship between these two measures. To this end, we created a spherical ROI around the peak coordinate located within the PCC, derived from the RSA contrast durable $>$ forgotten (Fig. 3A; MNI coordinates: $x = -3$, $y = -48$, $z = 33$; radius, 8 mm; see Results, Pattern similarity during encoding). We then extracted parameter estimates from the univariate activation analysis, as well as pattern similarity scores from searchlight images within this ROI. The relationship of these two measures was tested using across-subject correlations (Pearson's r ; Bonferroni corrected for multiple comparisons). We identified three outliers (mean \pm 3 SDs) across the three different conditions, in three different subjects (Fig. 4B and C, marked in red). Analysis was performed both including and excluding these values.

Picture categories and locations. We ran additional analyses to test whether specific perceptual features of the stimulus material (picture categories and locations) rather than mnemonic processing during encoding might drive the reported pattern similarity effects (Results, Pattern similarity during encoding). First, we grouped picture–location associations based on their picture category and computed the single-trial similarities of animals \times plants ("animate") and objects \times buildings ("inanimate"). Each run contained pictures of one animate and one inanimate category (see Materials and methods, Stimulus material and randomization; and Fig. 1E). Thus, we again considered only similarities between runs. Searchlight maps for animate and inanimate pattern similarities were postprocessed as described above and were compared with a paired-sample *t* test.

Second, we grouped encoding trials based on their locations (Fig. 1F) and calculated the similarities across trials with the same location (lower left \times lower left, upper left \times upper left, lower right \times lower right, upper right \times upper right), as well as between locations (all six possible combinations, not listed here). Again, we tested only similarities between runs. We then calculated average within- and between-location pattern similarity searchlight images and tested them with a paired-sample *t* test. We reasoned that if specific neuronal representation for each location were held by, for example, occipital regions, this should be reflected in stronger within- than between-location similarities. Further, we examined representations for the different locations during the immediate test (day 1). Single trials were modeled from trial onset until a button press occurred. Further analysis was conducted as described above.

Activation and pattern similarity during memory retrieval. To investigate whether our subsequent memory effects were spatially overlapping with processes occurring at retrieval, we also analyzed the fMRI data obtained during the immediate test (day 1) in a separate analysis step. For univariate activation analysis, retrieval data were modeled from trial onset until a button press occurred (thus, the duration was equal to the reaction time) but was otherwise performed identically to the analysis of the encoding data (see Materials and methods, Univariate activation analysis). Again, unless stated otherwise, group effects were tested using a second-level random-effects one-way ANOVA with memory durability (forgotten, weak, durable) as a within-subject factor, and *post hoc* paired-sample *t* tests. Also, for RSA, single trials were

modeled from trial onset until a button press occurred. The remaining analysis was identical to the RSA for the encoding data (see Materials and methods, Representational similarity analysis).

Results

Memory performance

Memory performance for both weak and durable associations was significantly above chance level (number of associations, mean \pm SEM: weak: 41.4 ± 2.2 , $t_{(33)} = 4.3$, $p < 0.0005$; durable: 86.2 ± 6.5 , $t_{(33)} = 8.3$, $p < 0.0005$; chance level, 32). Approximately one-third of the associations were forgotten (61.3 ± 5.5).

Activation during encoding

Next, we turned to the fMRI data and investigated brain activation during the encoding of durable relative to forgotten associations. This contrast showed increased activation in bilateral hippocampus, parahippocampal cortex, ventromedial and lateral prefrontal cortex, as well as bilateral fusiform gyrus, angular gyrus, and occipital cortex (durable > forgotten; Fig. 2A, Table 1; see Table 1 also for main effect of memory durability). A comparison of durable and weak associations exhibited a similar activation profile (durable > weak; Fig. 2B, Table 1). No region, however, showed more activation for weak relative to forgotten associations during encoding (weak > forgotten). Thus, as expected, the formation of durable memories (as opposed to weak or no memory formation) involved activation increases within the MTL and the prefrontal cortex, as well as inferior temporal, parietal, and occipital regions.

Pattern similarity during encoding

We hypothesized that durable memory formation, relative to weak or no memory formation, would depend on increased pattern similarity of unique picture–location associations in the hippocampus, surrounding MTL structures, LPFC, and the PCC. To test this, we performed single-trial RSA using a whole-brain searchlight approach (see Materials and methods, Representational similarity analysis).

Results revealed increased pattern similarity for durable relative to forgotten associations in the PCC, medial and lateral prefrontal cortex, fusiform gyrus, and angular gyrus (durable > forgotten; Fig. 3A, Table 2; see Table 2 also for main effect of memory durability). Moreover, we found increased pattern similarity in the left parahippocampal cortex when applying SVC ($p < 0.05$, FWE-corrected at cluster level; peak MNI coordinates of local maximum: $x = -27$, $y = -33$, $z = -18$; z -value, 3.3, 3 voxels; Fig. 3A). We did not find any significant effects in the hippocampus.

Relative to the weak associations, we found increased pattern similarity in the right precentral gyrus (durable > weak; Fig. 3B,

