PHILOSOPHICAL TRANSACTIONS B

royalsocietypublishing.org/rstb

Research

Cite this article: Raykov PP, Gurunandan K, Moniri M, Jain A, Greve A, Henson RN. 2026 Predictions and declarative memory encoding: two fMRI paradigms provide slim pickings for SLIMM. *Phil. Trans. R. Soc. B* **381**: 20250250. https://doi.org/10.1098/rstb.2025.0250

Received: 21 August 2025
Accepted: 24 February 2026

One contribution of 18 to a theme issue 'The role of hippocampal predictions in cognition: bridging perception and memory'.

Subject Areas:
cognition, neuroscience

Keywords:
prediction, memory, expectancy, hippocampus, SLIMM, MPFC

Author for correspondence:
Petar P. Raykov
e-mail: Petar.Raykov@mrc-cbu.cam.ac.uk

†These authors contributed equally to the study.

Supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.8547874.

THE ROYAL SOCIETY PUBLISHING

Predictions and declarative memory encoding: two fMRI paradigms provide slim pickings for SLIMM

Petar P. Raykov[1,†], Kshipra Gurunandan[1,†], Mahtab Moniri[1], Avnee Jain[1], Andrea Greve[1] and Richard N. Henson[1,2]

[1]MRC Cognition and Brain Sciences Unit, Cambridge CB2 7EF, UK
[2]Department of Psychiatry, University of Cambridge, Cambridge, UK

PPR, 0000-0002-2124-2995

Memory is often better for both highly unexpected and highly expected information. To explain this U-shape relationship between memory and expectancy, the SLIMM framework (schema-linked interactions between medial prefrontal and medial temporal regions) appeals to two memory systems: one involving the hippocampus and another involving the medial prefrontal cortex (MPFC). Specifically, SLIMM predicts that these two brain regions support opposite ends of the U-shape, with the hippocampus supporting memory for unexpected events and the MPFC underpinning memory for expected events. We tested these predictions in two functional magnetic resonance imaging (fMRI) paradigms: one examined memory for the location of objects within scenes (e.g. a toothbrush in a bathroom, either on a sink, as expected, or on top of a bin, as not expected); the second tested memory for words that completed well known idioms either in an expected or unexpected fashion (e.g. 'Don't count your chickens before they … hatch/cook'). Both paradigms replicated the predicted U-shaped relationship between expectancy and memory performance. However, neither study found the predicted interaction between expectancy and subsequent memory (remembered versus forgotten trials) in terms of fMRI activation at encoding in the *a priori* defined hippocampal and MPFC regions of interest. We discuss the implications of these findings for SLIMM, and how our results fit within the broader literature on expectancy and memory.

This article is part of the theme issue 'The role of hippocampal predictions in cognition: bridging perception and memory'.

1. Introduction

Memory shapes our understanding of the world, enabling us to learn from past experiences and anticipate future ones. One factor thought to influence learning and memory is prediction error—when incoming information deviates from what we expect (e.g. see for reviews [1–4]). While this view is supported by some studies [5–12], it contrasts with extensive evidence that memory is often better for expected information [13–24]. This puzzle may be reconciled by recent behavioural work, which has demonstrated a nonlinear, U-shaped relationship between expectancy and memory, with advantages for both highly unexpected and expected events [25,26]. However, it remains unclear what brain mechanisms support this behavioural pattern. Here, we introduce two novel functional magnetic resonance imaging (fMRI) paradigms designed to compare brain correlates during the encoding that

predict subsequent memory—or 'subsequent memory effects' (SMEs; [27,28])—as a function of expectancy.

The SLIMM framework (schema-linked interactions between medial prefrontal and medial temporal regions; [29]) proposes two brain systems to explain the behavioural U-shape: the medial prefrontal cortex (MPFC) is hypothesized to support encoding of expected information, by rapidly integrating it into existing knowledge structures, while the medial temporal lobe (MTL)—particularly the hippocampus—are hypothesized to support encoding of unexpected, schema-incongruent information. This framework predicts that MPFC should show a larger SME for expected events, while MTL regions should show a larger SME for unexpected events. Additionally, SLIMM predicts that MPFC inhibits MTL in the presence of a schema, and so MPFC–MTL functional connectivity should also vary with expectancy.

Previous fMRI studies have provided some support for SLIMM's predictions. The most relevant is from van Kesteren *et al*. [30], who found greater SMEs in MPFC for congruent object–scene pairs (e.g. a tennis racquet and a tennis court) and greater SMEs in MTL for incongruent pairings (e.g. a tennis racquet and a swimming pool). However, participants' memory performance increased linearly from incongruent to intermediate to congruent pairings. One reason for this lack of the behavioural U-shape predicted by SLIMM could be that participants adopted a 'generate-and-recognize' strategy, in which they generated scenes congruent with the object presented at test, and then tried to assess whether they had seen those scenes. Even though this strategy would not uniquely identify the correct scene name (given that one of the two foils was also congruent), it would still help discriminate them from the other foils in the case of congruent pairs, but not in the case of incongruent pairs. This is why previous studies revealing a U-shape in memory performance have controlled for such retrieval-based strategies [25,26].

Other fMRI studies have provided indirect support for SLIMM. For example, Brod *et al*. [31] tested learning in medical students and found that the SME in the hippocampus decreased with learning of schema-based associations (face–diagnosis pairs), but not for arbitrary associations (face–name pairs) (see also [19]). More generally, Kumaran & Maguire [32] demonstrated increased hippocampal activity for unexpected events. They showed sequences of eight objects in an ABCD–ABCD format, and found greater hippocampal activation to the seventh object when it was unexpected (based on the first four, e.g. ABCD–AB**D**C) than when it was expected (i.e. ABCD–AB**C**D). Although they did not test subsequent memory, this finding is consistent with claims that the hippocampus is a generic mismatch or novelty detector [7,12,33–35]. Given this evidence, our study specifically focuses on the hippocampus here, rather than MTL more generally.

The role of the MPFC in memory is well established in rodent and patient studies (see also [36]). For example, MPFC lesions in rats eliminate the usual schema advantage for spatial learning [37], and patients with MPFC damage often struggle to use script knowledge appropriately [17,38–43]. Human fMRI work similarly implicates the MPFC in processing schema-relevant information [31,44–49]. While expectancy reliably activates the MPFC, fewer studies have tested how SMEs vary with expectancy. One exception is van Kesteren *et al*. [50], who found greater MPFC activation during encoding of schema-consistent versus inconsistent facts (see also [23,30,31,51,52]).

Finally, several studies have reported increased functional connectivity between the MPFC and MTL for expected or unexpected information [51,53,54]. For instance, Bein *et al*. [51] found enhanced MPFC–hippocampus connectivity for unexpected items that were later remembered (see also [50]). Conversely, other studies have observed stronger MPFC–hippocampal connectivity for expected items [54], or failed to find connectivity differences across varying levels of expectancy [55]. SLIMM predicts that the activation of a schema should cause the MPFC to inhibit encoding processes in the MTL, such that MPFC–hippocampal connectivity should be reduced during encoding of high-expectancy items.

