Retrieval aids the creation of a generalised memory trace and strengthens episode-unique information

Ferreira, C.S.*, Charest, I., & Wimber, M.

School of Psychology and Centre for Human Brain Health

University of Birmingham, Edgbaston, B15 2TT, United Kingdom

E-mails:

Catarina S. Ferreira: a.c.sanchesferreira@bham.ac.uk

Ian Charest: i.charest@bham.ac.uk

Maria Wimber: m.wimber@bham.ac.uk

***Corresponding Author:** Correspondence concerning this article should be addressed to Catarina S. Ferreira, School of Psychology, University of Birmingham, Edgbaston, Birmingham, B15 2TT (United Kingdom). E-mail: a.c.sanchesferreira@bham.ac.uk

Declarations of interest: None.

Abstract

Generalised knowledge can adaptively guide our behaviour and help us navigate the world. In this study, we aim to test the role of memory retrieval in promoting such generalisation of memories. Retrieval is known to be a powerful memory enhancer. Both cognitive and neurobiological theories of retrieval-mediated learning propose that this benefit is due to the co-activation of related (semantic) information during retrieval, which strengthens this co-activated associative network. By doing so, retrieval might play an important role in the generalisation of the memory trace.

Here, we used univariate and pattern fMRI analyses to investigate whether memory representations that undergo retrieval (*vs.* restudy) become generalised over time. Participants encoded scene-object pairs and either retrieved or restudied the objects over two sessions, two days apart. We analysed univariate and multivariate changes in brain activity specific to retrieval but not restudy, and tested whether predicted changes occur rapidly within a session, or evolve slowly, across the two days.

Consistent with a role of retrieval in the *semanticisation* of memories, univariate analyses showed an increase in medial prefrontal cortex (mPFC) activation across consecutive retrieval attempts, and a multivariate increase in similarity between categorically related information. In addition to this *semanticisation*, we also observed that retrieval strengthened the patterns unique to the original study episodes. Semantic-categorical and episode-unique strengthening both evolved slowly, across two days, and were most pronounced in parietal areas. Our findings corroborate the hypothesis that retrieval supports the creation of a generalised memory trace, and show that this strengthening does not come at the expense of episode-unique information. Active remembering thus seems to promote a stable and adaptive memory that can be flexibly used to access both contextually specific and more abstract generalised information.

Keywords: Retrieval; episodic-memory; consolidation; hippocampus; mPFC; testing effect

1. Introduction

Generalised knowledge about the world can promote new learning, help us navigate the world and adaptively guide our behaviour. An overly precise memory may be desirable in some situations, but "mnemonic overfitting" can also be detrimental when the learned information needs to be transferred to new situations (Richards & Frankland, 2017). For a simple intuition, imagine you are getting a flight from an airport you have never flown from before. Even though you do not know this particular airport, previous "Airport" experiences can help guide your behaviour, so that you know you need to queue at the check-in, pass security, check the departure board to know which gate your flight leaves from, etc. Specific information such as the colour of the departure board is irrelevant and might even create undesired interference in this particular case. It is the generalised information about airports that will help you catch your flight.

In the present study, we aim to test the hypothesis that the repeated reactivation of a recently acquired memory plays an important role in promoting such generalisation. Retrieval practice is known to be a powerful memory enhancer. Actively and repeatedly retrieving a newly acquired memory promotes its long-term retention to a much greater extent than merely restudying the same information. This retrieval benefit, termed the testing-effect or retrieval-mediated learning, has been replicated in countless studies, using different types of materials, learners and contexts (Roediger & Karpicke, 2006; Karpicke, Lafayette, & States, 2017).

Several cognitive theories have been put forth to account for retrieval benefits. Importantly, some of these theories propose that retrieval leads to a memory boost due to the co-activation of related information during the retrieval process. Co-activated items can then serve as additional retrieval cues during a subsequent memory test, as proposed by the *elaborative retrieval hypothesis* (Carpenter, 2009), or as mediating information that promotes a link between the target and the cue (*mediator effectiveness hypothesis*; Pyc & Rawson, 2010).

In an attempt to bridge these cognitive theories and neuroscientific research, we recently proposed that memory retrieval might rely on mechanisms similar to those underlying offline

systems consolidation (Antony, Ferreira, Norman, & Wimber, 2017). Systems-level consolidation theories (McClelland et al., 1995) propose that during offline periods (such as sleep) memories are reactivated or replayed in hippocampal-neocortical networks. Reactivation presumably allows new memories to become embedded in pre-existing cortical knowledge structures, creating a neocortical trace more durable than the initial hippocampal one (Dudai et al., 2015). Importantly, this neocortical trace is thought to store generalised or gist-like representations (Lutz, Diekelmann, Hinse-Stern, Born, & Rauss, 2017; Rasch et al., 2007; Richards et al., 2014; Schapiro et al., 2017). The neural replay of recently acquired memories is a key characteristic of both retrieval and offline consolidation, and can thus be expected to produce parallel effects on the underlying memory trace (Antony et al., 2017). In the present work, we specifically test the hypothesis that retrieval supports the generalisation of novel episodic associations. Due to the imprecise nature of a replay event, the attempt to retrieve one memory will lead to the co-activation of related ones. This assumption, while neurobiologically motivated, is shared with cognitive theories of the testing effect (Carpenter, 2009; Pyc & Rawson, 2010). We propose that this reactivation of associated information allows neocortical regions (including the medial prefrontal cortex, mPFC, Nieuwenhuis & Takashima, 2011) to extract commonalities between memories and promote their generalisation, mirroring the effects of offline systems-consolidation.

Note that this generalisation should be specific to retrieval but not to other types of practice, such as restudy, that do not entail the replay of an associated network to the same extent (Antony et al., 2017). Thus, in this experiment, participants performed consecutive retrieval and restudy blocks over two sessions, two days apart, to track both global changes in activity and changes in neural patterns across practice cycles. Specifically, we tested the following three key predictions. First, we hypothesised that retrieval but not restudy will mimic the univariate increases in mPFC and decreases in hippocampal activation that have been described in the sleep literature (Nieuwenhuis & Takashima, 2011; Takashima et al., 2009). Secondly, we hypothesised that if retrieval embeds novel episodes in neocortex, leading to a *semanticisation* of memories, the neural

patterns of memories sharing semantic information (items belonging to the same semantic category) should become increasingly similar across repeated retrievals. We also investigated whether such an increase in categorical representational similarity would be paralleled by a loss of episodic-contextual details of the mnemonic trace. This question was motivated by the assumption of standard systems consolidation theory that as memories become more gist-like, they lose contextual detail (Cairney, Durrant, Musgrove, & Lewis, 2011, but see Jurewicz, Cordi, Staudigl, & Rasch, 2016). If generalisation via retrieval is paralleled by a loss of detail, we would expect that similarity between episode-specific study patterns and the patterns reactivated during retrieval should gradually decrease across repetitions. On the other hand, the testing effect literature shows that repeated retrieval practice boosts recollection (e.g. Roediger & Karpicke, 2006) and could thus be expected to produce a strengthening of episodic-contextual aspects of the memory.

Finally, our experimental setup also allowed us to test whether the predicted uni- and multivariate changes in brain activity occur on a fast time-scale (within the first session of the experiment) or rather develop at a slower time-scale (over the two recording days). This question is of interest since the behavioural benefits of retrieval practice typically evolve slowly and are only evident after longer delays (Roediger & Karpicke, 2006). For example, Roediger and Karpicke (2006) show that when tested immediately, participants showed better recall for prose passages that were studied twice than for those that had been studied and tested. However, this pattern inverted after two days, and retrieval strongly outperformed restudy after one week. This delayed retrieval boost was replicated in a second experiment and in several studies since (e.g. Coppens et al., 2011; Toppino and Cohen, 2009; van den Broek et al., 2014), and suggests that the benefits of retrieval-mediated learning evolve slowly over long delays, potentially interacting with offline consolidation (Himmer et al., 2019). However, and consistent with our fast consolidation framework, it has recently been proposed that durable neocortical engrams can develop within a single session when memories are repeatedly reactivated (Brodt et al., 2018) and retrieval benefits can often be observed already within the first session of a study (Brodt et al., 2016).

In the present study, we were specifically interested in memory stabilisation and generalisation, and whether repeatedly recalling the same memory enhances these processes as predicted by our fast consolidation framework (Anthony et al., 2017). Understanding the role of retrieval in memory generalisation, and the timescale at which any retrieval-specific effects unfold, provides valuable insight into the neural mechanisms underlying retrieval-mediated learning, and may offer a plausible neurobiological basis for existing accounts of the testing-effect (Carpenter, 2009; Pyc & Rawson, 2010).

2. Methods

2.1 Participants

Twenty-four volunteers (17 female, M_{age} = 23.3, SD = 4.0) completed the two sessions of the experiment. Two participants were excluded, one due to extreme movement (more than twice the voxel size) and the other for reporting falling asleep in the scanner. Due to technical errors, the third practice cycle of the experiment is missing in one participant, and the fourth cycle in another. The remaining data from these two participants was fully included in the analyses. All 22 participants in the final sample were right-handed, native or very fluent English speakers, had normal or corrected-to-normal vision and no history of neurological, psychological or psychiatric conditions. Participants received course credit or a monetary reward for taking part in the experiment. The experiment was approved by the STEM Ethics Committee of the University of Birmingham.

2.2 Materials

A set of 128 scenes and 128 objects were chosen as stimuli. Each object and scene was unique, but objects belonged to eight different semantic categories: animals, musical instruments, fruits, clothes, sports gear, office supplies, kitchen appliances and furniture, with 16 exemplars per category. The objects were chosen from the Bank of Standardized Stimuli (BOSS; Brodeur, Guérard, & Bouras, 2014; https://sites.google.com/site/bosstimuli/), resized to 170x170 pixels and modified

to greyscale. The scenes were drawn from the SUN database (Xiao, Hays, Ehinger, Oliva, & Torralba, 2010; http://groups.csail.mit.edu/vision/SUN/). All scenes were resized to 256x256 pixels and displayed in greyscale. Stimuli were presented using in-house Python code, running on PsychoPy v.1.84.2 (Peirce, 2006; http://www.psychopy.org/).

2.3 Procedure

Participants were asked to complete two sessions on two different days (Figure 1A). On the first session, participants were provided with task instructions and MRI information and asked for their informed consent.

In Session 1, participants performed a study phase twice: once outside and once inside the scanner. This was followed by three blocks of practice, each comprising one retrieval and one restudy cycle on separate sets of the studied material. Participants then returned after two days (approximately 48h) for Session 2. In this session they completed an additional practice block (one retrieval cycle and one restudy) and a functional localizer, all inside the scanner.

2.3.1 Session 1

2.3.1.1 Familiarisation Phase

Before starting the actual experiment (i.e., the study and practice phases), participants were given an opportunity to get familiar with the tasks. The trials during familiarisation followed the exact same structure as the ones presented later during the experiment, but used 10 stimuli only (5 scene-object pairings) that were used specifically for this phase and never again seen during the remainder of the experiment. None of the objects belonged to any of the semantic categories presented in the experiment itself.

2.3.1.2 Study Phase

At the beginning of the study phase, participants were informed that they would see a series of scene-object pairs that they needed to commit to memory the best they could. Participants were instructed that, to achieve this, they should link the two (scene and object) together as vividly as possible by, for instance, mentally integrating the object into the scene. They were explicitly told that although each object and scene was unique, the objects belonged to a number of different semantic categories. After a 5-trial practice (familiarisation phase), subjects performed the first study phase for all 128 scene-object pairs in a quiet room in our imaging facilities. Participants sat in front of a laptop to perform the task.

A study trial began with a black fixation cross on a white background (jittered 0.5-7.5 sec). This was followed by a scene-object pair, presented in random order, for 4.5 sec. The scene was always shown in the centre of the screen, while the object was presented in one of the four corners. The pair was shown on a different coloured background – pink, blue, green or yellow (Figure 1B). The position of each object and the colour of the background were pseudo-randomly assigned, so that each colour and position would be equally distributed across categories and later practice conditions. Spatial positions and background colours were used to create a more unique encoding context for each item. For each pair, participants were asked to press a key on the keyboard to indicate whether or not they had been able to come up with a vivid mental image connecting the scene and the object.

After the first study phase, participants were taken to the scanner where, after the acquisition of the structural images (~5 minutes), they performed the second study phase for all object-scene pairings, this time inside the scanner. Images were projected on a screen behind the scanner bore that participants saw through a mirror attached to the head coil. This phase followed the exact same procedure as the one outside the scanner, but stimuli were presented in a different random order. That is, the scene-object pairing, the object position and the background colours

were kept constant, but the order of stimulus presentation was changed. Inside the scanner, responses were made on a button box that participants held in their right hand.