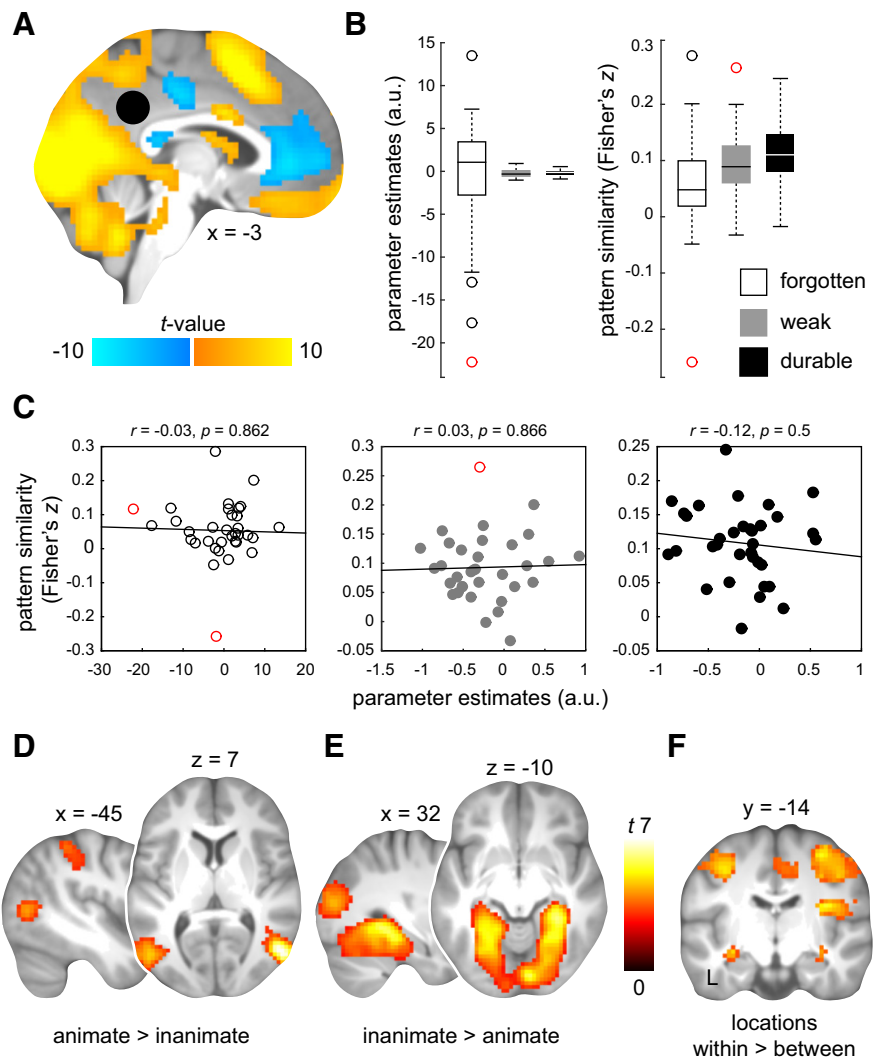


Figure 4. Activation and deactivation, pattern similarity, and perceptual features. **A**, General activation (warm colors) and deactivation (cool colors) during encoding, compared to a fixation baseline. The ROI in the PCC was based on the contrast durable > forgotten (Fig. 3A; 73 voxels, marked in black) and partly overlapped with deactivations (26% of all PCC voxels; not visible in the figure). **B**, Box plots show parameter estimates (arbitrary units, a.u.) and pattern similarity (Fisher's z) per memory durability condition, extracted from the PCC ROI. **C**, Correlations (Pearson's r) between PCC parameter estimates and pattern similarity, including the three outliers (marked in red; mean \pm 3 SDs; see also **B**). Results remained stable when outliers were removed (forgotten: $r = 0.64$, $p = 0.734$; weak: $r = 0.099$, $p = 0.597$; durable: $r = -0.166$, $p = 0.373$). **D**, **E**, RSA of picture categories: increased pattern similarities for the contrasts animate > inanimate (**D**) and inanimate > animate (**E**). **F**, Location RSA for retrieval data (immediate test, day 1; for encoding results, please see Table 3). All results are shown at $p < 0.001$ ($p < 0.05$, FWE-corrected at cluster level; Table 3). L, Left.

Table 2), possibly reflecting the association with a later motor response (note that subjects did not provide button presses during encoding). Finally, we compared the pattern similarities of weak and forgotten associations. The only region that showed increased pattern similarity during weak memory encoding was the PCC (weak > forgotten; Fig. 3C, Table 2).

In summary, results showed that the multivoxel patterns related to durable memory formation were more similar than later forgotten material within neocortical regions, including the PCC, medial and lateral prefrontal cortex, fusiform gyrus, parahippocampal cortex, and angular gyrus. These findings are largely different from results of the univariate activation analysis (Fig. 3D, both results are overlaid). Notably, the PCC showed increased pattern similarity during the encoding of both weak and durable associations, suggesting that this region promotes mem-

ory formation during encoding, irrespective of the eventual durability of a memory trace.

Activation and pattern similarity in the PCC

The results above suggest that the PCC facilitates memory formation through enhanced pattern similarity, regardless of the subsequent memory durability. Pattern similarity, however, might be increased together with univariate activation levels, assuming that enhanced (or decreased) activation amplifies the measured signal. To test this, we examined the relationship of general activations and deactivations with the reported pattern similarity during durable relative to forgotten memory encoding (durable > forgotten). We focused on the peak coordinate that was located within the PCC (see above; MNI coordinates: $x = -3$, $y = -48$, $z = 33$; see Materials and methods, Representational similarity analysis, Activation and pattern similarity). The PCC ROI did not overlap with regions that showed a general increase during encoding, and only partly overlapped with voxels that showed decreased activation compared with the fixation baseline (Fig. 4A). Activation and similarity signals extracted from the PCC ROI (Fig. 4B) were not significantly correlated, and these results remained stable after the exclusion of three outliers (Fig. 4C). Therefore, we argue that pattern similarity (in the PCC) reflects a measure of processing during encoding that is complementary to processes revealed by univariate analyses.

Pattern similarity of picture categories and locations

Potentially, perceptual features of the stimulus material rather than mnemonic processing might drive the reported pattern similarity effects. If this was the case, brain regions that coded for specific picture categories or locations should spatially overlap with regions that showed increased similarity for associations with the same memory durability. Therefore, we repeated the searchlight RSA procedure but grouped picture–location associations with respect to (1) their picture categories and (2) their locations in order to identify the specific representations (see Materials and methods, Representational similarity analysis and Picture categories and locations; Fig. 1E,F).

Results of the “picture category” RSA revealed that animate stimuli, compared to inanimate stimuli, showed stronger pattern similarity within left and right lateral occipital cortices, and left superior parietal cortex during encoding (animate > inanimate; Fig. 4D, Table 3). Inanimate stimuli elicited increased pattern similarity in bilateral fusiform gyrus and parahippocampal gyrus, extending into the posterior hippocampus and occipital cortex, compared with animate stimuli (inanimate > animate; Fig. 4E, Table 3). The “location” RSA showed that the four different locations were represented in occipital cortex during encoding (Table 3).