In the present study, we ran two incidental encoding paradigms that manipulated expectancy with the same group of participants, to examine whether SLIMM's predictions generalized across different types of stimuli and tasks. To minimize retrieval-based strategies, like the generate-and-recognize strategy mentioned above, or simple familiarity of repeated objects, we used *N*-alternative forced choice (AFC) memory tasks, with the AFC lures chosen to have comparable levels of expectancy (semantic congruency) to the target, and to be matched for its familiarity.

In the encoding phase of paradigm 1 (figure 1, top left), participants had to find objects placed in locations that varied in expectancy based on prior knowledge, e.g. an alarm clock on a bedside shelf, as expected, or on top of a pillow, as unexpected. This was a two-dimensional (2D) adaptation of a virtual reality paradigm that we used previously to demonstrate a U-shaped function of expectancy [26]. Memory for the location of the object was tested with a 3-AFC task (figure 1, bottom left), where the target scene was shown from a different angle (to minimize familiarity of the exact image) and the foils contained the target object in different locations, but matched for expectancy. For instance, if a toothbrush were seen in an incongruent location (e.g. on top of a bin), the lures would also show the toothbrush in other, equally incongruent locations (e.g. on top of the toilet, or on the floor). This prevented participants from guessing the correct answer based on remembering whether the trial had been surprising or expected. After their 3-AFC responses, participants provided subjective memory ratings, which were used to test a secondary prediction of SLIMM, namely that recognition memory performance for unexpected items, but not expected items, is driven by recollection processes rather than familiarity. Finally, participants rated the expectancy of each object–location they had seen during the encoding phase, and these ratings were used to obtain participant-specific levels of expectancy.

In the encoding phase of paradigm 2 (figure 1, top right), participants read sentences (two-thirds of which were common idioms, and one-third were neutral sentences) that were missing their final word. The final word was then depicted by an image that showed either expected, unexpected or neutral content (e.g. 'I've always believed a bird in the hand is worth two in the … bush/car'). Participants evaluated the congruency of the picture and sentence. In the retrieval phase (figure 1, bottom right), they completed a 3-AFC task that included foil images depicting the same word (e.g. different pictures of bushes, so matched for congruency), and that had appeared in other trials/conditions (so were matched for familiarity). Source memory for the sentences (the speaker of the sentence) was also tested after the study phase to again examine the secondary prediction of SLIMM, namely that source memory would be better for unexpected sentences.

To test the key predictions of SLIMM, we examined mean activation for each trial-type across voxels within two, *a priori* regions of interests (ROIs): the hippocampus and MPFC (defined from independent data). More specifically, for each ROI, we predicted an interaction between expectancy and the SME (the difference in activation according to whether a trial was later remembered or forgotten), but in opposite directions: greater encoding-related activation for unexpected than expected trials in the hippocampus, and greater encoding-related activation for expected than unexpected trials in MPFC. Furthermore, we tested SLIMM's prediction that the MPFC inhibits hippocampus during encoding of expected trials using beta-series regression (BSR), a trial-level analysis of functional connectivity between ROIs.

To foreshadow our results, the behavioural data showed the predicted U-shape function of memory accuracy against expectancy in both paradigms. However, we found no clear evidence for SLIMM's predictions of an interaction between expectancy and SME in neither the MPFC nor hippocampus. There was also no evidence that functional connectivity between MPFC and hippocampus was modulated by expectancy. We discuss potential alternative explanations.

2. Methods

(a) Participants

The study recruited 46 participants aged 18–35 (mean age = 22.63 ± 4.10; 17 males and 29 females) who did both paradigms in a single session. All participants were right-handed and native English speakers. Informed consent was obtained from all participants and they were paid £40 for participation, in accordance with local University of Cambridge ethics approval CPREC 2020.018.

Our power analyses showed that 34 participants would provide 80% power to detect a within-participant medium effect size of Cohen's $d = 0.5$ when using a one-tailed test at $\alpha = 0.05$.

(b) Stimuli

Each paradigm is summarized in figure 1, along with some example stimuli.

(i) Paradigm 1

This paradigm is a 2D adaptation of a three-dimensional (3D) virtual reality task [26]. We created 90 unique scene–object pairings using 'The Sims' video game engine. The 90 semantically distinct objects were placed in 90 unique scenes (e.g. kitchen/living room/bathroom), in locations that varied in their expectancy. For example, an alarm clock on top of a pillow within a bedroom scene was typically rated as unexpected. For the memory test, we created two alternative lures for each object–scene pairing and tried to match them on expectancy (e.g. alarm clock on a lamp and alarm clock on a high shelf), based on ratings from an independent group of participants. Note, however, that the foils may not be perfectly matched from the idiosyncratic perspective of any single participant in the present study (which would require the same participants to also rate the congruency of every foil too, which would have taken a prohibitive amount of time). We created the stimuli in order to have equal numbers of objects in three levels of expectancy: congruent, neutral or incongruent. Nonetheless, given likely individual differences in expectancy, our analyses used each participant's own congruency ratings (from a final debriefing phase) as continuous predictors of memory.

(ii) Paradigm 2

We used 180 sentences (120 idioms and 60 non-idioms) such that we could manipulate the expectancy of the final word. That final word was shown as a picture, which could be either expected, unexpected or neutral, based on the participant's own rating during the encoding phase. The sentences contained 8–15 words (or 51–62 characters), and were paired with images of famous people as the supposed speakers of the sentence. There were three different images corresponding to each final word, with one image appearing in each condition, so as to match the familiarity of the images used in the later AFC test. Speaker–sentence pairs were fixed for all participants, but their assignment to the congruent/incongruent condition and their sentence final images were counterbalanced across participants (six counterbalances).

(c) Procedure

In the scanner, the encoding phase of paradigm 1 was run first, followed by the encoding phase of paradigm 2. After removal from the scanner, the retrieval phase of paradigm 2 was run first, followed by the retrieval phase of paradigm 1. While this meant that the retention interval for paradigm 1 (approx. 150 min) was longer than for paradigm 2 (approx. 15 min), this retrieval order was chosen because separate piloting of each paradigm had shown that paradigm 2 was slightly harder than paradigm 1, so increasing the retention interval for paradigm 1 was expected to more closely match overall performance for the two paradigms (though see §3).

(i) Paradigm 1

The study phase consisted of 90 trials divided into two runs. In each trial, participants viewed a target object (e.g. clock) presented on a pedestal in an empty room for 2 s. The object label was presented below it. This was followed by a presentation of a scene that did not contain the target object for 2 s. The target object then appeared in a pre-defined location within the scene. The object stayed on the screen for 2 s and the participant's task was to press key 1 as soon as they had located the target object. To ensure engagement with the task, there were four catch trials during which a red dot was overlaid on top of the object. For these trials, participants were instructed to press key 2 instead. Finally, a green circle surrounded the object to ensure participants saw the location of the object. This was followed by a fixation cross for 1.5 s, which then turned red for 0.5 s in order to pre-empt the start of a new trial.