2.3.1.3 Practice Phase

The practice phase immediately followed Study. In Session 1, this phase comprised three retrieval and three restudy cycles. The order of conditions was counterbalanced across subjects, with half of the subjects performing retrieval-restudy, restudy-retrieval, retrieval-restudy, and the other half following the opposite scheme. Assignment of stimuli to conditions was also counterbalanced across subjects. Each subject performed one of four possible counterbalancing schemes. In each scheme, half of the stimuli (64), belonging to 4 semantic categories, were attributed to retrieval. The other 4 categories (64 stimuli) were assigned to restudy. Stimuli were allocated to retrieval or restudy pseudo-randomly so that each semantic category (with its 16 exemplars) was assigned to the retrieval condition in two of the counterbalancing schemes, and to the restudy condition in the remaining two schemes. Each given scene-object pair remained in either the retrieval or restudy condition across all four retrieval or restudy repetitions and across the two days. The order of stimulus presentation was randomized within participant for each cycle.

A retrieval trial (Figure 1B) started with a black fixation cross in the centre of a white screen (jittered 0.5-7.5 sec) and was followed by the presentation of the scene, as a cue to retrieve the object. The scene was presented for 3 sec over a white background, centred in the upper part of the screen with a black question mark below. Participants were instructed to think back to the item associated with this particular scene, and visualize it as vividly as possible in their mind for the full duration of the trial. Afterwards, participants saw four possible categories written on the screen (black font over white background) for 1 sec, and were asked to indicate the category to which the object associated with the scene they had just seen belonged to. For a given participant, the categories shown at retrieval and restudy, as well as the buttons associated to the response, were kept constant throughout the experiment. After the 64 trials of a retrieval cycle, there was a massed

feedback phase, where participants saw all the pairs together again, presented for 2 sec each. The scene was presented on the upper part of the screen, with the object centred below it, both over a white background. Participants were asked to press a key during these two seconds to indicate whether the object corresponded to the one they had thought of earlier or not. Massed feedback was included since it has been shown to enhance the retrieval-mediated learning effect (Roediger & Karpicke, 2006). Moreover, this manipulation allowed us to collect an additional (subjective) measure of participants' performance. Note that with the addition of the massed feedback phase, retrieval practice cycles are, effectively, cycles of alternated retrieval and restudy (feedback).

Restudy trials were very similar to retrieval ones. The only difference was that instead of the scene being paired with a question mark, participants saw the whole pair intact again (the scene in the upper part of the screen, with its corresponding object below, over a white background). Participants were instructed to take these trials as an opportunity to relearn the pairs, by again mentally integrating the object into the scene as vividly as possible. When the four categories appeared after each trial, they were asked to choose the category that the object of the current trial belonged to (1 sec). After 64 restudy trials there was also a "feedback" phase, where participants saw all the restudy pairs again. In this case, they were asked to press a button to indicate whether they still found it easy or hard to link the object and scene together. The trial structure of the restudy condition was thus highly similar to the retrieval condition, but involved no active retrieval demand at any point.

After each block of retrieval + restudy, there was a 2 minute break where participants were told to close their eyes if needed and rest. The first session ended after three cycles of retrieval and three of restudy practice. Participants were taken out of the scanner and sent home, with a reminder of the second session 48h later.

2.3.2 Session 2

Participants came back after two days to perform the second part of the experiment. Nineteen of the 24 came back exactly 48h later. For the remaining five, this was not possible due to personal or scanner booking constraints. Three of them came back on the same part of the day (e.g. tested in the morning on both sessions or in the afternoon on both sessions). The remaining two were tested in the morning on Session 1 but tested in the afternoon after two days.

On Session 2 subjects were first reminded of the practice phase instructions, and then given instructions for the localizer phase. Both phases were performed inside the scanner.

2.3.2.1 Practice Phase

The practice phase was identical to the one performed in Session 1. Participants that ended Session1 with a retrieval cycle started with a retrieval cycle, and those that ended with restudy started with restudy.

2.3.2.2 Functional Localizer

After the practice phase, participants were presented with all the object pictures again, with five repetitions per picture. Stimuli were presented at fovea on a white background for a duration of 2 sec, preceded by a black jittered fixation cross (0.5-7.5 sec). To make sure participants kept attending to the stimuli, a yes/no catch question appeared on the screen unpredictably. In order to not produce a high memory load, the question was always a simple question about the object shown on screen immediately before (e.g. “was the last object an instrument?”, “was the last object round?”). The localizer consisted of 730 trials: 640 of these were stimulus presentations (128 stimuli, repeated 5 times each) and 90 were catch questions. The localizer phase was divided into 5 continuous runs (not obvious to the participants). In each run the set of 128 stimuli and 18 questions were presented in a random order.

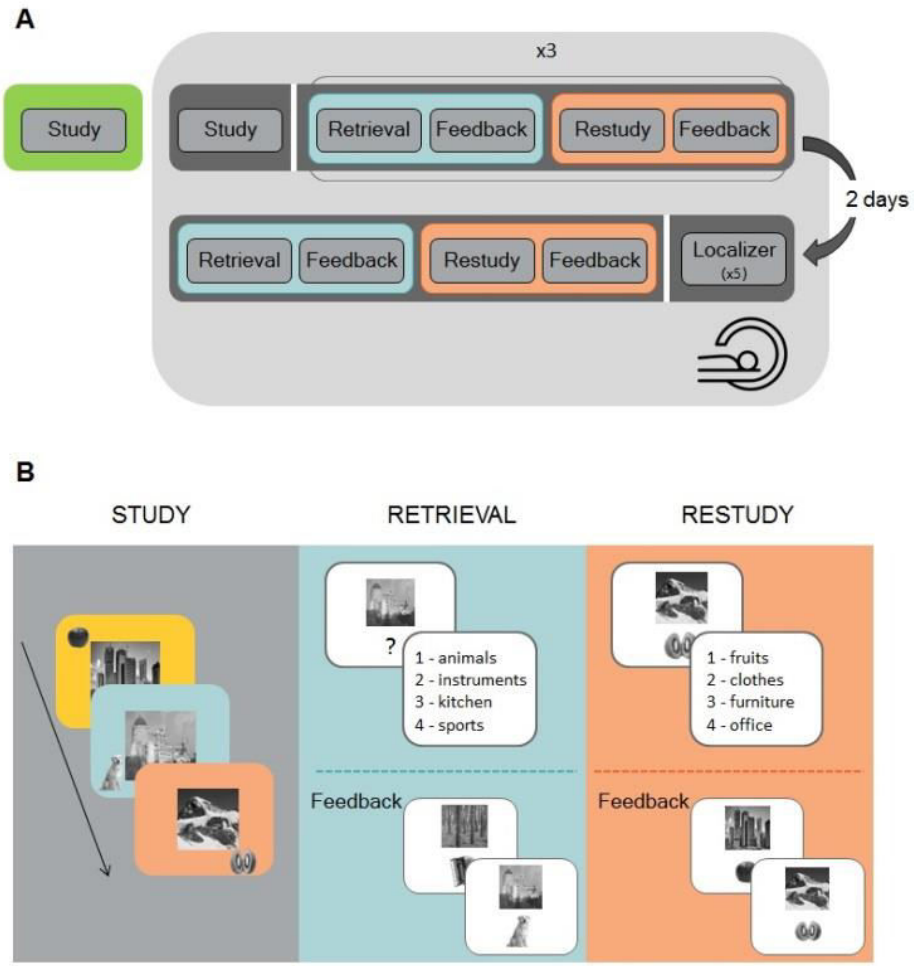


Figure 1: Experimental Procedure. **A.** General timeline of the experiment. Participants were tested on two different days, two days apart. In the first session, participants performed a study phase outside the scanner and another inside the scanner, followed by three cycles of retrieval and restudy trials. Each retrieval/restudy cycle was completed by a massed feedback phase. In the second session, participants performed an additional retrieval and restudy cycle followed by a functional localizer. All the phases encompassed by the light grey background took place inside the MRI scanner. **B.** Detailed depiction of the main phases of the experiment. During the study phase, participants saw different object-scene pairs and were asked to mentally link them together as vividly as possible. The scenes and objects were unique but, importantly, the objects belonged to a number of different semantic categories. During a retrieval trial, participants saw a scene and were asked to think back, as vividly as possible, to the object associated with that scene, and to then indicate its semantic category by a button press. Restudy trials unfolded in a similar way, the only difference being that the scene-object pair was presented intact on the screen and thus there was no need for participants to search their memory for a target object. A massed feedback phase followed each cycle, where subjects saw all the intact pairs again, presented in the centre of the screen.

2.4 fMRI acquisition and pre-processing

Images were acquired on a 3T Philips Medical Systems Achieva, at the Birmingham University Imaging Centre (BUIC), using a 32-channel head-coil. Participants were instructed to avoid movement as much as possible, and head motion was further restricted by using foam pads inside the RF coil.

High resolution (1x1x1 mm) T1-weighted images were acquired for each participant at the beginning of each session, using an MPRAGE sequence (with TR =7.4 msec; TE = 3.5 msec; flip angle = 7 degrees, and FOV = 256x256x176mm).

Functional images were acquired parallel to the longitudinal axis of the hippocampus, with isotropic voxels of 3mm, a TR of 2 sec, a TE of 30 sec, and a flip angle of 80 degrees. Each volume was comprised of 38 slices with no spatial gap between them. Slices were acquired in descending order and the first five volumes of each run discarded to allow for magnetic field stabilization.

Pre-processing of the images was done using SPM12 (University College London, London, UK; http://www.fil.ion.ucl.ac.uk/spm/). Motion and time correction (slices corrected to the middle slice) were performed in this order. Data was then linearly detrended, using a Linear Model of Global Signal algorithm (Macey et al., 2004) to remove global session and voxel effects. Functional and anatomical images were co-registered. For univariate analyses, the data was further normalized to an MNI template, and finally, images were spatially smoothed with an 8-mm FWHM Gaussian kernel. The multivariate analyses were performed in the participants' native space, that is, with no normalization or smoothing of the data.

2.5 Univariate analyses

Events of interest were modelled as stick functions and convolved with a canonical hemodynamic response function (HRF). The study, practice and localiser runs were modelled in the same general linear model (GLM). For study runs, we modelled the fixation cross, the onset of the scene object-pair and the onset of the response. For practice runs, fixation, scene onset, categories

onset, response, feedback pair onset and feedback response were used as regressors. The four breaks after each Retrieval + Restudy cycle were also included as regressors. Finally, the localiser regressors included fixation, object onset, question onset and response onset.

Since runs from the two sessions were included in the same model, session regressors were added to the GLM to account for differences between the two sessions. Likewise, motion parameters from spatial realignment were included as nuisance variables.

For each participant, beta values from the first level GLM were then extracted from one pre-defined anatomical regions of interest (ROIs), and one functionally defined ROI based on existing literature (see ROI definition below). The average beta values from each ROI were statistically analysed with participants as random effects as described in the following. To test one of our hypotheses, we assessed fast and slow changes in the brain. For fast-changing effects, we calculated whether univariate activity in a given ROI linearly increased or decreased within the first session by fitting a linear slope to the average beta values of the three retrieval and the three restudy cycles separately, individually per participant. The average slope was then tested against zero using a one-tailed dependent sample *t*-test, or compared between retrieval and restudy using a paired-sample *t*-test. This analysis identified neural changes that occurred across repeated retrieval or restudy cycles within the first day. To test for slow-changing effects across the two day delay, the average activation from repetition 3 (last cycle on Session 1) and repetition 4 (Session 2) in each condition were subjected to a 2x2 (cycle x condition) repeated measures ANOVA. Specific effects predicted a priori (i.e., planned comparisons) were then tested using dependent sample *t*-tests (one-tailed). This latter analysis allowed us to identify neural activity that was not yet present at the end of the first scanning session but then slowly evolved across days.

For completeness, whole-brain analyses were also conducted to (i) compare the average retrieval-related with the average restudy-related activity (see Supplementary Figure 1A and Supplementary Table 1); and (ii) to compute within- and across-session changes (Supplementary Figure 1B). These analyses reveal that, whereas retrieval tends to engage later visual and superior

parietal lobe areas, restudy engages relatively earlier visual and more inferior parietal areas, a pattern consistent with the imagery literature (Dijkstra et al., 2019) and the view that restudy re-imposes a specific visual representation on the brain, while retrieval engages higher-level, more abstract visual representations. Contrasts for within and between sessions are also shown for the retrieval condition, again highlighting the role of parietal regions. The within session effect, although small, show superior parietal changes congruent with previous literature (e.g. Karlsson Wirebring, 2015); cross session effects were more robust and show increased retrieval-related activity not only in parietal cortex, but also in inferior frontal gyrus, which is often present in studies investigating the neural processes supporting retrieval (e.g. Jonker et al., 2018; Wimber et al., 2015).

An additional univariate analysis was performed to rule out baseline differences between to-be-retrieved and to-be-restudied items. A new GLM model was built, with separate regressors for subsequently retrieved and subsequently restudied items at study. No significant differences were found between these items at a $p<.001$ *uncorrected* (Supplementary Figure 2A).

2.5.1 ROIs

The anatomical regions of interest (ROIs) for the univariate analyses were based on the existing sleep literature, which has shown that as memories become consolidated (and consequently generalised) there is an increase in mPFC activation and a decrease in hippocampal one (Gais et al., 2007; Takashima et al., 2006). Based on this evidence, beta values were extracted from hippocampus and mPFC and tested for fast and slow changes (see above).