Above, we demonstrated increased similarity for durable associations in the right precentral gyrus (durable > weak; Fig. 3B, Table 2) and speculated that this might reflect the association with a later button press (button presses were only required during the test phases, not during encoding). To see whether this was the case, we repeated the location RSA for fMRI data from the immediate test (day 1). Pictures that were associated with the same location were more similar than pictures associated with different locations in bilateral precentral gyrus and the hippocampus during retrieval (Fig. 4F, Table 3).

In conclusion, pattern similarity effects in the posterior MTL, inferior temporal, and occipital regions were likely driven by perceptual features of the different picture categories. Additionally, results indicated that increased pattern similarity in the precentral

Table 3. Pattern similarity of picture categories and locations

| Contrast and brain region | MNI coordinates | | | z-value | Cluster size |
|--|-----------------|-----|-----|---------|--------------|
| | x | y | z | | |
| Picture category RSA: animate > inanimate | | | | | |
| R middle occipital gyrus | 57 | −69 | 3 | 5.86 | 395 |
| L middle temporal gyrus | −51 | −69 | 6 | 4.47 | 249 |
| L superior parietal gyrus | −39 | −42 | 51 | 4.40 | 205 |
| Picture category RSA: inanimate > animate | | | | | |
| R fusiform gyrus | 24 | −36 | −21 | 6.17 | 3631 |
| Location RSA, encoding: | | | | | |
| within > between similarity | | | | | |
| R cuneus | 6 | −78 | 12 | 7.10 | 2680 |
| Location RSA, retrieval (immediate test, day 1): | | | | | |
| within > between similarity | | | | | |
| R postcentral gyrus | 36 | −30 | 42 | 5.22 | 801 |
| L postcentral gyrus | −36 | −14 | 48 | 4.66 | 755 |
| R supramarginal gyrus | 36 | −15 | 21 | 4.63 | 164 |
| R hippocampus | 33 | −18 | −12 | 4.17 | 116 |
| Cerebellum | −12 | −54 | −24 | 3.89 | 193 |
| R superior frontal gyrus | 9 | −21 | 45 | 3.72 | 143 |

MNI coordinates represent the location of peak voxels. We report the first local maximum within each cluster. Effects were tested for significance using cluster inference with a cluster-defining threshold of $p < 0.001$ and a cluster probability of $p < 0.05$ FWE-corrected for multiple comparisons (critical cluster sizes: picture categories, 69 voxels; locations during encoding, 100 voxels; locations during retrieval, 90 voxels). L, Left; R, right.

tral gyrus (durable > weak; Fig. 3B) reflected the association with a motor response that was required only during later retrieval. To identify brain regions that facilitated memory durability but that were not solely driven by the category- or location-specific features of the task, we plotted our RSA effects for durable > forgotten again, but exclusively masked them with the results of both the picture category and location RSA. Results suggest that the PCC, as well as medial and lateral prefrontal cortex (Table 2) promoted memory durability at encoding rather than solely representing the specific perceptual features of the task.

Memory retrieval

It is possible that a comparable pattern similarity profile is also present during retrieval. In this case, our effects would not be specific for (durable) memory formation at encoding. Therefore, we investigated pattern similarity during the immediate test (day 1) where subjects were asked to retrieve previously studied picture–location associations (Fig. 1C; see also Materials and methods, Activation and pattern similarity during retrieval). Additionally, we analyzed univariate activation to test how activity during retrieval varied with memory durability.

Reaction times for incorrect responses (forgotten associations, mean \pm SEM, 1739.6 ± 58 s) were significantly prolonged compared to correctly retrieved associations that were weak (1504.4 ± 49.7 s) or durable (1291.5 ± 44 s; main effect of memory durability: $F_{(1,4,47,2)} = 66.7$, $p < 0.0005$; *post hoc* paired-sample *t* tests: forgotten vs weak, $t_{(33)} = 6.2$, $p < 0.0005$; forgotten vs durable, $t_{(33)} = 9.2$, $p < 0.0005$; weak vs durable, $t_{(33)} = 7.9$, $p < 0.0005$).

In terms of pattern similarity, we did not find any significant increases during retrieval, also not at a more lenient threshold ($p < 0.005$, uncorrected; no positive effect of memory durability: contrasts durable > forgotten, durable > weak, or weak > forgotten; for main effect, see Table 4). We next tested univariate levels of activation during retrieval and found increased activity in the hippocampus and surrounding MTL, medial and lateral prefrontal cortex, fusiform gyrus, angular gyrus, and PCC during the retrieval of durable compared to forgotten (incorrect) associations (durable > forgotten; Fig. 5A, Table 4, see Table 4 also for main effect of memory durability). Furthermore, durable

Table 4. Activation and pattern similarity during memory retrieval

| Contrast and brain region | MNI coordinates | | | z-value | Cluster size |
|---------------------------------|-----------------|-----|-----|---------|--------------|
| | x | y | z | | |
| Activation, main effect* | | | | | |
| L superior frontal gyrus | −3 | 15 | 48 | | 42,614 |
| Activation, durable > forgotten | | | | | |
| R putamen | 12 | 9 | −12 | 6.25 | 17,018 |
| L middle orbitofrontal gyrus | −24 | 36 | −12 | 4.35 | 80 |
| R inferior frontal gyrus | 57 | 36 | −3 | 4.23 | 69 |
| Activation, durable > weak | | | | | |
| R middle occipital gyrus | 18 | −96 | 0 | 4.78 | 1358 |
| L insular cortex | −36 | −9 | 6 | 4.40 | 77 |
| L fusiform gyrus | −30 | −54 | −12 | 4.31 | 1425 |
| L posterior cingulate cortex | −3 | −33 | 30 | 4.09 | 134 |
| R superior parietal gyrus | 33 | −51 | 33 | 4.08 | 214 |
| Thalamus | −9 | −12 | 3 | 4.07 | 80 |
| Precuneus | −3 | −63 | 30 | 3.85 | 69 |
| Activation, weak > forgotten | | | | | |
| L putamen | −15 | 6 | −15 | 6.53 | 3119 |
| R putamen | 15 | 9 | −12 | 5.91 | 3533 |
| Cerebellum | −27 | −78 | −39 | 5.61 | 1762 |
| R inferior frontal gyrus | 57 | 36 | −3 | 5.26 | 87 |
| RSA, main effect* | | | | | |
| L inferior occipital gyrus | −27 | −84 | −6 | | 49,458 |

MNI coordinates represent the location of peak voxels. We report the first local maximum within each cluster. Effects were tested for significance using cluster inference with a cluster-defining threshold of $p < 0.001$ and a cluster probability of $p < 0.05$ FWE-corrected for multiple comparisons (critical cluster size: 65 voxels). L, Left; R, right.