During the test phase outside the scanner, participants completed a 3-AFC associative memory test for all 90 object–scene pairs. Each trial presented three versions of the same object–scene pairing, differing only in the object's location within the scene. To prevent reliance on gist memory, lure locations approximately matched the original trial's congruency (i.e. incongruent targets had incongruent lures). Furthermore, the target location (e.g. clock on pillow) was shown from a different camera angle, to minimize familiarity for a specific image, i.e. particular viewing angle. Participants had 6 s to respond, and then rated their response as 'Remember', 'Familiar' or 'Guess' (based on instructions from [56]).

In a final debriefing, participants viewed the original object–scene pairs again, and rated expectancy on a visual analogue scale (−100 = unexpected, to 100 = expected). These values were used to define expectancy for each trial and participant.

(ii) Paradigm 2

The study phase consisted of 180 trials divided into four runs. On each trial, participants saw first a fixation cross (1 s) and then the image of a famous person in the centre of the screen (1 s), followed by the speaker and sentence frame together, with the final word blanked out (6 s). Participants then saw an image representing the sentence final word and had to decide whether the ending was 'expected', 'unexpected' or 'neither' (4 s).

Outside the scanner, participants performed two 3-AFC memory tasks with 180 trials each. On each trial of the first 'associative memory' task, participants saw a previous sentence frame and had to choose the correct sentence final image from three previously seen images of the same word (i.e. matched for familiarity), and indicate high or low confidence in their response. In the second 'source memory' task, participants saw the previous sentences again, and had to pick the correct speaker from three previously seen speakers.

Note that expectancy ratings were acquired after the encoding and retrieval phases in paradigm 1, but during encoding in paradigm 2. A problem with acquiring ratings after retrieval (as in paradigm 1) is that they might be influenced by memory for the encoding trial. However, a problem with acquiring ratings during encoding (as in paradigm 2) is that making the expectancy dimension explicit might modulate its effects on memory (e.g. if the rating detracts from normal encoding, or affects the amount of semantic elaboration). This is why we deemed it important to find the U-shape in both paradigms, suggesting that it is not contingent on the manner in which expectancy was defined.

(d) Magnetic resonance image acquisition

Only the encoding phases were performed in the scanner, with paradigm 1 preceding paradigm 2. All data were acquired on a 3 T Siemens Prisma scanner with a 32-channel head coil. A T1-weighted high-resolution structural image was acquired with a 3D magnetization-prepared rapid gradient echo (MPRAGE) sequence and parameters repetition time (TR) = 2.25 s, echo time (TE) = 3.02 ms, 9° flip angle, field of view (FOV) = 192 × 256 mm, and 1 mm isotropic voxels. This was followed by two runs for the study phase of paradigm 1, each lasting approximately 8 min, and four runs for the study phase of paradigm 2, each lasting approximately 9 min. Blood oxygen level-dependent (BOLD)-weighted functional images were acquired with a gradient-echo echo planar imaging (EPI) sequence with 58 near-axial slices with in-plane resolution of 3 × 3 mm and thickness of 2 mm with no slice gap and interleaved acquisition (TR = 1.87 s, TE = 25 ms, flip angle = 75°, FOV = 192 × 192 mm, multiband acceleration factor of 2). Fieldmap data were acquired using a dual-echo gradient echo sequence (TR = 561 ms, TE1 = 4.86 ms, TE2 = 7.32 ms, 3 × 3 × 2 mm voxels), collected in-between the two paradigms.

(e) Image pre-processing

All imaging data were pre-processed using standard pre-processing pipelines as implemented in fMRIPrep 21.0.1 [57], which is based on Nipype 1.6.1 [58]. In short, T1- and BOLD-weighted images were skull-stripped and the T1-weighted image was registered to the MNI152 template. The BOLD-weighted images were slice-time corrected, motion corrected, susceptibility distortion corrected, co-registered with the high-resolution structural and registered to the MNI152 template. BOLD-weighted images were not smoothed for the main ROI analyses but a 6 mm full width at half maximum (FWHM) Gaussian smoothing kernel was applied for whole-brain exploratory analyses.

(f) Regions of interest definition

Given the predictions of SLIMM, we focused our analyses on two ROIs: hippocampus and MPFC (see figure 4). Both ROIs were defined *a priori*, and used for both paradigms. The hippocampal ROI was defined based on a high-field sub-cortex atlas

that used resting state functional connectivity gradients to delineate sub-cortical nuclei boundaries [59]. For the MPFC ROI, we created an 8 mm sphere around the peak coordinates of the van Kesteren *et al*. [30] study, which found interaction between congruency and subsequent memory effects.

(g) Behavioural analysis

For behavioural data, we fitted a logistic mixed effect model predicting memory accuracy on each trial from a second-order polynomial expansion of the subjective expectancy ratings (continuous in paradigm 1, and three levels in paradigm 2), along with random intercepts and slopes for each participant, i.e.:

$$\text{Memory}_{pt} \sim f\{\text{poly}(\text{Expectancy}_{pt}, 2) + (\text{poly}(\text{Expectancy}_{pt}, 2) \mid \text{Participant}_{p})\},$$

where p indexes participant, t indexes trial and f is a logistic linking function.

When breaking down memory by Remembered versus Familiar responses in paradigm 1, we estimated recollection by replacing 'Memory' above with 'Remembered', and estimated familiarity by replacing 'Memory' with 'Familiar' and removing Remembered trials (as in [26], i.e. assumed recollection and familiarity are independent processes).

We also examined whether the order of item presentation in paradigm 1 was associated with expectancy ratings and memory, in case expectancy ratings decreased over time as the experiment progressed. We found no evidence that the order of presentation was associated with the expectancy ratings (see electronic supplementary material).

(h) Functional magnetic resonance imaging time-series models

We used a general linear model (GLM), in which neural components for each trial-type were convolved with a canonical haemodynamic response function (HRF) using SPM (paradigm 1) and Nilearn (paradigm 2). Details of the neural components of each trial are described for each paradigm in the following sections. Additional confound regressors included a high-pass filter with a cut-off of 1/128 Hz, the six motion parameters, framewise displacement (FD) and the average signal in white matter (WM) and cerebrospinal fluid (CSF) masks as computed by fMRIPrep.

(i) Paradigm 1

To model the neural effects of interest, target object onsets were modelled by delta functions, separately for remembered and forgotten trials. The height of these delta functions was modulated by each participant's Z-scored expectancy rating. These were then convolved with a canonical HRF. Additional regressors of no interest were created from convolving delta functions at the onset of the empty scene at the start of each trial (regardless of object memory) and the onset of the red dot for rare catch trials.

(ii) Paradigm 2

To model the neural effects of interest, a delta function was used at the onset of the picture that ended the sentence, split according to whether it was remembered or forgotten, and whether it was expected, unexpected, or neither. All regressors were convolved with a canonical double gamma HRF function matching the default option used in SPM. Additional regressors of no interest were created for the speaker onset, modelled with delta function, and the sentences, modelled with 6 s boxcar regressors, separately for idiom and non-idiom sentences.