The hippocampi were manually traced on each participant's structural image, using ITK-SNAP (Yushkevich et al., 2006; http://www.itksnap.org). The mPFC mask was built from a human atlas as implemented in WFUpickatlas 3.0.5b (Maldjian, Laurienti, Kraft, & Burdette, 2003; http://fmri.wfubmc.edu/software/PickAtlas). The mask, composed of Brodmann Area (BA) 10, was then back-projected to the participants' native space, using the inverse normalization parameters obtained from SPM during the segmentation step. The BA10 mask was chosen for being an area

commonly found in literature referencing mPFC. However, it encompasses lateral portions of the brain that might not be relevant for memory consolidation. To address this, and the fact that peak coordinates of mPFC vary across different fields of research relevant to our hypotheses (e.g. across studies looking at offline consolidation, schema formation, memory integration and retrieval-mediated memory enhancement) we used the Marsbar toolbox (Brett, Anton, Valabregue, & Poline, 2002; http://marsbar.sourceforge.net/) to create a second functional mPFC ROI. We drew an 8mm sphere around the peak coordinates reported in several studies in these different sub-fields of memory (for a list of these studies and peak coordinates used, see Supplementary Table 2) and built a ROI composed of the sum of these spheres (Figure 2A, lower panel).

2.6 Multivariate analyses

Two GLM models were created in order to allow comparisons between neural patterns. The first included the onsets from the study phase, with one onset per stimulus (onset of the scene-object pair; 128 regressors), plus one regressor modelling all response onsets. These were convolved with a canonical HRF. Motion regressors were also included, being treated as nuisance variables. The second GLM was comprised of 512 regressors, one for the onset of each scene in each practice cycle, convolved with a canonical HRF. Motion regressors, as well as session regressors were added to the GLM to account for movement and differences between the two sessions, respectively. These were treated as nuisance variables.

Multivariate analyses were conducted using the RSA toolbox (Nili et al., 2014; http://www.mrc-cbu.cam.ac.uk/methods-and-resources/toolboxes/). We first assessed whether there were any differences in overall similarity between to-be-retrieved and to-be-restudied items, by computing the average correlation between patterns at study for subsequently retrieved and restudied items. No significant differences were found ($p>.1$; see Supplementary Figure 2B).

Since we were mainly interested in areas where multivariate patterns representing our initially studied stimuli changed across subsequent retrieval or restudy trials, all subsequent RSA

analyses compared the patterns elicited during initial study with the patterns elicited during each subsequent retrieval and restudy cycle. For instance, we compared the pattern elicited when encoding a dog paired with its respective scene at study, with the pattern elicited each time the participant saw this scene and retrieved the dog from memory (or saw the pair together, for the restudy condition).

Specifically, we tested whether a given brain region increasingly (or decreasingly) coded for semantic categorical patterns, or for episodic (i.e., episode-unique) patterns across repeated practice (retrieval or restudy) cycles, compared to the initial study. The model matrices used for these two types of analyses are explained in more detail in the next sections. Again, we were interested in neural pattern changes that occurred either fast within a session, or slowly across sessions.

We first looked at semantic and episode-unique pattern changes in the two ROIs (hippocampus and mPFC). These analyses yielded no significant results in either of the comparisons of interest; that is, we found no categorical or episode-unique increases/decreases, neither within-session or across sessions at the preset uncorrected p-level of $p<0.001$. Since we had no other specific *a priori* hypotheses regarding where in the brain the semantic and episodic-specific pattern changes should take place, we conducted two separate searchlight analyses. Searchlight analyses are ideally suited to search for regions in the brain where information is represented in a specific representational geometry (Kriegeskorte, Goebel, & Bandettini, 2006).

2.6.1 Semantic Searchlight

To investigate how the semantic structure of memory representations changed from the original memory trace (at study) across consecutive retrieval and restudy repetitions, we conducted a searchlight analysis to find, for each cycle and each condition, where in the brain there was evidence for an increase in the coding of semantic category.

The first step of this analysis consisted of building a model matrix that reflected the expected patterns of results if *semanticisation* took place (Figure 3A). Stimuli were arranged according to their semantic category membership, in the same order across the rows and columns of the matrix (i.e. if the first four rows were dog-elephant-trumpet-accordion during encoding, the first four columns would be these same items in this same order during retrieval or restudy). Each pair of items was assigned a value, according to how similar we hypothesised them to be. Pairs of items belonging to the same semantic category (e.g. dog-elephant) were assigned a value of 1 (similar) and items belonging to different categories (dog-accordion) a value of -1 (dissimilar). Same item cells (dog-dog) were set to NaN in order to exclude any item-specific effects from this categorical analysis.

The searchlight ran in each participant was a sphere with a 9mm radius, sampling the voxel-wise patterns of activity within the sphere. A similarity matrix was computed for each sphere using simple Pearson correlations as a metric. The matrix at each point was then correlated with the previously defined model matrix (representing the hypothesis for the categorical searchlight; Figure 3A), to determine what regions in the brain behaved in accordance with the model. The resulting correlation was assigned to the centre voxel in each given sphere. As a first step, a separate searchlight was run for each retrieval and restudy cycle. From these searchlight analyses, an activation map (*r*-map) was obtained for each subject and each retrieval or restudy cycle, depicting the degree to which a given brain area showed a representational geometry (Kriegeskorte et al., 2008) that was similar to that of the model matrix. The maps were normalized to MNI space (using the parameters from the segmentation of the T1-weighted image), smoothed with an 8-mm FWHM Gaussian kernel, and statistically compared in a random-effects analysis (see next paragraph). This first analysis step resulted in 8 searchlight maps per participant, 4 representing categorical patterns during each retrieval cycle, and 4 representing categorical pattern similarity during each restudy cycle.

As a second step, group analyses were conducted on the normalized and smoothed searchlight maps from each subject within a second-level GLM. Similarly to the univariate ROI

analyses, one contrast was established to assess fast-changing (within session) effects, and another one to assess slow-changing effects between the two sessions. For the within session effects we looked for the interaction of regions that increased linearly across retrieval but not restudy repetitions. For between session effects we contrasted regions that increased from repetition 3 (last cycle on day 1) to repetition 4 (day 2) in the retrieval but not the restudy condition. The results are reported at an uncorrected p-level of $p<0.001$, with a minimum extent threshold of $k=10$ voxels. We further assessed whether these contrasts survived correction for multiple comparisons (at p<.05, FWE).

Our main hypotheses all concerned similarity changes that were significantly more pronounced in the retrieval compared to the restudy condition. For reasons of completeness, corresponding contrasts were also created for regions showing an increase in similarity for restudy but not retrieval. These contrasts are not reported since they yielded no significant results in any of the group-level comparisons.

2.6.2 Episode-Unique Searchlight

Episode-unique effects were assessed in a similar way to categorical ones. The major difference was the definition of the model matrix. In this case, we were interested in areas in the brain that coded the similarity between each item's unique representation at study, and the same item's subsequent retrieval or restudy representation, beyond their shared category. Accordingly, same item cells (dog-dog) were set to 1 (high similarity) whereas cells of items belonging to the same category (dog-elephant) were set to -1 (low similarity). Between category cells (dog-accordion) were set to NaN (Figure 3B).

The rest of the analysis followed the same procedure as the semantic searchlight. Results are reported at an uncorrected p-level of $p<0.001$, with a minimum extent threshold of $k=10$ voxels. Clusters that survived an FWE correction (p<.05) are also shown in Figure 3D.

3. Results

3.1 Behavioural performance during practice

Behavioural analyses were conducted to assess performance during practice cycles. We first analysed the choice of the correct semantic category after the presentation of each stimulus pair in retrieval or restudy cycles (Table 1). A 2x4 (condition x cycle) repeated measures ANOVA was conducted and yielded a significant main effect of condition [$F(1,19) = 31.56$, $p=.000$, $\eta^2_{partial} =.63$], a significant main effect of cycle [$F(3,57) = 18.42$, $p=.000$,$\eta^2_{partial} =.49$] and a significant condition x cycle interaction [$F(3,57) = 11.83$, $p=.000$, $\eta^2_{partial} =.38$]. Not surprisingly, the proportion of correct responses for restudy trials was higher than for retrieval trials ($M_{restudy} = 0.89$, $SD= 0.08$; $M_{retrieval}= 0.79$, $SD=0.13$). Note that although performance in restudy trials is very high, it was hard for participants to reach 100% accuracy, given the time limit imposed for their response (1 second, see Methods section), which led to some missing responses. The main effect of cycle reflects the fact that, regardless of condition, performance on the first cycle was significantly lower than performance in all subsequent ones. No other significant differences were found across cycles. The interaction shows that performance during restudy cycles remained relatively constant, while retrieval performance linearly increased from one cycle to the other across the first three repetitions and decreased again on the fourth (after two days; see Table 1).

To ensure participants were not only retrieving object category, but actually thinking back to the item associated with the retrieval cue, we measured the proportion of trials where after choosing the correct semantic category participants also reported (during the massed feedback phase) that they had thought of the correct item. In this analysis, we found a main effect of cycle [$F(3,42) = 6.00$, $p=.002$, $\eta^2_{partial} =.30$]. Performance increased from the first retrieval attempt to subsequent ones (Table 1).

3.2 Univariate changes in mPFC and the hippocampus

Regarding univariate effects, we hypothesised that if retrieval acts as a consolidation-like mechanism, promoting memory's generalisation, neural changes specific to this condition should parallel the effects reported in the sleep-dependent consolidation literature (Gais et al., 2007; Takashima et al., 2006), namely an increase in mPFC and a decrease in hippocampal engagement. To test for fast changes in global activity we computed the linear slopes of mPFC and hippocampal activation for retrieval and restudy across the first three practice cycles (Session 1 of the experiment; see Figure 1A), and tested whether retrieval slopes show a steeper increase (mPFC) or decrease (hippocampus) across retrievals compared to restudy repetitions. To test for slower changes that occur across sessions, we compared activations on the last cycle of Session 1 (repetition 3) with the practice trials in Session 2 (repetition 4).

Table 1: Behavioural results during the practice phase. Restudy 1 was significantly different from 2 [t(21) = -2.73, p=.012)] and 4 [t(20) = -4.32, p=.000)]. No other significant differences were found in restudy. Retrieval 1 was significantly different from all the others (Ret1-Ret2 t(21) = -7.70, *p*=.000; Ret1-Ret3 t(20) = -7.13, *p*=.000; Ret1-Ret4 t(20) = -3.72 *p*= .001). Retrieval 2 differed from 3 (t(20) = -2.26, *p*=.035) and 3 from 4 (t(19) = 2.87, *p*=.010).

	Overall	Restudy	Retrieval
Correct Semantic Category	*M (SD)*	*M (SD)*	*M (SD)*
Cycle 1	0.79 (0.09)	0.87 (0.08)	0.70 (0.13)
Cycle 2	0.85 (0.08)	0.90 (0.08)	0.81 (0.12)
Cycle 3	0.87 (0.10)	0.89 (0.10)	0.86 (0.11)
Cycle 4	0.85 (0.10)	0.92 (0.07)	0.78 (0.15)
		Vivid | Correct Category	*Correct Trial | Correct Category*
Cycle 1	---	0.73 (0.18)	0.75 (0.22)
Cycle 2	---	0.74 (0.24)	0.88 (0.10)
Cycle 3	---	0.72 (0.23)	0.86 (0.15)
Cycle 4	---	0.71 (0.20)	0.87 (0.10)

3.2.1 Medial Prefrontal Cortex

In the BA10 ROI (Figure 2A, upper panel), the retrieval slope significantly differed from zero ($t(19) = 3.66$, $p = .001$), showing a within-session increase from first to third repetition. The restudy slope did not differ significantly from zero ($t(19) = -.65$, $p = .26$). The linear slopes of retrieval and restudy significantly differed from each other ($t(19) = 2.60$, $p = .009$). Similar results were obtained for the functional mPFC ROI based on the literature (Figure 2A, lower panel). For this ROI, whereas the retrieval slope significantly differed from zero ($t(19) = 3.99$, $p = .000$), the restudy slope did not ($t(19) = -.32$, $p = .49$). These slopes significantly differed from each other ($t(19) = 1.7$, $p = .05$). The within-session univariate results thus suggest a rapid change in activity that is specific to retrieval (Figure 2A).

No evidence for slowly evolving changes was found, with no significant effects across sessions, either for the BA10 ROI (main effect of condition: $F(1,19)=1.58$, $p = .224$; main effect of session and condition x session interaction both $F<1$, n.s) or for the literature-based mPFC (main effect of condition, main effect of session and condition x session interaction all $F<1$, n.s).

3.2.2 Hippocampus

In the hippocampus (Figure 2B), the retrieval slope was significantly different from zero ($t(19) = -2.39$, $p = .01$), whereas the restudy one was not ($t(19) = -.70$, $p = .25$). We found no significant differences between conditions within the first session ($t(19) = -1.69$, $p = .107$).