*Results for main effects were thresholded at $p < 0.05$ FWE-corrected for multiple comparisons.

memories showed stronger activation in the fusiform gyrus, occipital regions, and PCC relative to weak associations (durable > weak; Fig. 5B, Table 4). Lastly, weak associations were related to increased activation in the hippocampus and surrounding MTL, medial prefrontal cortex, fusiform gyrus, and in the PCC during retrieval (weak > forgotten; Fig. 5C, Table 4), and this was comparable to the contrast durable > forgotten.

In summary, we did not find any increases in pattern similarity for durable or weak memories during retrieval, suggesting that the observed pattern similarity effects during encoding (see above) were specific for memory formation. In terms of univariate levels of activation, we found stronger activity during correct (durable and weak) retrieval in the MTL, medial prefrontal cortex, angular gyrus, and in the PCC. Durable relative to weak memories were characterized by additional retrieval-related activation increases in the fusiform gyrus and in the PCC. Interestingly, these regions largely overlapped with the regions that showed increased pattern similarity during encoding (Fig. 5D).

Discussion

Here, we investigated the neuronal correlates of durable memory formation at encoding. We found that the PCC processed both durable and weak associations similarly, regardless of the eventual durability. Durable memories, as opposed to weak or forgotten material, were associated with increased activation in medial temporal and prefrontal regions. This was paralleled by heightened pattern similarity in prefrontal and parietal regions, but also anterior and posterior midline structures—regions that were engaged during subsequent retrieval (Fig. 6). Furthermore, pattern similarity effects were not influenced by general levels of activation and deactivation, category-, or location-specific features of the task, and were specific to memory encoding, not memory retrieval. Altogether, these results corroborate our hypothesis that consistent processing at encoding underlies associative memory formation.

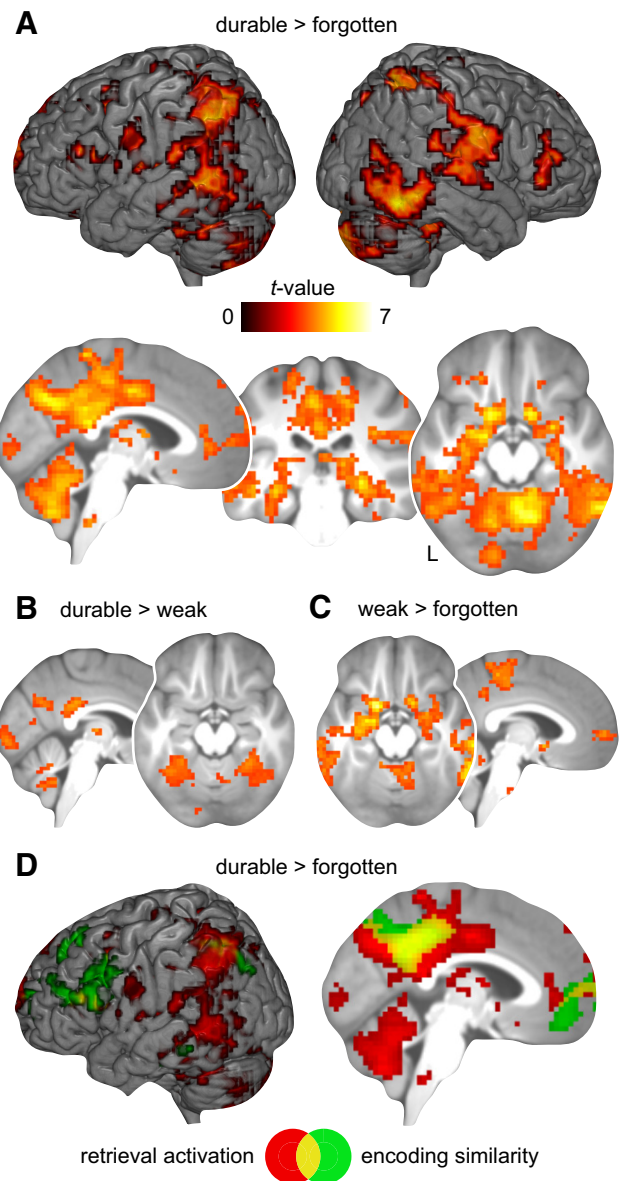


Figure 5. Activation during memory retrieval, and overlap with pattern similarity during encoding. **A–C**, Increased BOLD activation during retrieval (immediate test, day 1) of durable > forgotten (**A**), durable > weak (**B**), and weak > forgotten (**C**) associations. Results are shown at $p < 0.001$ ($p < 0.05$, FWE-corrected at cluster level; see also Table 4), and at MNI coordinates: $x = -3$, $y = -27$, $z = -17$. **D**, Overlay of univariate activation during retrieval (durable > forgotten; **A**) and pattern similarity during encoding (same contrast; Fig. 3A). L, Left.