(i) Whole-brain and regions of interest (least squares unitary trial-averaged encoding models)

For the whole-brain analysis, we used a standard 'least squares unitary' (LSU) model, estimating mean responses for each trial-type, weighted across runs by trial count. Planned contrasts tested: (i) main effect of Remembered versus Forgotten (i.e. SME); (ii) main effect of expectancy (the parametric modulation in paradigm 1, and the Expected versus Unexpected contrast, averaged over memory, in paradigm 2), and (iii) interaction between SME and expectancy. The resulting contrast value for each participant was entered into a one-sample t-test that treated the participant as the only random effect. We applied threshold-free cluster enhancement (TFCE) using 5000 permutations through the 'randomize' function [60] to correct for multiple comparisons (see figure 3).

For the ROI analyses, the same factorial contrasts were estimated, and Bayes factors computed with the BayesFactor package, using a default Cauchy prior with a scale of 0.707 [61]. These produced the plots shown in figure 4.

(j) Supplementary regions of interest analyses (least-squares separate trial-wise encoding and decoding models)

The standard LSU approach averages over trial-level variability and predicts brain activity from psychological factors, as is typical in 'encoding' models. For the beta-series regression (BSR), we estimated trial-wise activity using a 'least-squares separate' (LSS) approach, where a separate GLM is fitted for each trial [62,63]. The resulting trial-wise estimates allowed us not only to examine functional connectivity, but also to run mixed effect encoding models that include both trial-level and participant-level variance (see Supplementary figure S3). All predictors in all models were Z-scored.

However, the SLIMM's main prediction is that the U-shape function of memory is explained by two brain systems, which is essentially a decoding model, with the outcome variable being behaviour rather than brain activity. Therefore, in the electronic supplementary material, we also report a trial-wise, logistic mixed-effect model, where we predicted memory performance as a function of both MTL and MPFC activity and their interactions with expectancy (Memory ~ MTL ∗ Expectancy + MPFC ∗ Expectancy + (Expectancy | Participant)). While none of these models furnished support for the SLIMM predictions, the electronic supplementary material results demonstrate different effects depending on whether one uses encoding or decoding models (owing to covariance between multiple independent variables and different error terms; see Discussion (§4) for further consideration of this issue).

(k) Beta-series regression

Finally, to test SLIMM's prediction that MPFC inhibits MTL as expectancy increases, we fitted the BSR model:

$$\mathrm{MTL}_{pt} \sim \mathrm{MPFC}_{pt} * \mathrm{Expectancy}_{pt} + (\mathrm{Expectancy}_{pt} \mid \mathrm{Participant}_{p}).$$

The main effect of MPFC on MTL is likely to be positive (e.g. owing to variations in attention affecting activity in both ROIs), but SLIMM predicts a less positive relationship when expectancy is high, because MPFC suppresses MTL (i.e. a negative interaction term). Note that the main effect of expectancy is not of interest here, since it is captured by the encoding model above.

3. Results

(a) Behavioural results

We excluded 11 participants in paradigm 1 for whom permutation testing showed that their AFC memory performance was not above chance (leaving $n = 35$). This relatively large proportion of exclusion is likely because the scanner-version of this paradigm entailed a longer retention interval than its piloted version (owing to the interjection of encoding and test phases of paradigm 2; see Procedure §2c), resulting in worse performance than expected. While this may affect generalization of the results of paradigm 1 to a more typical population, this exclusion is important for trial-based subsequent memory analyses that remembered responses include more than just guesses.

For paradigm 2, six participants were excluded: three with technical issues, two who did not complete the task, and one for whom permutation testing showed chance-level AFC performance (leaving $n = 40$).

(i) Paradigm 1

Responses on the 3-AFC task that participants reported as guesses were treated as forgotten trials. For the primary outcome of 3-AFC associative memory, participants showed average memory performance of $M = 55.9\%$ (s.d. = 0.07%), which was significantly above the chance level of 33%, $t_{34} = 18.84$, $p < 0.001$ (though this is biased by removal of participants who performed at chance level; see Methods (§2)). The logistic mixed effects models revealed the predicted U-shaped function of memory against expectancy (figure 2, left), with a significant quadratic effect of expectancy ($\beta = 0.12$; 95% CI [0.04, 0.23]; $z = 3.04$; $p = 0.002$), but no significant linear component ($\beta = 0.11$; 95% CI [−0.005, −0.23]; $z = 1.87$; $p = 0.06$). In the electronic supplementary material, we show that participants had quicker reaction times during encoding for expected trials, but this effect did not differ between remembered and forgotten trials, and the U-shape function of memory remained significant even when including Reaction Time (RT) as a confounding variable.

We additionally tested the secondary prediction of the SLIMM model that recollection processes would be particularly engaged for unexpected items. For the subjective memory ratings provided after each trial, 27% of responses were Remember, 41% were Familiar and 32% were Guess. Assuming recollection and familiarity are independent (see Behavioural analysis), we found a U-shaped effect of expectancy for recollection, but no effect of expectancy on familiarity (electronic supplementary material, figure S1). This does not correspond to SLIMM's predictions, but is consistent with our previous study [26].

(ii) Paradigm 2

Mean accuracy on the 3-AFC memory task, after recoding guesses as forgotten, for sentence-ending images was 53.0% (s.d. = 12.6%), significantly above chance, $t_{39} = 26.5$, $p < 0.001$. As predicted, memory showed a significant quadratic effect of expectancy ($\beta = 0.25$, 95% CI [0.19, 0.32], $z = 7.48$, $p < 0.001$). The linear component was also significant ($\beta = 0.42$, 95% CI [0.32, 0.51], $z = 8.49$, $p < 0.001$), with generally better memory for Expected trials (see figure 2). Pairwise comparisons confirmed that Expected trials were remembered better than Neither trials ($\beta = 1.09$, $z = 9.83$, $p < 0.001$), although evidence that Unexpected trials were also remembered better than Neither trials did not reach significance ($\beta = 0.13$, $z = 1.54$, $p = 0.12$). As in paradigm 1, including reaction time at encoding did not affect the memory results (see electronic supplementary material).

Mean accuracy on the 3-AFC source ('speaker') memory did not show any effects of expectancy (see electronic supplementary material, figure S2).