We did, however, find a decrease across sessions when comparing the third practice trial (end of day 1) to the fourth practice trial (day 2). The decrease was significant for retrieved items ($t(19)=2.02$, $p = .029$; Figure 2B) but not for restudied items ($t(19)=-.374$, $p = .357$). There was a trend towards a condition x cycle interaction that did not reach statistical significance [$F(1,19)=3.48$, $p = .078$]. The pattern is therefore indicative of a slow retrieval-specific hippocampal disengagement across days.

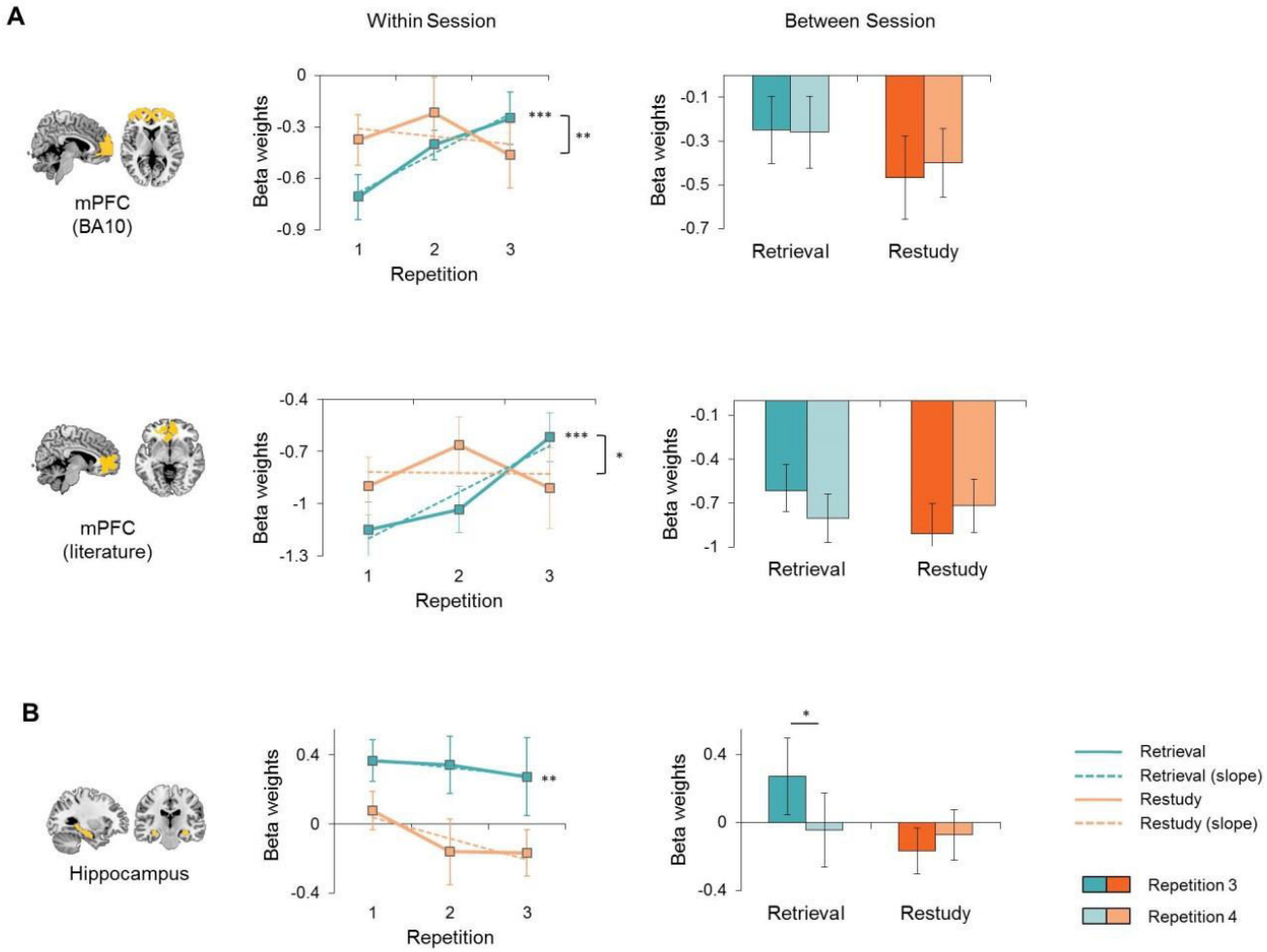


Figure 2: Univariate results within (left) and between (right) sessions. Retrieval results are presented in blue and restudy in orange, with the third repetition in darker colours than the fourth. Error bars represent standard errors of the mean across subjects. Significant differences are denoted with an * (*p<.05, **p<=.01, ***p<=.001). **A.** mPFC results: mPFC activation increased within the first session of the experiment for retrieval trials, both for a ROI composed of Broadmann area 10 (upper figure) and for a functional ROI built based on previous studies in the literature (see Supplementary Table 2). For both ROIs, the difference between retrieval and restudy slopes (dashed line) was statistically significant. No significant differences were found for either condition across sessions. **B.** Hippocampus results: hippocampus activation did not decrease for retrieval within the first session, and there was no difference between the slopes of both conditions. A significant difference was found, however, between the third and fourth retrieval repetitions, with hippocampus activation decreasing across the two days. No such difference was found for restudy trials. The session x condition interaction did not reach significance.

3.3 Multivariate effects

To investigate what regions in the brain increasingly coded for semantic and episode-unique effects, we ran two independent whole-brain searchlights. These searchlights compared the representational geometry at each searchlight location with a conceptual model matrix formalising our hypotheses regarding the retrieval-specific changes in category- and episode-level coding (Figures 3A and B), and specifically the level of similarity between representations at study and at each subsequent cycle of practice (retrieval or restudy). These analyses resulted in 8 activation maps (*r*-maps) per participant, representing regions in the brain that behave in accordance with the model matrix during each individual retrieval and restudy cycle. These *r*-maps were then subjected to second-level group analyses, following a similar logic as for the univariate analyses. For within-session effects, we computed a linear contrast encompassing the first three repetitions of retrieval or restudy (Session 1). To assess between-session effects, we contrasted regions where similarity with the model increased from the third to the fourth cycle (i.e., across days) for retrieval but not restudy, paralleling the univariate analyses reported above.

3.3.1 Semantic searchlight

Our main hypothesis predicts that retrieval should lead to an increase in categorical similarity, reflecting the *semanticisation* of items, across consecutive retrieval cycles. If such *semanticisation* occurs rapidly, as predicted by the fast consolidation account (Antony et al., 2017), it should be observed in neural pattern changes within Session 1 (linear contrast). If the process is dependent on subsequent consolidation, it would be expected to evolve slowly across the two scanning days.

Comparing our semantic category model matrix (Figure 3A) to the empirical similarity matrices acquired from each point of the searchlight, we found no significant changes in semantic structure within the first session. We did, however, find that across sessions, parietal regions (Figure 3C) increasingly coded for semantic category, including clusters in bilateral cingulate gyrus,

precuneus, left superior temporal gyrus and angular gyrus (Table 2). These neural changes were significant at a threshold of $p<.001$ *uncorrected* ($k=10$ voxels), but did not survive corrections for multiple comparisons (FWE, $p<.05$).

3.3.2 Episode-unique searchlight

Standard systems consolidation accounts assume that as memories become more gist-like, they lose contextual detail (Cairney, Durrant, Musgrove, & Lewis, 2011). Accordingly, we expected to find a decrease in similarity between episode-unique information (similarity between an item's study patterns and subsequent retrievals, compared with items from the same semantic category) across cycles, within the first session.

We found no significant increases or decreases in similarity for episode-specific information within Session 1. However, strong item-unique effects were found across sessions, with mostly parietal regions encoding episode identity significantly more on cycle 4 (day 2) compared to cycle 3 (day 1) of retrieval, with no difference between restudy cycles (Figure 3D). Regions implicated in episode- and retrieval-specific pattern changes ($p<.001$, *uncorrected*, $k=10$ voxels) included bilateral superior and inferior parietal lobe, as well as bilateral precuneus, middle occipital gyrus and right supramarginal gyrus (Table 2). One of these clusters survived multiple comparisons correction (FWE $p<.05$; see blue cluster in Figure 3D). The regions found in this searchlight did not overlap with those found in the semantic searchlight (see Supplementary Figure 3).

Table 2: Searchlight results. Significant clusters and peak coordinates are shown from a contrast testing for an interaction between condition (retrieval vs. restudy) and practice repetition (3 vs. 4; [Retrieval3<Retrieval4] > [Restudy3<Restudy4]).

	Hemisphere	BA	Cluster label(s)	XYZ	t-value	k
Semantic Effects						
	R	31	Cingulate gyrus Superior parietal	3 -40 41	3.95	117
	L	39	Superior temporal gyrus Angular gyrus	-60 -58 26	3.76	79
	R		Uvula	33 -67 -25	3.59	23
Episode-specific Effects						
	R	7/40	Superior parietal Precuneus	30 -58 41	4.37	522
	L	7	Inferior parietal	-33 -58 44	4.32	289
	L		Thalamus	-21 -25 11	3.54	27
	L	17	Cuneus Middle occipital gyrus	-21 -79 14	3.38	23
	R	40	Inferior parietal gyrus Supramarginal gyrus	48 -37 35	3.29	13

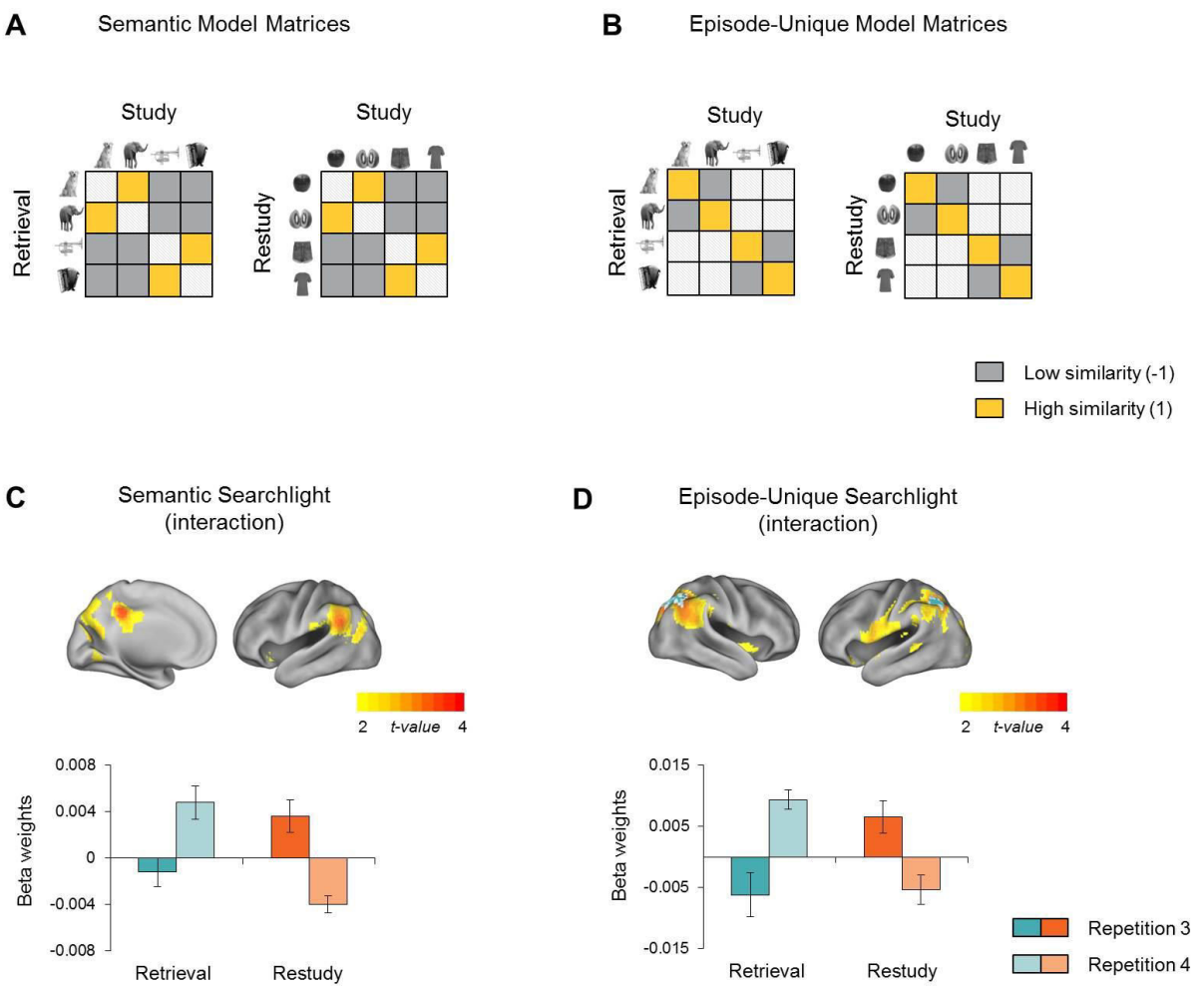


Figure 3: Searchlight model matrices and results for semantic information (left panel) and episode-unique information (right panel). **A.** Model matrices used to investigate regions in the brain coding for categorical representations. We looked for brain areas where similarity between items belonging to the same category (yellow cells) was higher than that of items belonging to different semantic categories (filled grey cells). Cells that are not filled were set to NaN. This matrix was correlated with the activation patterns obtained from each searchlight point to determine areas in the brain that code for categorical information in a similar fashion. **B.** Model matrices used to assess which neural regions code for episode-unique, that is, regions where similarity between the same item (yellow cells) was higher than similarity between items belonging to the same semantic category (filled grey cells). Cells that are not filled were set to NaN. **C.** Interaction analysis ([Retrieval3 > Retrieval4] > [Restudy3 > Restudy4]) of semantic effects. We found left parietal regions (upper figure) to increasingly code for category structure during retrieval but not restudy. The lower figure shows beta weights extracted from the all the regions depicted above, at a $p<.001$ *uncorrected* ($k=10$ voxels). Note that this graph has no statistical value and is used for visualisation purposes only, to assess what drives the interaction between the two conditions. **D.** Interaction analysis of episode-unique effects. Bilateral parietal regions were found to increasingly code for episode unique information (upper figure), showing a strengthening of the original study trace. This pattern was found for retrieved but not restudied items (lower figure: beta weights extracted from the regions depicted above). Error bars indicate standard errors of the mean across participants.