We hypothesized that the encoding of durable associations compared to weak and forgotten associations would be related to enhanced activation in the MTL and LPFC. Indeed, we observed increased activation in the hippocampus and surrounding MTL, inferior temporal regions, as well as in prefrontal and parietal cortex (durable > forgotten, similar to durable > weak; Fig. 2). This is partly converging with previous reports. Uncapher and Rugg (2005), for example, found increased LPFC activation during durable memory formation. In their study, however, weak memory formation was associated with stronger activation in the fusiform gyrus. This difference in findings is likely due to the visual nature of our design, leading to extensive fusiform and visual cortex activation at durable memory encoding compared with the semantic stimulus material used by Uncapher and Rugg

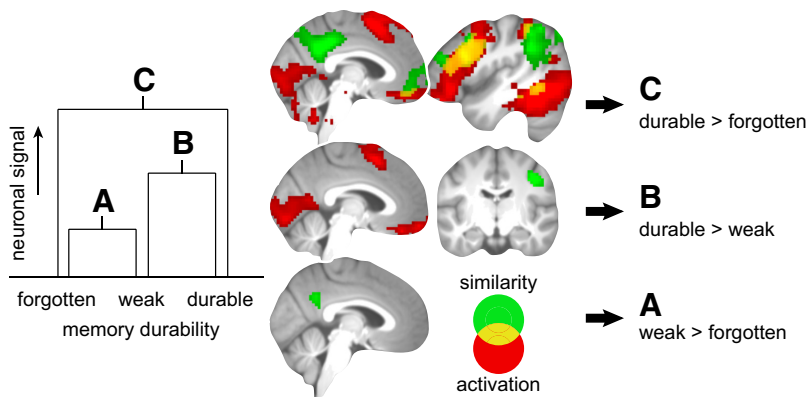


Figure 6. Durable memory encoding. Left, Overview of the different contrasts between memory durability conditions. The y-axis represents the “neuronal signal” (i.e., univariate levels of activation and pattern similarity). Right, Pattern similarity (green) in the PCC is required for a memory to be formed; contrasts weak > forgotten (A) and durable > forgotten (C). For durable > weak memory formation (B), widespread activation increases (red) and pattern similarity in the precentral gyrus appear necessary. The contrast durable > forgotten (C) yielded the largest increase in activation and pattern similarity. While activation was enhanced in regions typically associated with the encoding of information, pattern similarity was increased in regions that were engaged during later retrieval (see also Fig. 5), and signals were partly overlapping (yellow) in prefrontal, parietal, and inferior temporal regions. These results were previously presented in Figures 2 and 3.

(2005). Importantly, we did not find activation differences between weak and forgotten associations during encoding. Previous studies that used a single memory test, often administered shortly following study, were unable to separate memories with different durability. Therefore, a contrast of remembered over forgotten material likely included both durable and weak memories and produced the typical subsequent memory effects (Spaniol et al., 2009; Kim, 2011). Here, the comparison of weak and forgotten associations clearly excluded durable memories, which could explain our null result. Similarly, Carr et al. (2010) tested all stimulus material immediately and at a delayed test. The authors reported increased MTL activation only for durable recollection. Hence, increased hippocampal–neocortical activation appears necessary for durable memory formation, whereby weak memories might rely on other factors during encoding, such as consistent processing.

Next, we hypothesized that consistent processing at encoding would mirror the buildup of associative memory traces. We quantified consistency as pattern similarity of picture–location associations that shared the same memory durability, and expected increased pattern similarity for durable associations relative to weak and forgotten associations within the MTL, LPFC, and PCC. Results revealed heightened pattern similarity in the parahippocampal cortex, medial and lateral prefrontal cortices, inferior temporal regions, angular gyrus, and in the PCC during the formation of durable memories (durable > forgotten; Fig. 3A). For durable over weak encoding, increased pattern similarity was found in the right precentral gyrus (Fig. 3B), which also coded for specific picture locations, or the button press, during later retrieval (Fig. 4F). This is likely specific to the task material at hand, and other representational regions might be involved for different stimulus material. Altogether, we suggest that durable memories entailed a specific representation of the picture–location associations already at encoding. Potentially, along with increased activation in a medial temporal–neocortical set of regions, this is what makes spatial associative memories longer lasting.

Overall pattern similarity, or global similarity, was formerly related to subsequent recognition memory (LaRocque et al., 2013), confidence and categorization (Davis et al., 2014), as well as fear memory formation (Visser et al., 2013). Furthermore, Qin

et al. (2014) revealed that consistent hippocampal processing indicated the transition from procedure- to memory-based problem solving in children, and recent reports showed that global pattern similarity can be influenced through attention (Aly and Turk-Browne, 2016a,b) and reward (Wolosin et al., 2013). The present findings thus corroborate our hypothesis that consistent processing mirrors associative memory formation. Moreover, the reported pattern similarity results (in the PCC) were not driven by general levels of activation or deactivation during encoding (Jimura and Poldrack, 2012; Xue et al., 2013; Bird et al., 2015), and did not merely represent the specific picture categories or locations (Fig. 4).

Regions that showed stronger pattern similarity were largely distinct from areas associated with increased activation during encoding, but both measures—activation and pattern similarity—spatially overlapped in frontoparietal regions. This included the lateral prefrontal and posterior parietal cortex. While the LPFC coordinates the selection and organization of memories via top-down projections (Blumenfeld and Ranganath, 2007), the posterior parietal cortex is thought to regulate attention to mnemonic content (Cabeza et al., 2008; Ciaramelli et al., 2008). Interactions between the two regions are considered relevant for the modulation of pattern similarity during encoding. Recently, frontoparietal activity (Xue et al., 2013) and LPFC stimulation (Lu et al., 2015) were directly associated with increased pattern similarity and its beneficial effects on memory formation. Moreover, these regions were speculated to suppress irrelevant information processing by coordinating spatial attention (Kastner et al., 1998; Lu et al., 2011; Rabinowitz et al., 2015) and by guiding task relevance (Jehee et al., 2011; Poort et al., 2015). Although suppression has so far mainly been studied in the visual system, similar mechanisms might operate also here.