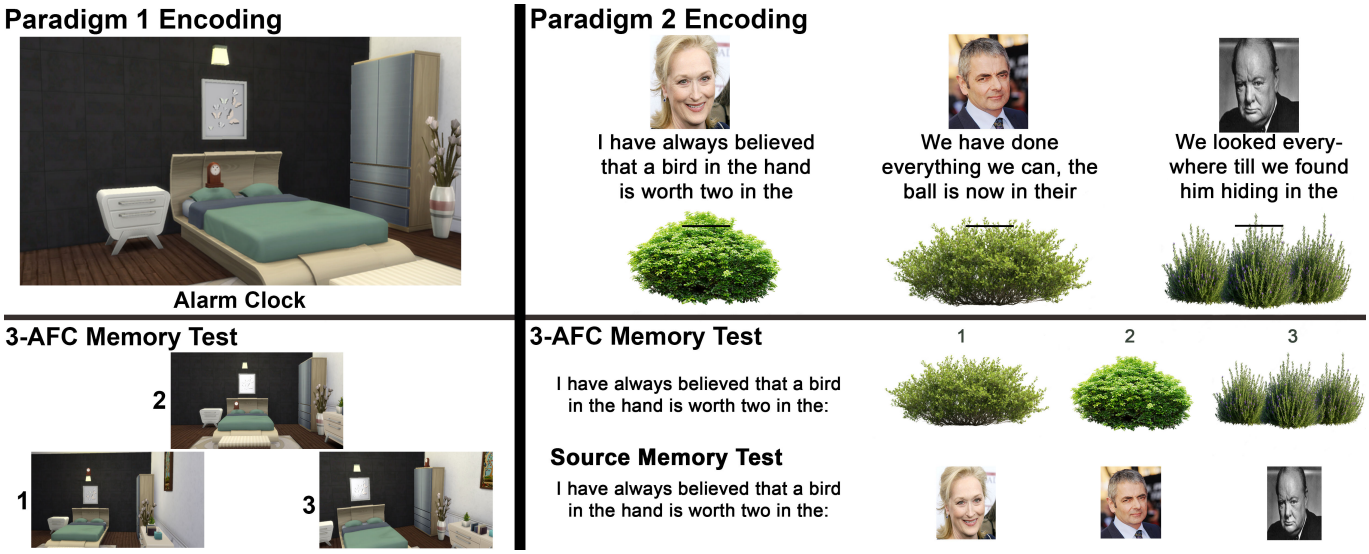

Figure 1. Design schema for both paradigms. In paradigm 1 (left), participants learned objects in scene pairs with varied levels of expectancy (unexpected in this example). Memory was tested with a 3-AFC recognition memory task that had expectancy-matched lures. To reduce the influence of familiarity with the target image itself, the target scene–object pair was presented from a different viewing angle. In paradigm 2 (right), participants first saw a picture of a famous person, acting as a speaker, followed by an incomplete sentence, which was completed after 6 s by a nameable image depicting a known concept (e.g. bush), which acted as the final word of the sentence. Participants were instructed to predict the ending of the sentence. Sentences were taken from idioms or non-idioms. Participants rated whether the image completing the sentence (e.g. bush) was 'expected', 'unexpected' or 'neither'. For instance, 'bush' might be expected ending for the first idiom ('two in the bush'), unexpected as the ending of the wrong idiom 'ball is in their bush', or neither/neutral for the last sentence 'hiding in the bush'. The bush images all acted as lures for the 3-AFC tests, in order to match for familiarity. Assignment of sentence-ending images to different conditions was counterbalanced between participants. Outside the scanner, participants performed two surprise 3-AFC tasks with lures matched for familiarity. Participants provided continuous expectancy ratings for the object–scene pairs in paradigm 1 after all memory tests were completed.

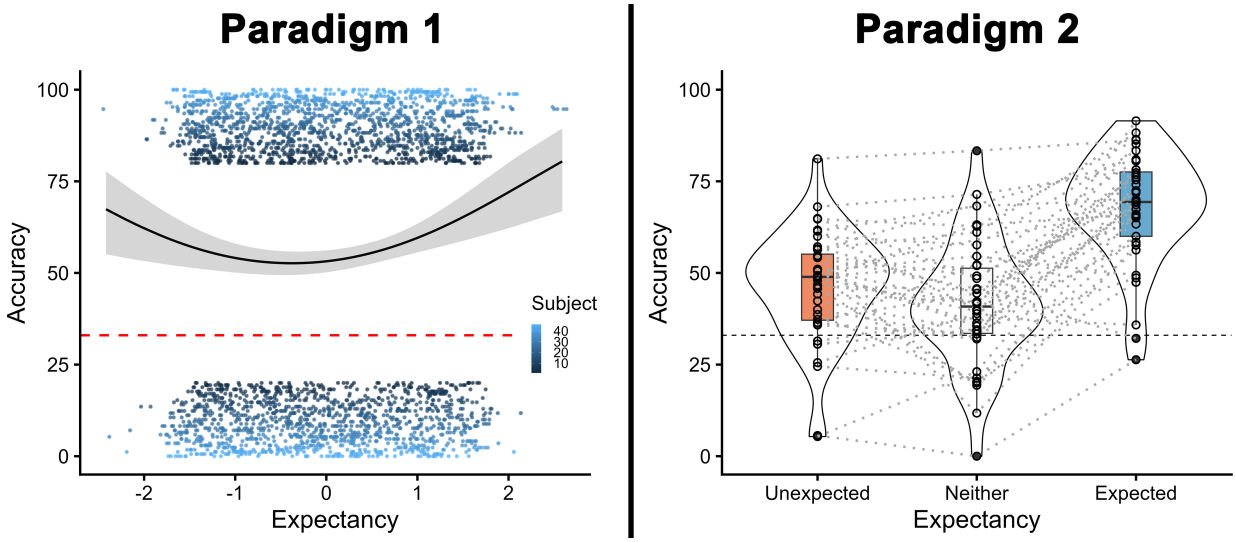

Figure 2. Behavioural memory performance. Left subplot shows raw accuracy (dots, jittered vertically by subject) and predicted (mean plus 95% confidence intervals) memory performance for paradigm 1 using second-order polynomial expansion of expectancy. Red dashed line shows chance performance. Right subplot shows condition-averaged memory performance for paradigm 2. In both paradigms, memory was better for highly unexpected and expected items compared with items with mid-level expectancy.

(b) Whole-brain analyses

Figure 3 shows the supra-threshold clusters from the whole-brain analysis for the interaction between memory and expectancy—the main prediction of SLIMM—as well as the main effects of memory and expectancy.

No cluster showed a significant interaction between memory and expectancy in paradigm 1. In paradigm 2, there were bilateral clusters in the temporal sulcus and supramarginal cortex (plus more ventral temporal cortex and superior parietal cortex in the right hemisphere; top right panel of figure 3), which showed greater SMEs for expected than unexpected trials. A cluster in the anterior right hippocampus (10 voxels, with peak MNI coordinates of (32, −18, −16)) also showed the same interaction, but this was in the opposite direction to that predicted by SLIMM.