4. Discussion

Active retrieval is known to promote the long-term retention of newly learned episodic memories. According to cognitive theories of the testing-effect (Carpenter, 2009; Pyc & Rawson, 2010), the co-activation of related memories during retrieval underlies these benefits. This is also one of the central tenets of a framework in which we recently proposed that retrieval acts as a fast-consolidation mechanism, with memory reactivation in hippocampal-neocortical circuits presumably leading to the online stabilisation of memories across retrieval attempts (Antony et al., 2017). We here tested the central prediciton that co-activated related information facilitates memory integration during retrieval, and thus leads to a generalisation or *semanticisation* of the memory trace. In addition, we also investigated whether a potential generalisation is associated with a simultaneous change in episode-unique mnemonic representations.

First we tested whether univariate changes in brain activity, induced by repeated retrieval, are consistent with a consolidation-like mechanism. Paralleling what has been reported in the sleep literature (Gais et al., 2007; Takashima et al., 2006, 2009), we found an increase in mPFC activation across successive retrieval attempts within the first session, with no corresponding increase in the restudy condition. This pattern was evident using an anatomic definition (BA10), as well as a functional definition based on the literature (Figure 2A). Although indirect, this observation suggests that in terms of global engagement of brain regions, retrieval mimics the changes observed with increasing systems consolidation (Takashima et al., 2006, 2009). Memory consolidation models have attributed a specific role to mPFC in supporting schemas (Robin & Moscovitch, 2017; Sekeres et al., 2018), which represent common elements among exemplars within a class of events. Consequently, mPFC engagement in offline consolidation studies has often been interpreted as indicating a memory trace that is more generalised in nature than the original hippocampus dependent one (e.g. Gais et al., 2007). In line with this idea, we also found a decrease in hippocampal activation across days: hippocampal activity significantly decreased for retrieved, but not restudied, items. Note, however, that there was no statistically robust practice cycle by condition interaction, and evidence

for this hippocampal decrease being specific to retrieval is thus weak. Contrary to mPFC activity, the hippocampal change was observed across sessions, not within the first day, suggesting that at short delays, retrieval continues to rely on the hippocampus. However, actively retrieving an episode multiple times immediately after learning appears to benefit memory at longer delays of several days, where the memory can now be accessed with little hippocampal involvement[1].

Taken together, our univariate results show that retrieval produces a rapid increase in mPFC engagement and a slow disengagement of the hippocampus. These changes are not present after repeated restudy, and mimic a pattern that has been described in studies of sleep-dependent memory consolidation, commonly interpreted as reflecting the creation of a more generalised neocortical memory (Lutz et al., 2017; Rasch et al., 2007; Richards et al., 2014; Schapiro et al., 2017). At shorter delays, both regions appear to support memory retrieval, indicating that the neocortical and hippocampal traces co-exist (as previously suggested in relation to multiple trace theory; see Winocur, Moscovitch and Bontempi, 2010). This might create the opportunity for these areas to communicate and consolidate the memory trace, even though access to the memory will still be supported by the hippocampus at short delays, when the hippocampal trace is presumably dominant (Winocur et al., 2010).

More direct support for the role of retrieval in generalising memories comes from our multivariate analyses. Here we were interested in specific qualitative changes in representational geometries, induced by repeated retrieval at faster and slower timescales. We predicted that similarity between items belonging to the same semantic category (e.g. animals) should increase across retrieval but not restudy attempts. Changes in similarity were not found in our pre-defined ROIs, which might indicate that these regions play a role in coordinating the communication between different brain networks (Himmer et al., 2019; Reagh and Ranganath, 2018), rather than storing the mnemonic representations themselves. We did, however, find these retrieval-specific categorical changes when using a whole-brain searchlight, most pronouncedly in parietal regions

(see discussion further below) that are part of a core recollection network (Rugg & Vilberg, 2012), and exclusively across sessions.

Contrary to what standard systems consolidation theories would predict (Cairney, Durrant, Musgrove, & Lewis, 2011), this strengthening of semantic information did not come at the cost of episodic detail. Instead, we found that idiosyncratic study traces became increasingly represented in brain activity across repeated mental reinstatements. This effect was most pronounced in superior parietal regions, and again only found across a two-day delay, indicating a slowly evolving change in the underlying memory trace.

Taken together, our pattern fMRI results indicate that retrieval, when followed by a 48h delay, concurrently strengthens semantic-categorical and episodic-contextual aspects of memories. This observation contradicts standard consolidation accounts that would predict a loss of episodic detail as memories become more gist-like (Cairney, Durrant, Musgrove, & Lewis, 2011). Notably, however, previous studies in the offline literature also found no evidence for sleep-dependent decontextualisation (Jurewicz et al., 2016), and showed that episodic and semantic effects are not mutually exclusive but can co-occur (Schlichting, Mumford, & Preston, 2015; Schapiro et al., 2017; Tompary & Davachi, 2017). Again, the multivariate effects induced by retrieval in the present study thus mimic the changes found in memory representations after offline consolidation, and add to a growing body of evidence supporting the co-existence of two cortical traces, consistent with a multiple trace view (Winocur et al., 2010).

Going beyond previous work, our multi-day design allowed us to explicitly investigate the timescale at which these retrieval-induced neural pattern changes occur. We observed that most changes (apart from the rapid mPFC increase) evolved slowly and were only significant after two days. While this observation is surprising given our original fast-consolidation hypothesis (Antony et al., 2017), it resonates well with the cognitive testing effect literature which consistently reports the most robust memory benefits after delays of several days (Roediger & Karpicke, 2006), and suggests that retrieval exerts its long-term benefits by interacting with sleep- or time-dependent processes.

At this point we can only speculate about the nature of such interactions. It is possible that repeated retrieval "tags" new memories for prioritized replay during subsequent offline periods, just like reward or saliency would (Antony and Paller, 2017; Dudai et al., 2015), leading to more stabilisation and in turn less forgetting than for non-retrieved or restudied memories. These immediate "tags" might be invisible, or at least impossible to detect with the pattern tracking methods we used in the present study. Having said that, a recent study by Brodt et al. (2018) found that in parietal lobe (precuneus in particular), repeated learning is associated with changes in diffusion-based measures that are assumed to reflect local plasticity. These changes evolved relatively rapidly during the first learning day, and were related to the long-term retention of the learned information. DTI-based measures might thus provide a useful proxy for immediate plasticity. A follow-up fMRI study by the same group (Himmer et al., 2019) indicated that repeated rehearsal indeed produces increases in parietal lobe engagement, but sleep is required to stabilise these activity changes. These findings are consistent with a tagging account, where repeated retrieval after learning is necessary but not sufficient for creating a stable, long-lasting memory trace. We would like to emphasize, however, that empirical evidence exists that is inconsistent with such a tagging idea (Bäuml et al., 2014), and that the nature and locus of such hypothesised "priority tags" are currently not understood.

In terms of location, both of our searchlights produced effects in the parietal cortex, generally highlighting the role of posterior association cortices in representing reactivated mnemonic content during recall (Lee et al., 2017) and imagery (Dijkstra et al., 2019), and reflecting the dynamic change of these mental representations over time (Sommer, 2016). Interestingly, we found distinct sub-regions within the parietal lobe to track the changes on the categorical-semantic and the episodic-contextual level, with very little overlap between the two searchlight analyses (see Supplementary Figure 3). Semantic effects were mainly found in medial (posterior cingulate and precuneus) and inferior lateral parietal regions (Figure 3C). This is consistent with these areas' role in the retrieval of semantic and schema-relevant information (Binder et al., 2009; Binder and Desai, 2011; Linden et al., 2017; Sommer, 2016), with some authors specifically highlighting these regions

as hubs responsible for recombining or integrating schema-consistent information (Gilboa & Marlatte, 2017; Himmer et al., 2019; Wagner et al., 2015). Categorical reactivation in the same set of areas has previously been observed in a study investigating repeated memory retrieval (Lee et al., 2017), and the location of our *semanticisation* effects is thus largely consistent with the existing literature.

Less evidence exists for the role of superior parietal lobe in representing reinstated content. Our own findings provide strong evidence for a strengthening of the episode-unique aspects of an item over repeated retrievals, in these more superior aspects of parietal lobe. Many studies that tracked item-level reactivation used a ROI-based approach and focused specifically on the medial and inferior portions of parietal lobe described in the previous paragraph (e.g. Jonker et al., 2018; Kuhl and Chun, 2014; Lee et al., 2017). Other studies do provide evidence for episode-specific reinstatement in more superior regions (Favila et al., 2018). In the context of repeated testing, it was shown that in superior parietal lobe, a gradual increase in item distinctiveness (i.e., a decrease in between-item similarity), predicts long term retention, highly consistent with our present finding of an increased "episodicness" for retrieved items in the same area (Karlsson Wirebring et al., 2015). Interestingly, it has also been demonstrated that superior parietal areas are more dominant in representing the information during memory recall and mental imagery than during actual perception (Dijkstra et al., 2019; Favila et al., 2018). Moreover, the degree to which relevant mental contents are activated in these areas positively relates to memory strength during retrieval (Ye et al., 2016), and to the saliency and vividness of mental images in imagery tasks (Dijkstra et al., 2019). Studies in several domains, including memory retrieval, therefore suggest that superior parietal areas contribute to representing vivid mental images, and might thus naturally relate to episodic-recollective aspects of memory reinstatement.

We believe that our findings can help to shed light onto the neurobiological mechanisms underlying retrieval-mediated learning, and to bridge the existing gap to more cognitively framed theories of the testing-effect (Karpicke et al., 2017). Many of these theories, including the

elaborative retrieval and semantic mediator hypotheses (Carpenter, 2009; Pyc & Rawson, 2010), share one central assumption: that retrieval has a tendency to co-activate associatively linked information, providing an opportunity to interlink a new memory with existing knowledge or other episodes, thereby creating additional access routes to the target memory. These theories motivated our proposal that retrieval acts as a consolidation event (Anthony et al., 2017), and are highly consistent with the observation that repeated retrieval generalises memories. Even though previous studies on retrieval-mediated learning did not directly test for *semanticisation* via repeated testing, they too fit well with such an interpretation. Wing et al. (2013) showed stronger functional connectivity between hippocampus and mPFC after retrieval compared with relearning, consistent with a consolidation-like effect (Gais et al., 2007). Lee and colleagues (2017) investigated how multivariate patterns during retrieval practice predicted memory on a final recognition test. In line with our own results, they found an increase in categorical reactivation in parietal cortex. Interestingly, this increase was associated with a higher likelihood that participants behaviourally endorsed a perceptually similar lure as old, indicating a role of retrieval in generalising memories.

On the other hand, contextual reinstatement accounts (Rowland & DeLosh, 2014) have been put forward which assume that retrieval facilitates the reactivation of previously stored contextual information. This assumption is well in line with behavioural work showing that testing boosts recollection (see Karpicke et al., 2017 for a recent review), and our finding that episode-unique information is increasingly activated across repeated retrievals. Jonker et al. (2018) directly compared pattern changes across retrieval and restudy opportunities, and found that retrieval strengthens the neural representations of the target objects along with contextually linked objects, again in parietal cortex. This study is thus consistent with a contextual reinstatement account (Rowland & DeLosh, 2014) and provides evidence that retrieval-mediated generalisation can affect contextually related items as much as semantically related ones. The above mentioned work by Lee and colleagues (2017) also found that in addition to the categorical effects, item-specific reactivation in parietal lobe was related to later correct item memory. Like the present study, this study thus

found simultaneous categorical and episodic strengthening via retrieval, and suggests that these neural effects are associated with specific behavioural consequences of repeated testing: better memory for the specific episode, but along with a tendency to generalise this benefit to related items and thus a corresponding increase in false alarms.

Together with our searchlight results, existing investigations of retrieval-mediated learning suggest that retrieval shapes representations by extracting commonalities between new memories and previously stored knowledge, while preserving relevant stimulus-specific information. Retrieval can thus affect a memory representation at multiple levels, from global-categorical to idiosyncratic, episode-specific features. These neural changes are not mutually exclusive, an observation that might help unify different accounts of the testing effect.