The PCC showed increased pattern similarity during successful memory formation, independent of the eventual memory durability (Fig. 3). Additionally, brain regions with enhanced pattern similarity at encoding overlapped with regions that showed activation increases during later retrieval (Fig. 5). Akin to our results, pattern similarity in the PCC was shown to promote memory during encoding (Xue et al., 2013) and consolidation (Bird et al., 2015). Typically, the PCC is regarded as central for memory retrieval (Rugg and Vilberg, 2013; Watrous et al., 2013; King et al., 2015; Wagner et al., 2015), and pattern similarity during encoding might thus represent retrieval processes to flexibly embed novel associations into pre-existing memories (Maguire et al., 1999; Bird et al., 2015). For instance, when studying picture–location associations, one may create stories with pictures that belong to the same location (indeed, some subjects reported the use of this strategy). The encoding of a novel picture–location association might thus trigger the retrieval of other pictures associated with the same location. Furthermore, our pattern similarity effects were present only during encoding, not during retrieval; hence, we concluded that consistent processing is necessary for associative memory formation. Accordingly, a recent study reported greater dissimilarity of associations during retrieval (Karlsson Wirebring et al., 2015). Better memory might

thus depend on consistent processing (similarity) at encoding, but on differentiated representations (dissimilarity) during retrieval. This could be seen as an “encoding/retrieval-flip” in the pattern similarity domain (in contrast to univariate activation; Daselaar et al., 2004, 2009; Huijbers et al., 2012) and should be the subject of further research.

We found encoding-related increases in pattern similarity for durable over weak memories in the parahippocampal cortex, but not in the hippocampus (Fig. 3A). The formation of associations, especially of a spatial nature, is a central feature of hippocampal function (Brown and Aggleton, 2001; Burgess et al., 2001; Sperling et al., 2003; Kirwan and Stark, 2004; Mayes et al., 2007; Staerens and Davachi, 2008), possibly achieved by pattern separation, or the orthogonalization of memory representations (Yassa and Stark, 2011). Specifically, LaRocque et al. (2013) reported that successful recognition memory was predicted by dissimilar hippocampal representations at encoding (pattern separation), while representations in surrounding MTL regions were found to be more similar (pattern completion; LaRocque et al., 2013). The latter finding resonates with our results.

Studies that have investigated memory durability (Uncapher and Rugg, 2005; Carr et al., 2010; Liu et al., 2013; Sneve et al., 2015) used the so-called “remember/know procedure” (Tulving, 1985; Yonelinas, 2002), since recollection (remember response) and familiarity (know response) display differential decay (Yonelinas and Levy, 2002). In contrast with this, we aimed at targeting associative, recollection-based memory rather than familiarity-based recognition. Furthermore, to preclude testing effects that could boost weak memories in their durability (Karpicke and Roediger, 2008; Carpenter, 2009; Pyc and Rawson, 2009; Roediger and Butler, 2011), most studies tested half of the stimulus material at a first test and the second half later (Uncapher and Rugg, 2005; Sneve et al., 2015). This comes at the expense of being able to explicitly define memory durability, as it is unclear whether material remembered at a first test (and defined as weak) will be forgotten at a second test. To our knowledge, only one study so far has tested all stimulus material twice (Carr et al., 2010; but see also Liu et al., 2013). We favored this latter approach since it allowed us to link encoding-related brain activity (patterns) with the prospective memory durability of each unique picture–location association, as determined from both tests. We acknowledge that retesting could have boosted some of the weak memories so that they were falsely classified as durable. This shortcoming could have affected only the trial sorting for the subsequent memory analysis, diminishing differences between durable compared to weak memory encoding. However, despite this potential limitation, we found reliable activation and pattern similarity effects for durable over weak memories at encoding (Figs. 2B, 3B).

Initially, we expected enhanced activation and pattern similarity in the hippocampus and in specific neocortical regions for durable relative to weak or no memory formation (durable > weak > forgotten). Our results only partly confirmed this hypothesis of a stepwise, linear relationship, and indicated that activation and pattern similarity did not follow the same model of signal increase with increasing memory durability. First, we suggest that pattern similarity in the PCC might represent an initial “threshold” for associative memory formation that needs to be overcome. Second, pattern similarity appears to linearly increase with memory durability, while activation seems mostly relevant for the formation of durable memories (Fig. 6). Thereafter, successful encoding is likely followed by synaptic (Redondo and Morris, 2011) and systems consolidation (Frankland and Bontempi, 2005), and we speculate that these consolidation processes are generally more pronounced for durable than for weak (or forgotten) associations. However, also in line with the testing effect (Karpicke and Roediger, 2008; Roediger and Butler,

2011), another possibility is that initially weak memories, compared with strong associative material, may benefit more from the additional effort and elaboration required to successfully retrieve those memories during the initial test (Carpenter, 2009; Pyc and Rawson, 2009), which could potentially promote them in their durability (see above).

To conclude, we showed that pattern similarity, or consistent processing, in the PCC predicts the formation of spatial associative memories at encoding. If this is augmented by additional activation increases in regions typically related to encoding, and by consistent processing in regions involved in later retrieval, formed memories appear durable for at least 48 h. Thus, consistent processing underlies the emergence of associative memory traces.