Despite the lack of strong evidence for the interactions predicted by SLIMM, the main effects of subsequent memory showed several large clusters that were similar across paradigms (middle row of figure 3), with encoding-related activations in bilateral

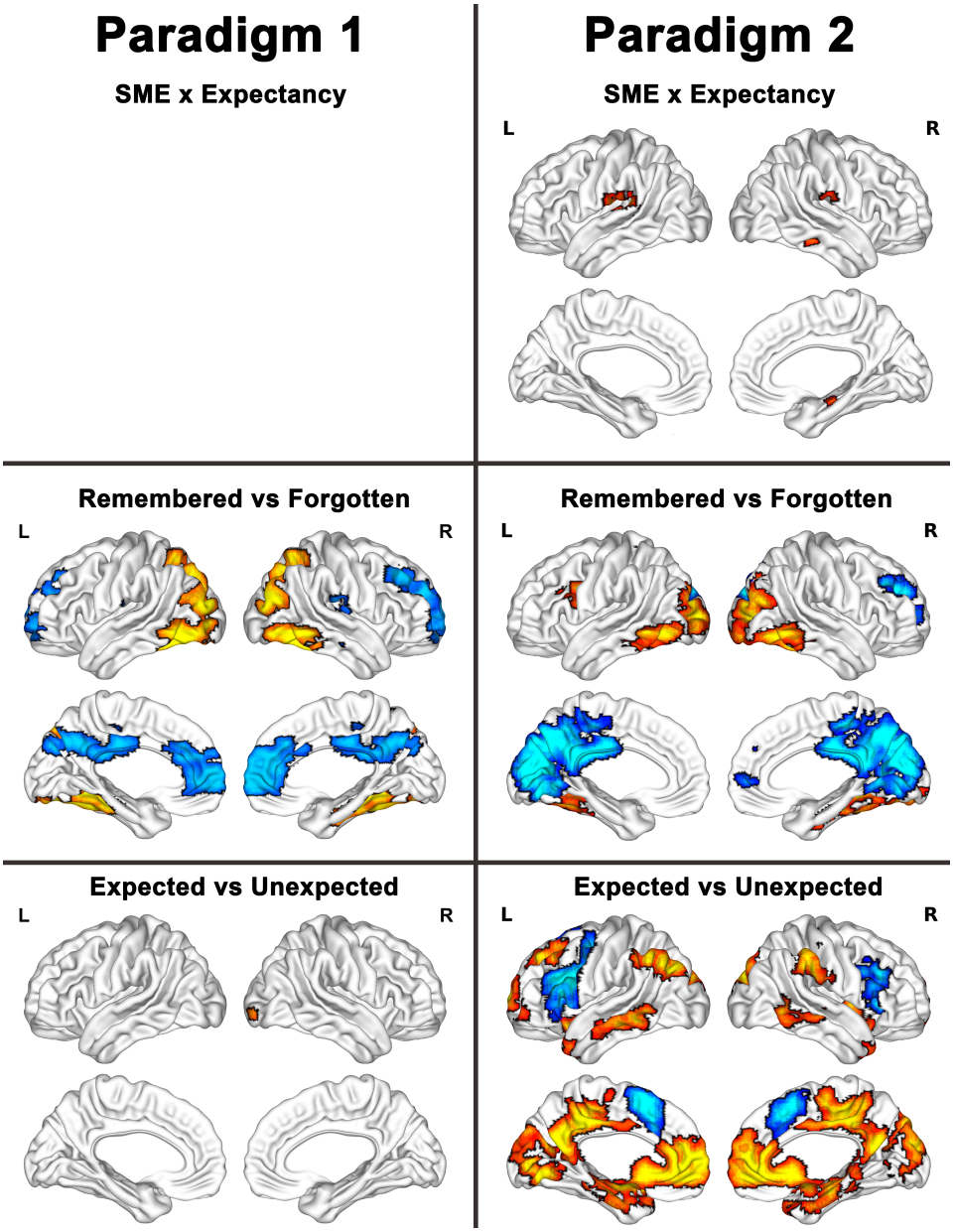

Figure 3. Whole-brain clusters surviving threshold-free cluster enhancement correction for paradigm 1 (left) and paradigm 2 (right), for each of the three factorial effects of interest (rows). Top left is blank as no regions survived correction. Top row corresponds to the critical schema-linked interactions between medial prefrontal and medial temporal regions (SLIMM) hypothesis of memory-by-expectancy interaction, where warm colours in paradigm 2 indicate a greater subsequent memory effect (SME) for expected trials; no clusters survived correction in paradigm 1. The second row shows clusters with greater activation (warm colours) for subsequently remembered trials (averaged across expectancy), and greater activation (cool colours) for subsequently forgotten trials. The final row shows clusters whose activation increased (warm colours) or decreased (cool colours) with expectancy (averaged across memory).

ventral (e.g. fusiform) and dorsal (e.g. posterior lateral parietal) visual streams (plus left prefrontal cortex in paradigm 2), and encoding-related deactivations in medial frontal and medial parameter clusters. These patterns are similar to those expected from the literature (e.g. [27,28]), suggesting that both paradigms and their data were sufficient to identify memory-related activity.

Finally, paradigm 1 did not show much for the main effect of expectancy (apart from a small cluster in early visual cortex showing greater activation for expected trials), while paradigm 2 showed stronger effects of expectancy, the most notable of which were activations in lateral parietal and lateral temporal regions, and deactivations in bilateral posterior prefrontal regions.

(c) Activation differences within regions of interests (trial-averaged encoding models)

Here we focused on two *a priori*, independently defined ROIs key to the SLIMM framework. figure 4 shows results from contrasts from the trial-averaged encoding model, in which the mean activation across trials of each type (and across voxels within each of the two ROIs) is shown for each of the three factorial effects of interest. Results for the analogous trial-wise encoding model (based on single-trial estimates) are shown in the electronic supplementary material.