A few limitations to our study should be mentioned. First, our study was purely focused on neural changes via repeated retrieval, and did not include a final behavioural memory test. Consequently, we are unable to claim that the generalisation and episodic strengthening effects observed in our study will translate into behaviour in the short- or long-term. As mentioned above, however, there are studies indicating that repeated retrieval enhances the strength of the target memory (Lee et al., 2017) and produces an increase in contextual recollection (Chan & McDermott, 2007), but concurrently also increases the likelihood of endorsing perceptually (Lee et al., 2017) or semantically (Roediger et al., 1996) related lure items as old. The pattern of behavioural changes that has been reported in the existing literature is thus well aligned with our finding that retrieval concurrently strengthens semantic and episodic aspects of a memory trace.

A second potential limitation concerns stimulus selection, and in particular the fact that items in our eight semantic categories were not matched for lower-level visual properties. Dynamic changes in categorical similarity might thus at least partly be driven by shared perceptual (e.g. shape) rather than conceptual features within a category, and the use of the term *semanticisation* does not imply that we measured purely semantic, abstract stimulus properties. In fact, we would like to emphasize that objects in the real world often do share low-level properties, that humans and

machines tend to use these properties for object categorization (Wardle and Ritchie, 2014), and that perceptual features can be represented at very late stages of visual processing (Martin et al., 2018). While these low-level features might not be sufficient to classify objects (Clarke et al., 2015; Kaiser et al., 2016), we believe that they are naturally confounded with semantic category membership, and eliminating them would create an artificial stimulus set that might not capture real-world semantics well.

A final potential limitation is that out of the two searchlight results, only the episode-unique effects in parietal lobe are present when using a conservative correction for multiple comparisons, while evidence from the categorical searchlights in support of our major hypothesis was less strong. We believe that the moderate effects in both analyses could be due to the fact that our searchlight approach uses an unusually specific contrast, looking for *differential* semantic/item-unique study-retrieval similarity relative to non-related pairs in the first stage, and then comparing these differential effects across repetitions and between conditions. Such an approach might make it difficult to observe *any* effect on a whole brain level (i.e., low sensitivity), but we are confident that the contrasts appropriately reflect the relevant cognitive comparisons, and that the analyses should thus produce results with high specificity.

In sum, in the present study we empirically tested whether retrieval (*vs.* restudy) promotes the generalisation of novel memory traces. We show that retrieval leads to a rapid increase in mPFC activation. This is accompanied by a slower cross-day decrease of hippocampal activation, congruent with a consolidation-like mechanism, although not necessarily a fast acting one. Moreover, in searchlight analyses we found a retrieval-specific strengthening of categorical-semantic features of a memory, concurrent with a strengthening (rather than a loss) of episode-unique information. These findings provide a neurobiological substrate for accounts of retrieval-mediated learning positing that retrieval's benefits on long-term retention depend on the co-activation of semantically and contextually related information. Creating an enduring engram that contains generalised information in addition to episodic detail is highly adaptive in everyday life where contexts change

frequently, and knowing the commonalities between past experiences can guide behaviour in novel situations.

Footnotes

1. Note that these univariate changes cannot be easily explained by effort or difficulty based interpretations. Such interpretations would predict an increase rather than decrease in hippocampal engagement after a delay of several days, since retrieval is rendered more difficult then (as evidenced by the drop in behavioural performance on day 2). Similarly, it could be argued that the retrieval-specific increase in mPFC activation is simply reflecting retrieval being initially more demanding than restudy, but becoming easier across repetitions, associated with an increase in default mode regions (Fox et al., 2005). Contrarily, we found that mPFC activity remained elevated across the two-day delay, when the task is again rendered more difficult.

   Along similar lines, it could be argued that motivational differences are at the base for the different neural dynamics found here. Although we cannot completely rule out this explanation, we do believe participants' motivation during restudy trials was still high. An indication is the fact that across restudy cycles, behavioural performance did not decline and was kept at a high level, averaging at 90%.

Acknowledgments

We thank Alexandru-Andrei Moise for his work tracing the participants' hippocampi.

This work was supported by grant ES/M001644/1, from the Economic and Social Research Council UK, awarded to M.W and by the British Academy Postdoctoral Fellowship PF2\180009, awarded to C.S.F.

References

Antony, J.W., Ferreira, C.S., Norman, K.A., Wimber, M., 2017. Retrieval as a Fast Route to Memory Consolidation. Trends Cogn. Sci. 21, 573–576. https://doi.org/10.1016/j.tics.2017.05.001

Antony, J.W., Paller, K.A., 2017. Hippocampal Contributions to Declarative Memory Consolidation During Sleep, in: The Hippocampus from Cells to Systems. Springer International Publishing, Cham, pp. 245–280. https://doi.org/10.1007/978-3-319-50406-3_9

Bäuml, K.-H., Holterman, C., Abel, M., 2014. Sleep can reduce the testing effect: It enhances recall of restudied items but can leave recall of retrieved items unaffected. J. Exp. Psychol. Learn. Mem. Cogn. 40, 1568–1581. https://doi.org/10.1037/xlm0000025

Binder, J.R., Desai, R.H., 2011. The neurobiology of semantic memory. Trends Cogn. Sci. 15, 527–536. https://doi.org/10.1016/j.tics.2011.10.001

Binder, J.R., Desai, R.H., Graves, W.W., Conant, L.L., 2009. Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. Cereb. Cortex 19, 2767–2796. https://doi.org/10.1093/cercor/bhp055

Brett, M., Anton, J.-L., Valabregue, R., Poline, J.-B., 2002. Region of Interest Analysis Using the MarsBar Toolbox for SPM 99. NeuroImage, 16, 497.

Brodeur, M.B., Guérard, K., Bouras, M., 2014. Bank of Standardized Stimuli (BOSS) Phase II: 930 New Normative Photos. PLoS One 9, e106953. https://doi.org/10.1371/journal.pone.0106953

Brodt, S., Gais, S., Beck, J., Erb, M., Scheffler, K., Schönauer, M., 2018. Fast track to the neocortex: A memory engram in the posterior parietal cortex. Science 362, 1045–1048. https://doi.org/10.1126/SCIENCE.AAU2528

Brodt, S., Pöhlchen, D., Flanagin, V.L., Glasauer, S., Gais, S., Schönauer, M., 2016. Rapid and independent memory formation in the parietal cortex. Proc. Natl. Acad. Sci. 113, 13251-13256. https://doi.org/10.1073/pnas.1605719113

Cairney, S.A., Durrant, S.J., Musgrove, H., Lewis, P.A., 2011. Sleep and environmental context: Interactive effects for memory. Exp. Brain Res. 214, 83–92. https://doi.org/10.1007/s00221-011-2808-7

Carpenter, S.K., 2009. Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. J. Exp. Psychol. Learn. Mem. Cogn. 35, 1563–1569. https://doi.org/10.1037/a0017021

Chan, J.C., McDermot, K.B., 2007. The testing effect in recognition memory: a dual process account. J. Exp. Psychol. Learn. Mem. Cogn. 33, 431-437. http://dx.doi.org/10.1037/0278-7393.33.2.431

Clarke, A., Devereux, B.J., Randall, B., Tyler, L.K., 2015. Predicting the Time Course of Individual Objects with MEG. Cereb. Cortex 25, 3602–3612. https://doi.org/10.1093/cercor/bhu203

Coppens, L.C., Verkoeijen, P.P.J.L., Rikers, R.M.J.P., 2011. Learning Adinkra symbols: The effect of testing. J. Cogn. Psychol. 23, 351–357. https://doi.org/10.1080/20445911.2011.507188

Dijkstra, N., Bosch, S.E., van Gerven, M.A.J., 2019. Shared Neural Mechanisms of Visual Perception and Imagery. Trends Cogn. Sci. 23, 423–434. https://doi.org/10.1016/j.tics.2019.02.004

Dudai, Y., Karni, A., Born, J., 2015. The Consolidation and Transformation of Memory. Neuron 88, 20–32. https://doi.org/10.1016/j.neuron.2015.09.004

Favila, S.E., Samide, R., Sweigart, S.C., Kuhl, B.A., 2018. Parietal Representations of Stimulus Features Are Amplified during Memory Retrieval and Flexibly Aligned with Top-Down Goals. J. Neurosci. 38, 7809–7821. https://doi.org/10.1523/JNEUROSCI.0564-18.2018

Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Van Essen, D.C., Raichle, M.E., 2005. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc. Natl. Acad. Sci. 102, 9673-9678. https://doi.org/10.1073/pnas.0504136102

Gais, S., Albouy, G., Boly, M., Dang-Vu, T.T., Darsaud, A., Desseilles, M., Rauchs, G., Schabus, M., Sterpenich, V., Vandewalle, G., Maquet, P., Peigneux, P., 2007. Sleep transforms the cerebral trace of declarative memories. Proc. Natl. Acad. Sci. 104, 18778–18783. https://doi.org/10.1073/pnas.0705454104

Gilboa, A., Marlatte, H., 2017. Neurobiology of Schemas and Schema-Mediated Memory. Trends Cogn. Sci. 21, 618–631. https://doi.org/10.1016/j.tics.2017.04.013

Himmer, L., Schönauer, M., Heib, D.P.J., Schabus, M., Gais, S., 2019. Rehearsal initiates systems memory consolidation, sleep makes it last. Sci. Adv. 5, eaav1695. https://doi.org/10.1126/sciadv.aav1695

Jonker, T.R., Dimsdale-Zucker, H., Ritchey, M., Clarke, A., Ranganath, C., 2018. Neural reactivation in parietal cortex enhances memory for episodically linked information. Proc. Natl. Acad. Sci. 115, 11084-11089. https://doi.org/10.1073/pnas.1800006115

Jurewicz, K., Cordi, M.J., Staudigl, T., Rasch, B., 2016. No Evidence for Memory Decontextualization across One Night of Sleep. Front. Hum. Neurosci. 10, 7. https://doi.org/10.3389/fnhum.2016.00007

Kaiser, D., Azzalini, D.C., Peelen, M. V., 2016. Shape-independent object category responses revealed by MEG and fMRI decoding. J. Neurophysiol. 115, 2246–2250. https://doi.org/10.1152/jn.01074.2015

Karlsson Wirebring, L., Wiklund-Hörnqvist, C., Eriksson, J., Andersson, X., 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. https://doi.org/10.1523/JNEUROSCI.3550-14.2015

Karpicke, J.D., Lafayette, W., States, U., 2017. Retrieval-Based Learning : A Decade of Progress, Third Edit. ed, Learning and Memory: A Comprehensive Reference. Elsevier. https://doi.org/10.1016/B978-0-12-805159-7.02023-4

Kriegeskorte, N., Goebel, R., Bandettini, P., 2006. Information-based functional brain mapping. Proc. Natl. Acad. Sci. 103, 3863-3868. https://doi.org/10.1073/pnas.0600244103

Kriegeskorte, N., Mur, M., Bandettini, P., 2008. Representational similarity analysis – connecting the branches of systems neuroscience. Front. Syst. Neurosci. 2:4. 1-28. https://doi.org/10.3389/neuro.06.004.2008

Kuhl, B.A., Chun, M.M., 2014. Successful Remembering Elicits Event-Specific Activity Patterns in Lateral Parietal Cortex. J. Neurosci. 34, 8051-8060. https://doi.org/10.1523/JNEUROSCI.4328-13.2014

Lee, H., Samide, R., Richter, F.R., Kuhl, B.A., 2017. Decomposing parietal memory reactivation to predict consequences of remembering. Cereb. Cortex bhy200. https://doi.org/10.1101/208678

Linden, M. van der, Berkers, R.M.W.J., Morris, R.G.M., Fernández, G., 2017. Angular Gyrus Involvement at Encoding and Retrieval Is Associated with Durable But Less Specific Memories. J. Neurosci. 37, 9474–9485. https://doi.org/10.1523/JNEUROSCI.3603-16.2017

Lutz, N.D., Diekelmann, S., Hinse-Stern, P., Born, J., Rauss, K., 2017. Sleep supports the slow abstraction of gist from visual perceptual memories. Sci. Rep., 7:42950. https://doi.org/10.1038/srep42950

Macey, P.M., Macey, K.E., Kumar, R., Harper, R.M., 2004. A method for removal of global effects from fMRI time series. Neuroimage 22, 360–366. https://doi.org/10.1016/J.NEUROIMAGE.2003.12.042

Maldjian, J.A., Laurienti, P.J., Kraft, R.A., Burdette, J.H., 2003. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. Neuroimage, 19, 1233-1239. https://doi.org/10.1016/S1053-8119(03)00169-1

Martin, C.B., Douglas, D., Newsome, R.N., Man, L.L., Barense, M.D., 2018. Integrative and distinctive coding of visual and conceptual object features in the ventral visual stream. Elife 7:e31873. https://doi.org/10.7554/eLife.31873

McClelland, J.L., McNaughton, B.L., O'Reilly, R.C., 1995. 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. https://doi.org/10.1037/0033-295X.102.3.419