References

- Aly M, Turk-Browne NB (2016a) Attention stabilizes representations in the human hippocampus. *Cereb Cortex* 26:783–796. [CrossRef Medline](#)
- Aly M, Turk-Browne NB (2016b) Attention promotes episodic encoding by stabilizing hippocampal representations. *Proc Natl Acad Sci U S A* 113: E420–E429. [CrossRef Medline](#)
- Ashburner J (2007) A fast diffeomorphic image registration algorithm. *Neuroimage* 38:95–113. [CrossRef Medline](#)
- Bird CM, Keidel JL, Ing LP, Horner AJ, Burgess N (2015) Consolidation of complex events via reinstatement in posterior cingulate cortex. *J Neurosci* 35:14426–14434. [CrossRef Medline](#)
- Blumenfeld RS, Ranganath C (2007) Prefrontal cortex and long-term memory encoding: an integrative review of findings from neuropsychology and neuroimaging. *Neuroscientist* 13:280–291. [CrossRef Medline](#)
- Brewer JB, Zhao Z, Desmond JE, Glover GH, Gabrieli JD (1998) Making memories: brain activity that predicts how well visual experience will be remembered. *Science* 281:1185–1187. [CrossRef Medline](#)
- Brown MW, Aggleton JP (2001) Recognition memory: what are the roles of the perirhinal cortex and hippocampus? *Nat Rev Neurosci* 2:51–61. [CrossRef Medline](#)
- Burgess N, Maguire EA, Spiers HJ, O’Keefe J (2001) A temporoparietal and prefrontal network for retrieving the spatial context of lifelike events. *Neuroimage* 14:439–453. [CrossRef Medline](#)
- Cabeza R, Ciaramelli E, Olson IR, Moscovitch M (2008) The parietal cortex and episodic memory: an attentional account. *Nat Rev Neurosci* 9: 613–625. [CrossRef Medline](#)
- Carpenter SK (2009) Cue strength as a moderator of the testing effect: the benefits of elaborative retrieval. *J Exp Psychol Learn Mem Cogn* 35: 1563–1569. [CrossRef Medline](#)
- Carr VA, Viskontas IV, Engel SA, Knowlton BJ (2010) Neural activity in the hippocampus and perirhinal cortex during encoding is associated with the durability of episodic memory. *J Cogn Neurosci* 22:2652–2662. [CrossRef Medline](#)
- Ciaramelli E, Grady CL, Moscovitch M (2008) Top-down and bottom-up attention to memory: a hypothesis (AtoM) on the role of the posterior parietal cortex in memory retrieval. *Neuropsychologia* 46:1828–1851. [CrossRef Medline](#)
- Daselaar SM, Prince SE, Cabeza R (2004) When less means more: deactivations during encoding that predict subsequent memory. *Neuroimage* 23: 921–927. [CrossRef Medline](#)
- Daselaar SM, Prince SE, Dennis NA, Hayes SM, Kim H, Cabeza R (2009) Posterior midline and ventral parietal activity is associated with retrieval success and encoding failure. *Front Hum Neurosci* 3:13. [CrossRef Medline](#)
- Davis T, Xue G, Love BC, Preston AR, Poldrack RA (2014) Global neural pattern similarity as a common basis for categorization and recognition memory. *J Neurosci* 34:7472–7484. [CrossRef Medline](#)
- Fernández G, Effern A, Grunwald T, Pezer N, Lehnertz K, Dümpelmann M, Van Roost D, Elger CE (1999) Real-time tracking of memory formation in the human rhinal cortex and hippocampus. *Science* 285:1582–1585. [CrossRef Medline](#)
- Frankland PW, Bontempi B (2005) The organization of recent and remote memories. *Nat Rev Neurosci* 6:119–130. [CrossRef Medline](#)
- Huijbers W, Vannini P, Sperling RA, C M P, Cabeza R, Daselaar SM (2012) Explaining the encoding/retrieval flip: memory-related deactivations and activations in the posteromedial cortex. *Neuropsychologia* 50:3764–3774. [CrossRef Medline](#)