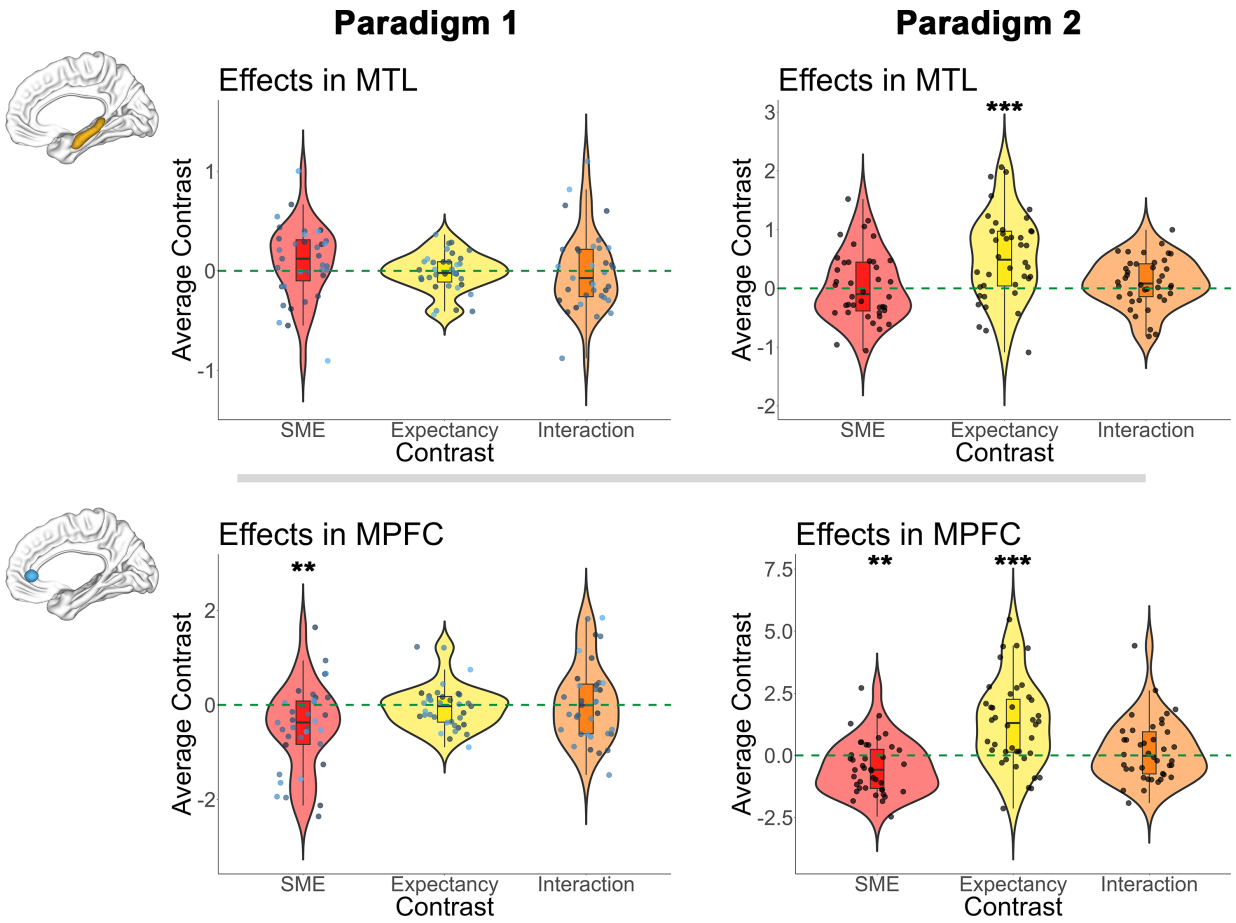

Figure 4. Planned comparisons of mean activation within hippocampal and medial prefrontal cortex (MPFC) regions of interest (rows) and paradigm (columns). Each point and line represents the contrast estimate for one participant for: (i) the main effect of expectancy, (ii) the main (linear) effect of memory (SME), and (iii) the interaction between these two factors, respectively. Note that for paradigm 1, the expectancy factor represents the linear slope across 200 levels, whereas in paradigm 2 it reflects the difference between expected and unexpected endings to idioms; both based on participant's individual ratings. MTL, medial temporal lobe. **: $p < 0.01$; ***: $p < 0.001$

(i) Hippocampal regions of interest

In paradigm 1, none of the three planned comparisons reached significance: there was no evidence of linear modulation of hippocampal activity by expectancy ($t_{34} = -0.53$, 95% CI [−0.08, 0.05], $p = 0.59$, $BF_{10} = 0.21$), no evidence for an SME ($t_{34} = 1.43$, 95% CI [−0.04, 0.22], $p = 0.16$, $BF_{10} = 0.46$), and critically for SLIMM, no evidence that the SME was bigger for surprising trials, i.e. no evidence for an interaction between expectancy and memory ($t_{34} = -0.09$, 95% CI [−0.14, 0.13], $p = 0.92$, $BF_{10} = 0.18$). In the latter case, the BF for the null (BF = 1/0.18 = 5.55) approached strong evidence for no interaction.

In paradigm 2, there was a significant main effect of expectancy, but with higher activation for expected over unexpected trials, contrary to SLIMM ($t_{39} = 4.53$, 95% CI [0.29, 0.76], $p < 0.001$, $BF_{10} = 430.25$). More importantly, there was no significant SME ($t_{39} = 0.37$, 95% CI [−0.15, 0.22], $p = 0.71$, $BF_{10} = 0.18$), nor significant interaction ($t_{39} = 1.60$, 95% CI [−0.03, 0.24], $p = 0.12$, $BF_{10} = 0.55$). If anything, there was a trend for the opposite interaction to that predicted by SLIMM, with larger SME for expected than unexpected trials (resembling the right hippocampal cluster in the whole-brain analyses).

(ii) Medial prefrontal cortex regions of interest

In paradigm 1, we again did not observe the critical interaction between memory and expectancy ($t_{34} = 0.30$, 95% CI [−0.24, 0.33], $p = 0.77$, $BF_{10} = 0.19$). However, we did observe a main effect of memory (Memory: $t_{(34)} = -3.12$, 95% CI [−0.78, -0.17], $p = 0.003$, $BF_{10} = 10.12$), with reduced activation for Remembered than Forgotten trials, but no linear effect of expectancy (Expectancy: $t_{34} = -0.57$, 95% CI [−0.21, 0.12], $p = 0.57$, $BF_{10} = 0.21$).

In paradigm 2, we again found a significant main effect of Memory ($t_{39} = -2.84$, 95% CI [−0.82, −0.13], $p = 0.007$, $BF_{10} = 5.49$), with reduced activation for remembered trials, like paradigm 1. Unlike paradigm 1, there was also a significant main effect of Expectancy ($t_{39} = 4.59$, 95% CI [0.71, 1.82], $p < 0.001$, $BF_{10} = 506.35$), with greater activation for expected trials. Again, however, we did not find a significant interaction between Memory and Expectancy ($t_{39} = 0.93$, 95% CI [−0.21, 0.58], $p = 0.36$, $BF_{10} = 0.25$).

(iii) Memory predicted by both regions of interests (decoding models)

We also fitted a decoding model in which we combined hippocampal and MPFC activations to predict trial-wise memory performance (using a logistic mixed effects model; see Methods (§2) and electronic supplementary material). This is a more relevant test of SLIMM's prediction that the U-shaped function of memory against expectancy derives from two memory systems: an MTL one that encodes unexpected events and an MPFC one that encodes expected events. As in the previous analysis (see figure 4), there was no interaction effect. However, now both MPFC and MTL showed significant SME effects, unlike the above individual-ROI encoding models, highlighting the need for careful consideration when constructing statistical models of brain–behaviour relationships. For full statistics, see the electronic supplementary material.

(iv) Connectivity between regions of interests

Finally, SLIMM also predicts that functional connectivity between MPFC and hippocampus should differ as a function of expectancy: specifically, that the more expected a trial, the more MPFC should inhibit the hippocampus. We tested this using BSR, in which trial-wise activation in the hippocampal ROI was predicted by the trial-wise activation in MPFC, by expectancy, and by the interaction between these two.

In paradigm 1, there was a significant main effect of MPFC (β = 0.36, 95% CI [0.29, 0.43], $t_{29.11}$ = 10.54, $p < 0.001$), which reflected a positive correlation between ROIs, which is expected since there are many factors (such as attention) that vary over trials and affect many brain regions similarly. Importantly, there was no evidence that this connectivity decreased during expected trials, i.e. no evidence for an interaction (β = −0.002, 95% CI [−0.03, 0.03], $t_{1468.18}$ = −0.15, p = 0.88). The same pattern was seen in paradigm 2, with a significant positive main effect of MPFC (β = 0.45, 95% CI [0.38, 0.52], $t_{(39.7)}$ = 13.72, $p < 0.001$) but no evidence that this interacted with expectancy (β = −0.008, 95% CI [0.38, 0.52], $t_{(37.8)}$ = −0.57, p = 0.57).