Nieuwenhuis I.L., Takashima A., 2011. The role of ventromedial prefrontal cortex in memory consolidation. Behav. Brain Research, 218, 325-34. https://doi.org/10.1016/j.bbr.2010.12.009

Nili, H., Wingfield, C., Walther, A., Su, L., Marslen-Wilson, W., Kriegeskorte, N., 2014. A Toolbox for Representational Similarity Analysis. PLoS Comput. Biol. 10. https://doi.org/10.1371/journal.pcbi.1003553

Peirce, J.W., 2006. PsychoPy—Psychophysics software in Python. J. Neurosci. Methods 162, 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Pyc, M.A., Rawson, K.A., 2010. Why testing improves memory: mediator effectiveness hypothesis. Science 330, 335. https://doi.org/10.1126/science.1191465

Rasch, B., Büchel, C., Gais, S., Born, J., 2007. Odor Cues During Slow-Wave Sleep Prompt Declarative Memory Consolidation. Science 315, 1423–1426. https://doi.org/10.1126/science.1134457

Reagh, Z.M., Ranganath, C., 2018. What does the functional organization of cortico-hippocampal networks tell us about the functional organization of memory? Neurosci. Lett. 680, 69–76. https://doi.org/10.1016/j.neulet.2018.04.050

Richards, B.A., Frankland, P.W., 2017. The persistance and Transience of Memory. Neuron. 94, 1071-1086. https://doi.org/10.1016/j.neuron.2017.04.037

Richards, B.A., Xia, F., Santoro, A., Husse, J., Woodin, M.A., Josselyn, S.A., Frankland, P.W., 2014. Patterns across multiple memories are identified over time. Nat. Neurosci. 17, 981-986. https://doi.org/10.1038/nn.3736

Robin, J., Moscovitch, M., 2017. Details, gist and schema: Hippocampal–neocortical interactions underlying recent and remote episodic and spatial memory. Curr. Opin. in Behav. Sci. 17, 114-123. http://doi.org/10.1016/j.cobeha.2017.07.016

Roediger, H.L., Karpicke, J.D., 2006. The Power of Testing Memory Basic Research and Implications for Educational Practice. Perspect. Psych. Sci. 1, 181-210. https://doi.org/10.1111/j.1745-6916.2006.00012.x

Roediger III, H.L., Jacoby, J.D., McDermott, K.B., 1996. Misinformation Effects in Recall: Creating False Memories through Repeated Retrieval. J. Mem. Lang. 35, 300–318. https://doi.org/10.1006/jmla.1996.0017

Rowland, C.A., DeLosh, E., 2014. Benefits of testing for nontested information: Retrieval-induced facilitation of episodically bound material. Psychon Bull Rev., 21, 1516-1523. https://doi.org/10.3758/s13423-014-0625-2

Rugg, M.D., Vilberg, K.L., 2012. Brain Networks Underlying Episodic Memory Retrieval. Curr. Opin. Neurobiol. 23, 255-260. https://doi.org/10.1016/j.conb.2012.11.005

Schapiro, A.C., McDevitt, E.A., Chen, L., Norman, K.A., Mednick, S.C., Rogers, T.T., 2017. Sleep Benefits Memory for Semantic Category Structure while Preserving Exemplar-Specific Information. Sci. Rep. 7, 1–22. https://doi.org/10.1038/s41598-017-12884-5

Schlichting, M.L., Mumford, J.A., Preston, A.R., 2015. Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. Nat. Commun. 6, 1–10. https://doi.org/10.1038/ncomms9151

Sekeres, M., Winocur, G., Moscovitch, M., 2018. The hippocampus and related neocortical structures in memory transformation. Neurosci. Letters 680, 39-53. https://doi.org/10.1016/j.neulet.2018.05.006

Sommer, T., 2016. The Emergence of Knowledge and How it Supports the Memory for Novel Related Information. Cereb. Cortex 27, bhw031. https://doi.org/10.1093/cercor/bhw031

Takashima, A., Nieuwenhuis, I.L.C., Jensen, O., Talamini, L., Rijpkems, M, Fernández, G., 2009. Shift from hippocampal to neocortical centered retrieval network with consolidation. J. Neurosci. 29, 10087-10093. https://doi.org/10.1523/JNEUROSCI.0799-09.2009

Takashima, A., Petersson, K.M., Rutters, F., Tendolkar, I., Jensen, O., Zwarts, M.J., Mcnaughton, B.L., Fernández, G., 2006. Declarative memory consolidation in humans: A prospective functional magnetic resonance imaging study. Proc. Natl. Acad. Sci. 103, 756-761. https://doi.org/10.1073/pnas.0507774103

Tompary, A., Davachi, L., 2017. Consolidation Promotes the Emergence of Representational Overlap in the Hippocampus and Medial Prefrontal Cortex. Neuron 96, 228-241. https://doi.org/10.1016/j.neuron.2017.09.005

Toppino, T.C., Cohen, M.S., 2009. The Testing Effect and the Retention Interval Questions and Answers. Exp. Psych. 56, 252-257. https://doi.org/10.1027/1618-3169.56.4.252

van den Broek, G.S.E., Segers, E., Takashima, A., Verhoeven, L., 2014. Do testing effects change over time? Insights from immediate and delayed retrieval speed. Memory 22, 803–812. https://doi.org/10.1080/09658211.2013.831455

Wagner, I.C., van Buuren, M., Kroes, M.C., Gutteling, T.P., van der Linden, M., Morris, R.G., Fernández, G., 2015. Schematic memory components converge within angular gyrus during retrieval. Elife 4. https://doi.org/10.7554/eLife.09668

Wardle, S.G., Ritchie, J.B., 2014. Can object category-selectivity in the ventral visual pathway be

explained by sensitivity to low-level image properties? J. Neurosci. 34, 14817–14819. https://doi.org/10.1523/JNEUROSCI.3566-14.2014

Wimber, M., Alink, A., Charest, I., Kriegeskorte, N., Anderson, M.C., 2015. Retrieval induces adaptive forgetting of competing memories via cortical pattern suppresion. Nat. Neurosci. 18, 582-589. https://doi.org/10.1038/nn.3973

Wing, E.A., Marsh, E.J., Cabeza, R., 2013. Neural correlates of retrieval-based memory enhancement: An fMRI study of the testing effect. Neuropsychologia 51, 2360-2370. https://doi.org/10.1016/j.neuropsychologia.2013.04.004

Winocur, G., Moscovitch, M., Bontempi, B., 2010. Memory formation and long-term retention in humans and animals: Convergence towards a transformation account of hippocampal-neocortical interactions. Neuropsychologia 48, 2339–2356. https://doi.org/10.1016/j.neuropsychologia.2010.04.016

Xiao, J., Hays, J., Ehinger, K., Oliva, A., Torralba, A., 2010. SUN Database: Large-scale Scene Recognition from Abbey to Zoo., in: IEEE Conference on Computer Vision and Pattern Recognition.

Ye, Z., Zhu, B., Zhuang, L., Lu, Z., Chen, C., Xue, G., 2016. Neural Global Pattern Similarity Underlies True and False Memories. J. Neurosci. 36, 6792–802. https://doi.org/10.1523/JNEUROSCI.0425-16.2016

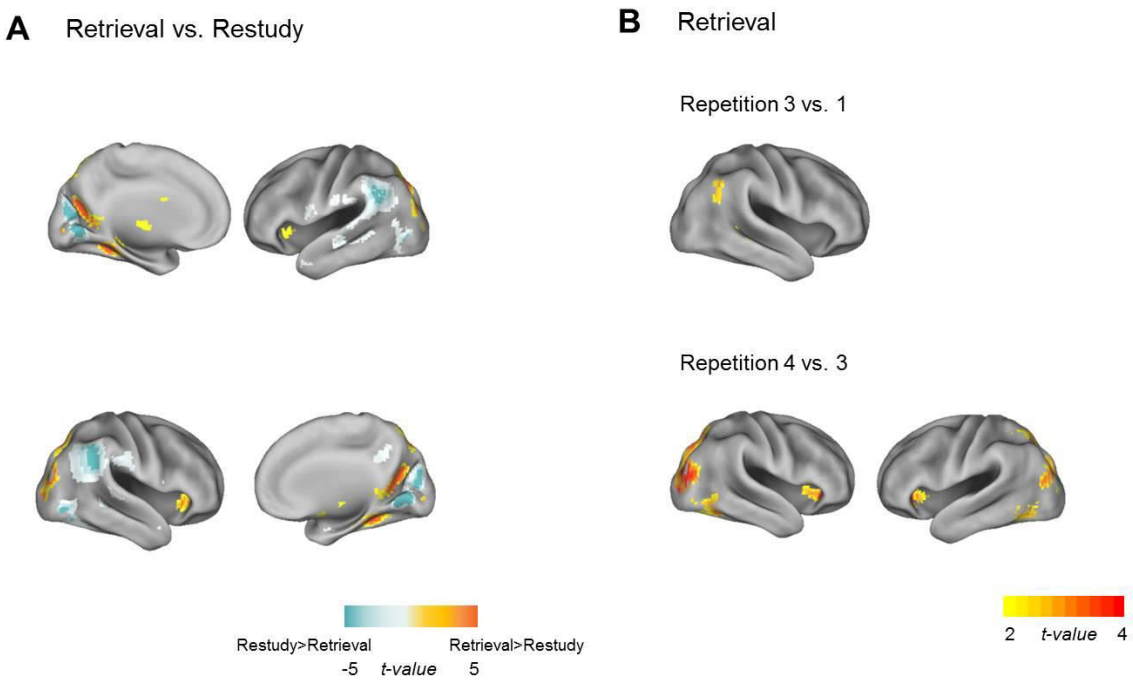
Yushkevich, P.A., Piven, J., Hazlett, H.C., Smith, R.G., Ho, S., Gee, J.C., Gerig, G., 2006. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. Neuroimage 31, 1116-1128. https://doi.org/10.1016/j.neuroimage.2006.01.015

Retrieval aids the creation of a generalised memory trace and strengthens episode-unique information

Ferreira, C.S.*, Charest, I., &Wimber, M.

School of Psychology and Centre for Human Brain Health

University of Birmingham, Edgbaston, B15 2TT, United Kingdom



Supplementary Figure 1: Whole-brain univariate analyses. **A.** Clusters that significantly differ between retrieval and restudy practice (across all repetitions). Areas more active for retrieval than restudy are shown in orange, and the opposite contrast is shown in blue, both at p<.001 *uncorrected* (k=10). Consistent with results found in the imagery literature (e.g. Dijkstra, Bosch, & van Gerven, *TICS,* 2019), while restudy engages relatively earlier visual and more inferior parietal areas, retrieval tends to engage later visual and superior parietal lobe areas. This is in line with the idea that restudy re-imposes a specific visual representation on the brain, while retrieval engages higher-level, more abstract visual representations. **B.** Retrieval specific univariate results. The upper panel shows regions whose activity increased within session (from the first to the third retrieval repetition). The lower panel shows regions that are more engaged across sessions (higher activity on the fourth than on the third repetition).

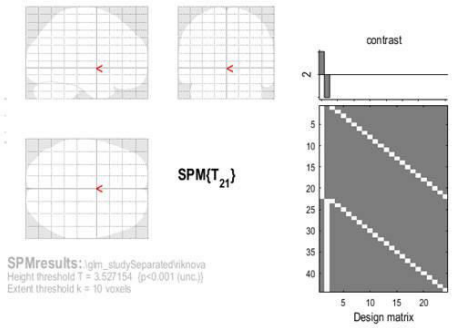
Supplementary Table 1. Whole brain univariate analyses.