- Jehee JF, Brady DK, Tong F (2011) Attention improves encoding of task-relevant features in the human visual cortex. *J Neurosci* 31:8210–8219. [CrossRef Medline](#)
- Jimura K, Poldrack RA (2012) Analyses of regional-average activation and multivoxel pattern information tell complementary stories. *Neuropsychologia* 50:544–552. [CrossRef Medline](#)
- Karlsson Wirebring L, Wiklund-Hörnqvist C, Eriksson J, Andersson M, Jonsson B, Nyberg L (2015) Lesser neural pattern similarity across repeated tests is associated with better long-term memory retention. *J Neurosci* 35:9595–9602. [CrossRef Medline](#)
- Karpicke JD, Roediger HL 3rd (2008) The critical importance of retrieval for learning. *Science* 319:966–968. [CrossRef Medline](#)
- Kastner S, De Weerd P, Desimone R, Ungerleider LG (1998) Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. *Science* 282:108–111. [CrossRef Medline](#)
- Kim H (2011) Neural activity that predicts subsequent memory and forgetting: a meta-analysis of 74 fMRI studies. *Neuroimage* 54:2446–2461. [CrossRef Medline](#)
- King DR, de Chastelaine M, Elward RL, Wang TH, Rugg MD (2015) Recollection-related increases in functional connectivity predict individual differences in memory accuracy. *J Neurosci* 35:1763–1772. [CrossRef Medline](#)
- Kirwan CB, Stark CE (2004) Medial temporal lobe activation during encoding and retrieval of novel face-name pairs. *Hippocampus* 14:919–930. [CrossRef Medline](#)
- Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping. *Proc Natl Acad Sci U S A* 103:3863–3868. [CrossRef Medline](#)
- Kriegeskorte N, Mur M, Bandettini P (2008) Representational similarity analysis—connecting the branches of systems neuroscience. *Front Syst Neurosci* 2:4. [CrossRef Medline](#)
- LaRocque KF, Smith ME, Carr VA, Witthoft N, Grill-Spector K, Wagner AD (2013) Global similarity and pattern separation in the human medial temporal lobe predict subsequent memory. *J Neurosci* 33:5466–5474. [CrossRef Medline](#)
- Liu Q, Dong Q, Chen C, Xue G (2013) Neural processes during encoding support durable memory. *Neuroimage* 88C:1–9. [CrossRef Medline](#)
- Lu Y, Wang C, Chen C, Xue G (2015) Spatiotemporal neural pattern similarity supports episodic memory. *Curr Biol* 25:780–785. [CrossRef Medline](#)
- Lu ZL, Li X, Tjan BS, Doshier BA, Chu W (2011) Attention extracts signal in external noise: a BOLD fMRI study. *J Cogn Neurosci* 23:1148–1159. [CrossRef Medline](#)
- Maguire EA, Frith CD, Morris RG (1999) The functional neuroanatomy of comprehension and memory: the importance of prior knowledge. *Brain* 122:1839–1850. [CrossRef Medline](#)
- Mayes A, Montaldi D, Migo E (2007) Associative memory and the medial temporal lobes. *Trends Cogn Sci* 11:126–135. [CrossRef Medline](#)
- Mumford JA, Turner BO, Ashby FG, Poldrack RA (2012) Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *Neuroimage* 59:2636–2643. [CrossRef Medline](#)
- Mumford JA, Davis T, Poldrack RA (2014) The impact of study design on pattern estimation for single-trial multivariate pattern analysis. *Neuroimage* 103:130–138. [CrossRef Medline](#)
- Poort J, Khan AG, Pachitariu M, Nemri A, Orsolich I, Krupic J, Bauza M, Sahani M, Keller GB, Mrsic-Flogel TD, Hofer SB (2015) Learning enhances sensory and multiple non-sensory representations in primary visual cortex. *Neuron* 86:1478–1490. [CrossRef Medline](#)
- Poser BA, Versluis MJ, Hoogduin JM, Norris DG (2006) BOLD contrast sensitivity enhancement and artifact reduction with multiecho EPI: parallel-acquired inhomogeneity-desensitized fMRI. *Magn Reson Med* 55:1227–1235. [CrossRef Medline](#)
- Pyc MA, Rawson KA (2009) Testing the retrieval effort hypothesis: does greater difficulty correctly recalling information lead to higher levels of memory? *J Mem Lang* 60:437–447. [CrossRef](#)
- Qin S, Cho S, Chen T, Rosenberg-Lee M, Geary DC, Menon V (2014) Hippocampal-neocortical functional reorganization underlies children's cognitive development. *Nat Neurosci* 17:1263–1269.
- Rabinowitz NC, Goris RL, Cohen M, Simoncelli EP (2015) Attention stabilizes the shared gain of V4 populations. *Elife* 4:e08998. [CrossRef Medline](#)
- Redondo RL, Morris RG (2011) Making memories last: the synaptic tagging and capture hypothesis. *Nat Rev Neurosci* 12:17–30. [CrossRef Medline](#)
- Roediger HL 3rd, Butler AC (2011) The critical role of retrieval practice in long-term retention. *Trends Cogn Sci* 15:20–27. [CrossRef Medline](#)
- Rugg MD, Vilberg KL (2013) Brain networks underlying episodic memory retrieval. *Curr Opin Neurobiol* 23:255–260. [CrossRef Medline](#)
- Sneve MH, Grydeland H, Nyberg L, Bowles B, Amlien IK, Langnes E, Walhovd KB, Fjell AM (2015) Mechanisms underlying encoding of short-lived versus durable episodic memories. *J Neurosci* 35:5202–5212. [CrossRef Medline](#)
- Spaniol J, Davidson PS, Kim AS, Han H, Moscovitch M, Grady CL (2009) Event-related fMRI studies of episodic encoding and retrieval: meta-analyses using activation likelihood estimation. *Neuropsychologia* 47:1765–1779. [CrossRef Medline](#)
- Sperling R, Chua E, Cocchiarella A, Rand-Giovannetti E, Poldrack R, Schacter DL, Albert M (2003) Putting names to faces: successful encoding of associative memories activates the anterior hippocampal formation. *Neuroimage* 20:1400–1410. [CrossRef Medline](#)
- Staresina BP, Davachi L (2008) Selective and shared contributions of the hippocampus and perirhinal cortex to episodic item and associative encoding. *J Cogn Neurosci* 20:1478–1489. [CrossRef Medline](#)
- Takashima A, Nieuwenhuis IL, Jensen O, Talamini LM, Rijpkema M, Fernández G (2009) Shift from hippocampal to neocortical centered retrieval network with consolidation. *J Neurosci* 29:10087–10093. [CrossRef Medline](#)
- Tulving E (1985) Memory and consciousness. *Can Psychol Can* 26:1–12. [CrossRef](#)
- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15:273–289. [CrossRef Medline](#)
- Uncapher MR, Rugg MD (2005) Encoding and the durability of episodic memory: a functional magnetic resonance imaging study. *J Neurosci* 25:7260–7267. [CrossRef Medline](#)
- van Dongen EV, Takashima A, Barth M, Fernández G (2011) Functional connectivity during light sleep is correlated with memory performance for face-location associations. *Neuroimage* 57:262–270. [CrossRef Medline](#)
- van Dongen EV, Thielens JW, Takashima A, Barth M, Fernández G (2012) Sleep supports selective retention of associative memories based on relevance for future utilization. *PLoS One* 7:e43426. [CrossRef Medline](#)
- Visser RM, Scholte HS, Beemsterboer T, Kindt M (2013) Neural pattern similarity predicts long-term fear memory. *Nat Neurosci* 16:388–390. [CrossRef Medline](#)
- Wagner AD, Schacter DL, Rotte M, Koutstaal W, Maril A, Dale AM, Rosen BR, Buckner RL (1998) Building memories: remembering and forgetting of verbal experiences as predicted by brain activity. *Science* 281:1188–1191. [CrossRef Medline](#)
- Wagner IC, van Buuren M, Kroes MC, Gutteling TP, van der Linden M, Morris RG, Fernández G (2015) Schematic memory components converge within angular gyrus during retrieval. *Elife* 4:e09668. [CrossRef Medline](#)
- Watrous AJ, Tandon N, Conner CR, Pieters T, Ekstrom AD (2013) Frequency-specific network connectivity increases underlie accurate spatiotemporal memory retrieval. *Nat Neurosci* 16:349–356. [CrossRef Medline](#)
- Wolosin SM, Zeithamova D, Preston AR (2013) Distributed hippocampal patterns that discriminate reward context are associated with enhanced associative binding. *J Exp Psychol Gen* 142:1264–1276. [CrossRef Medline](#)
- Xue G, Dong Q, Chen C, Lu Z, Mumford JA, Poldrack RA (2010) Greater neural pattern similarity across repetitions is associated with better memory. *Science* 330:97–101. [CrossRef Medline](#)
- Xue G, Dong Q, Chen C, Lu ZL, Mumford JA, Poldrack RA (2013) Complementary role of frontoparietal activity and cortical pattern similarity in successful episodic memory encoding. *Cereb Cortex* 23:1562–1571. [CrossRef Medline](#)
- Yassa MA, Stark CE (2011) Pattern separation in the hippocampus. *Trends Neurosci* 34:515–525. [CrossRef Medline](#)
- Yonelinas AP (2002) The nature of recollection and familiarity: a review of 30 years of research. *J Mem Lang* 46:441–517. [CrossRef](#)
- Yonelinas AP, Levy BJ (2002) Dissociating familiarity from recollection in human recognition memory: different rates of forgetting over short retention intervals. *Psychon Bull Rev* 9:575–582. [CrossRef Medline](#)