	Hemisphere	Cluster label(s)	XYZ	t-value	k
Retrieval>Restudy					
	L	Posterior cingulate Superior parietal	-18 -61 20	6.57	393
	R	Middle occipital gyrus Parahippocampal gyrus	30 -40 -10	5.89	93
	L	Parahippocampal gyrus	-21 -40 -10	5.69	124
	R	Parietal Lobe	18 -55 20	5.36	137
	R	Middle temporal gyrus Lingual gyrus	39 -79 26	5.16	141
	R	Middle occipital gyrus Precuneus	21 -73 50	5.14	98
	L	Cuneus	-12 -91 5	4.42	13
	R	Inferior frontal gyrus	33 26 -1	4.34	25
Restudy>Retrieval					
	L	Lingual gyrus			
	R	Lingual gyrus	-3 -70 5	9.31	624
	R	Cuneus			
	R	Inferior occipital gyrus	45 -70 -7	4.91	119
	L	Supramarginal gyrus Middle temporal gyrus	-54 -52 29	4.86	295
	R	Supramarginal gyrus	54 -43 32	4.81	298
	L	Inferior temporal gyrus	-45 -73 -1	4.74	44
	R	Middle temporal gyrus	63 -34 -1	4.66	52
	R	Culmen	39 -46 -19	3.64	12
	R	Postcentral gyrus	60 -19 29	3.62	26
Retrieval 3>Retrieval 1					
	R	Middle temporal gyrus	60 -28 -10	3.97	13
	R	Superior temporal gyrus Inferior parietal	60 -58 23	3.55	20
Retrieval 4>Retrieval 3					
	R	Middle temporal gyrus	36 -76 20	5.06	152
	L	Superior occipital lobe Middle temporal gyrus	-30 -73 23	4.31	119
	R	Middle occipital gyrus Inferior occipital gyrus	33 -85 2	4.17	28
	L	Inferior frontal gyrus	-33 29 -1	4.16	28
	R	Inferior frontal gyrus	33 32 -7	4.07	31
	R	Fusiform gyrus Culmen	27 -49 -10	4.00	89
	R	Superior parietal	15 -61 56	3.55	13
	R	Superior parietal	27 -73 44	3.44	10

A Univariate Analyses

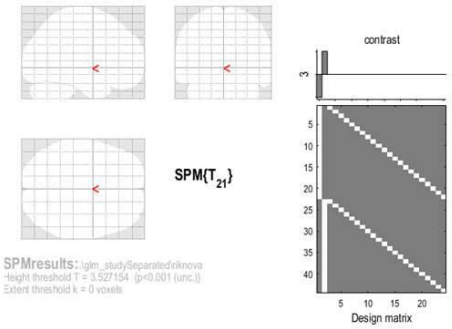
Later Retrieved > Later Restudied

p<.001 *uncorr*, k=10

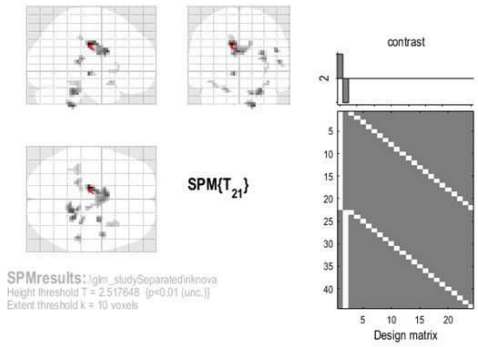


Later Restudied > Later Retrieved

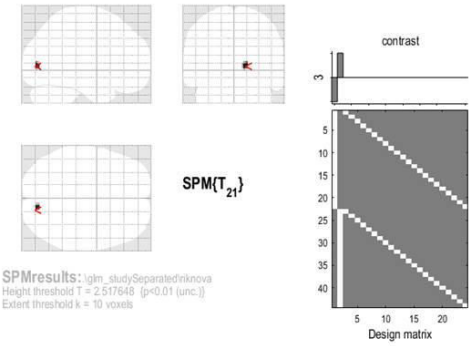
p<.001 *uncorr*, k=10

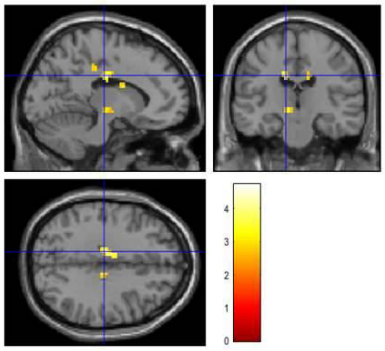


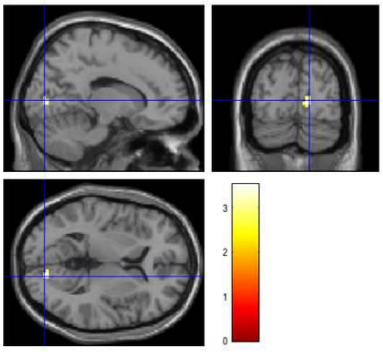
p<.01 *uncorr*, k=10



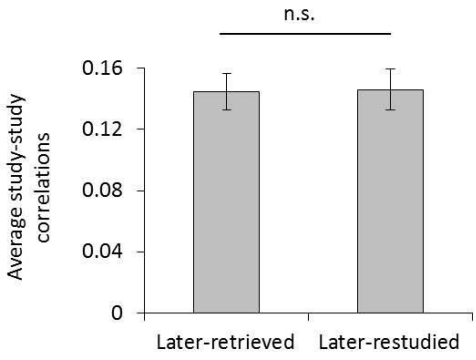
p<.01 *uncorr*, k=10







B Multivariate Analyses

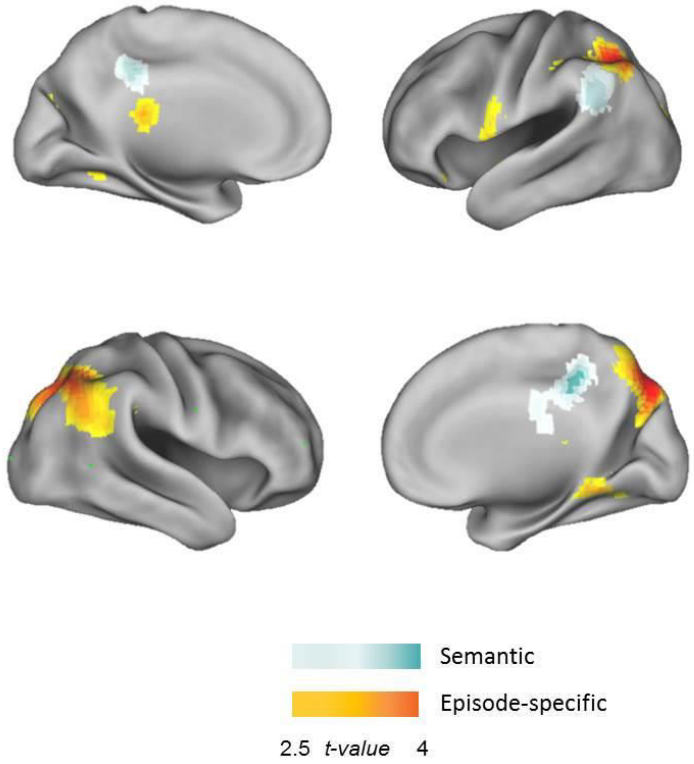


Supplementary Figure 2: Uni- and multivariate analyses at study, comparing to-be-retrieved and to-be-restudied items in order to check whether differences already existed at this stage. **A.** Univariate analyses at study. Left panel depicts differences at study between to-be-retrieved>to-be-restudied. The right panel presents the opposite contrast (to-be-restudied>to-be-retrieved). The upper glass brains show that there were no significant differences at study between subsequently retrieved and subsequently restudied items at a threshold of p<.001 *uncorrected*, *k*=10 voxels. At a more lenient threshold of *p*<.01 (lower panels) some clusters can be observed, although they appear to be artefacts. **B.** Multivariate analyses at study. Similarity (average correlations) between neural patterns of later retrieved and later restudied items at study did not reveal any significant difference between items.

Supplementary Table 2. Studies used to build the functional mPFC ROI, with their respective peak coordinates.

Study	BA	Hemisphere	MNI peak [X Y Z]		
Gais et al. (*PNAS,* 2007)	11	L	-6	26	-10
	11	L	-6	36	-18
Takashima et al. (*PNAS,*2006)	32	L	-2	32	-10
Takashima et al. (*Learning and Memory*, 2007)	32	R	2	48	2
	32	L	-12	48	-4
	10	R	6	52	8
	10	L	-4	52	6
Sterpenich et al. (*JNeuro*, 2009)	10	R	34	50	8
	10	L	-35	52	2
	24	R	6	38	6
Mack et al. (*PNAS*, 2016)	32	R	10	43	-6
Mack et al. (*International Workshop on Pattern Recognition*, 2017)	11	L	-5	51	20
Kumaran et al. (*Neuron*, 2009)	10	R	6	57	-6
van Kesteren et al. (*Neuropsychologia,* 2013)	32	R	2	46	0
	32	L	-2	40	2
van Kesteren et al. (*JoCN*, 2014)	32	R	14	42	4
Schlichting et al. (*CurrOpinBehav Sci.,*2015)	11	L	-8	44	-17
Bowman and Zeithamova (*JNeuro*, 2018)	11		0	44	-16

Wing et al. (*Neuropsychologia*, 2013)	11	R	23	38	-19
	11	R	11	41	-15
Karlsson Wirebring et al. (*JNeuro,* 2015)	10		0	60	14



Supplementary Figure 3: Overlap between semantic and episode-specific searchlights (left hemisphere on top, right on the bottom). Semantic searchlight clusters are depicted in blue tones, whereas episodic-specific clusters are presented in orange hues. For both searchlights, clusters obtained from the interaction between cycle and condition ([Retrieval3>4]>[Restudy3>4]) and are presented at a threshold of p<.001 *uncorrected* and k=10 voxels. Note that there is little to no overlap between the two searchlight results. Whereas the semantic searchlight evidences the role of medial and inferior parietal areas that have been argued to support integrative retrieval and rule based associations (Wagner et al., 2015), episode-specific results highlight the role of superior parietal regions, shown to track the reinstatement of specific memories (Karlsson Wirebring et al., 2015).

References

Bowman, C.R., Zeithamova, D., 2018. Abstract Memory Representations in the Ventromedial Prefrontal Cortex and Hippocampus Support Concept Generalization. J. Neurosci. 38, 2605–2614. https://doi.org/10.1523/JNEUROSCI.2811-17.2018

Dijkstra, N., Bosch, S.E., van Gerven, M.A.J., 2019. Shared Neural Mechanisms of Visual Perception and Imagery. Trends Cogn. Sci. 23, 423–434. https://doi.org/10.1016/j.tics.2019.02.004

Gais, S., Albouy, G., Boly, M., Dang-Vu, T.T., Darsaud, A., Desseilles, M., Rauchs, G., Schabus, M., Sterpenich, V., Vandewalle, G., Maquet, P., Peigneux, P., 2007. Sleep transforms the cerebral trace of declarative memories. Proc. Natl. Acad. Sci. 104, 18778–18783. https://doi.org/10.1073/pnas.0705454104

Karlsson Wirebring, L., Wiklund-Hörnqvist, C., Eriksson, J., Andersson, X., 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. https://doi.org/10.1523/JNEUROSCI.3550-14.2015

Kumaran, D., Summerfield, J.J., Hassabis, D., Maguire, E.A., 2009. Tracking the Emergence of Conceptual Knowledge during Human Decision Making. Neuron 63, 889–901. https://doi.org/10.1016/j.neuron.2009.07.030

Mack, M.L., Love, B.C., Preston, A.R., 2016. Dynamic updating of hippocampal object representations reflects new conceptual knowledge. Proc. Natl. Acad. Sci. 113, 13203–13208. https://doi.org/10.1073/pnas.1614048113

Mack, M.L., Preston, A.R., Love, B.C., 2017. Medial prefrontal cortex compresses concept representations through learning, in: 2017 International Workshop on Pattern Recognition in Neuroimaging (PRNI). IEEE, pp. 1–4. https://doi.org/10.1109/PRNI.2017.7981500

Schlichting, M.L., Mumford, J.A., Preston, A.R., 2015. Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. Nat. Commun. 6, 1–10. https://doi.org/10.1038/ncomms9151

Sterpenich, V., Albouy, G., Darsaud, A., Schmidt, C., Vandewalle, G., Vu, T.T.D., Desseilles, M., Phillips, C., Degueldre, C., Balteau, E., Collette, F., Luxen, A., Maquet, P., 2009. Sleep Promotes the Neural Reorganization of Remote Emotional Memory. J. Neurosci. 29, 5143–5152. https://doi.org/10.1523/JNEUROSCI.0561-09.2009

Takashima, A., Nieuwenhuis, I.L.C., Rijpkema, M., Petersson, K.M., Jensen, O., Fernández, G., 2007. Memory trace stabilization leads to large-scale changes in the retrieval network: a functional MRI study on associative memory. Learn. Mem. 14, 472–9. https://doi.org/10.1101/lm.605607

Takashima, A., Petersson, K.M., Rutters, F., Tendolkar, I., Jensen, O., Zwarts, M.J., Mcnaughton, B.L., Fernández, G., 2006. Declarative memory consolidation in humans: A prospective functional magnetic resonance imaging study. Proc. Natl. Acad. Sci. 103, 756-761. https://doi.org/10.1073/pnas.0507774103

van Kesteren, M.T.R., Beul, S.F., Takashima, A., Henson, R.N., Ruiter, D.J., Fernández, G., 2013. Differential roles for medial prefrontal and medial temporal cortices in schema-dependent encoding: From congruent to incongruent. Neuropsychologia 51, 2352–2359. https://doi.org/10.1016/j.neuropsychologia.2013.05.027

van Kesteren, M.T.R., Rijpkema, M., Ruiter, D.J., Morris, R.G.M., Fernández, G., 2014. Building on Prior Knowledge: Schema-dependent Encoding Processes Relate to Academic Performance. J. Cogn. Neurosci. 26, 2250–2261. https://doi.org/10.1162/jocn_a_00630

Wagner, I.C., van Buuren, M., Kroes, M.C., Gutteling, T.P., van der Linden, M., Morris, R.G., Fernández, G., 2015. Schematic memory components converge within angular gyrus during retrieval. Elife 4. https://doi.org/10.7554/eLife.09668

Wing, E.A., Marsh, E.J., Cabeza, R., 2013. Neural correlates of retrieval-based memory enhancement: An fMRI study of the testing effect. Neuropsychologia 51, 2360–2370. https://doi.org/10.1016/j.neuropsychologia.2013.04.004