4. Discussion

To investigate the predictions of the SLIMM framework at encoding, we used two fMRI paradigms with the same participants, testing encoding-related activation in the hippocampus and MPFC as a function of expectancy. In one paradigm, expectancy was manipulated spatially by presenting objects in either expected or unexpected locations within a room, while in the other, expectancy was manipulated through the predictability of final words in idiomatic sentences. Both paradigms used AFC recognition memory paradigms with lures matched as well as possible for congruency and familiarity, which is important to control for retrieval strategies that might bias effects of expectancy [64], and that may have obscured the U-shape function of expectancy/congruency in some other studies.

Consistent with the SLIMM framework and previous behavioural research, both paradigms revealed better memory for highly expected and highly unexpected associations compared with moderately expected ones, confirming a U-shaped relationship between memory and expectancy. However, we failed to find evidence that this behavioural pattern was driven by differential engagement of MPFC and hippocampus, i.e. SLIMM's predictions that MPFC supports encoding of expected associations and the MTL (hippocampus) supports encoding of unexpected associations. Neither ROI showed an interaction between expectancy and subsequent memory, regardless of analysis type, i.e. whether using encoding models (predicting fMRI activation from memory) or decoding models (predicting memory from fMRI activation), and regardless of analysis level, i.e. including both trial-averaged and trial-wise (mixed effects) models.

Whole-brain analyses replicated some well established subsequent memory effects (SMEs) in regions including MPFC, but interestingly not the hippocampus. This absence of hippocampal SME is surprising given that it is consistently observed in other studies (e.g. [27,65–67]). Considering our successful replication of SMEs in other regions—e.g. positive SMEs in ventral temporal regions and negative SMEs in posterior medial and MPFC regions [27,28,66–68]—our data and analyses seem valid. One possibility is that our anatomically defined hippocampal ROI was too broad, masking sub-regional effects. Nonetheless, in a combined decoding model in which memory activity was predicted by both hippocampal and MPFC ROIs, we did observe a significant SME of both ROIs: a positive effect of hippocampal activity on memory accuracy and a negative effect of MPFC activity, consistent with the prior literature. In other words, once other sources of variance were accounted for, the hippocampus did show a positive SME (see Methodological considerations (§4a) for further discussion of different statistical models).

The results for the main effect of expectancy differed between paradigms. In paradigm 1, neither ROI nor whole-brain analysis revealed clear effects of expectancy, possibly owing to any surprise diminishing over the course of the experiment as participants adapted to seeing objects in odd locations. However, we found no evidence for order effects in paradigm 1, suggesting expectancy ratings did not decrease as the experiment progressed (see electronic supplementary material, figure S4). Moreover, the expectancy range was sufficient to modulate behavioural memory performance (i.e. produce the U-shaped function). In contrast, paradigm 2 elicited strong expectancy effects in whole-brain and ROI analyses, likely owing to more salient violations (unexpected completions of idioms), a higher number of trials and/or a longer trial duration that allowed participants to form stronger predictions. In this paradigm, both MPFC and hippocampus showed greater activation for expected trials. While the MPFC findings are consistent with SLIMM, the hippocampal activation is opposite to what is predicted by its proposed role in detecting unexpected events.

Both paradigms used a relatively short delay between encoding and retrieval phases (maximum 2.5 h). Prior work has shown expectancy effects on memory after a delay of a day or more [21,29,69–73]. Some of this work found that memory for incongruent items was better than memory for congruent items after a delay [21]; however, other work found the opposite

effect, suggesting that longer delays increase the reliance on existing schemas [69,71,72,74–77]. Further examination of the interaction between delay and expectancy is needed, and would be theoretically informative, e.g. in terms of effects of expectancy on consolidation. Nonetheless, we note that, despite the relatively short delay used here, we were able to see the predicted behavioural U-shape function. This suggests our delay was sufficient to test the neurobiological predictions of the SLIMM framework.

One possible explanation for our lack of hippocampal activation to unexpected trials comes from a recent study by Aitken & Kok [78], who investigated how the hippocampus acquires predictive associations between visual shapes. In the early stages of learning, hippocampal activity predominantly reflected unexpected stimuli, consistent with the representation of prediction errors. However, when exposure to the predictive cues was extended, hippocampal representations gradually shifted to reflect the predicted shapes, regardless of whether they were actually presented, suggesting a transition from error-driven encoding to prediction-based retrieval. These findings suggest that the hippocampus flexibly balances encoding and retrieval depending on the stage of learning, and a similar shift may have occurred in our study, averaging out over time.

A related possibility is that the hippocampus responds differently depending on the nature of the violated expectations. Varga *et al*. [79] recently showed that hippocampal responses to prediction errors vary depending on whether the expectation was derived from episodic memory (i.e. prior experiences) or schematic knowledge. According to this view, hippocampal activation may reflect episodic retrieval triggered by a prediction error, rather than the error signal itself (e.g. [80]). This interpretation is consistent with our pattern of results and may reconcile some mixed findings in the literature, where some studies report stronger hippocampal responses to expected events, while others find the opposite [17,54,81].

Turning to functional connectivity, we were not able to find evidence for SLIMM's further prediction that coupling between MPFC and hippocampus varies as a function of expectancy. BSR revealed no modulation of connectivity by expectancy. This null finding may reflect an incorrect prediction of SLIMM, or limitations of BOLD fMRI in detecting underlying inhibitory mechanisms, or it could be that other unmeasured factors, such as trial-to-trial variability in attention, might have obscured subtle connectivity changes.

Finally, we tested SLIMM's additional behavioural prediction that unexpected events enhance recollection or source memory, but found no support for this. In paradigm 1, recollection was greater for both expected and unexpected trials compared with neutral ones, mirroring the U-shaped pattern. This replicates earlier findings [26] and may reflect the spatial demands of the object-location task, which may enforce recollection. In contrast, paradigm 2 showed no effect of expectancy at all, in terms of source memory for the speaker. One possibility is that participants encoded the speaker as an integral part of the core event, rather than as incidental context. Alternatively, these discrepancies might suggest that SLIMM's predictions are overly simplistic. Recent work (e.g. [82]) indicates that retrieval factors also shape the effects of expectancy on memory, which could explain the dissociation observed here between behavioural effects and encoding-related neural effects. Future studies could investigate neural effects of expectancy during retrieval to test this possibility more directly.

(a) Methodological considerations

An important methodological insight from this work was the realization that to test SLIMM's prediction that the U-shaped memory function arises from two distinct brain systems (one supporting encoding of expected events and another supporting the encoding for unexpected events) requires a decoding model (where behaviour is predicted from brain activity) rather than an encoding model (where brain activity is predicted from behaviour). More specifically, we needed a decoding approach where memory performance is predicted from neural activity across multiple ROIs. Interestingly, the combined-ROI decoding model successfully identified subsequent memory effects in both the MPFC and hippocampus, even though an SME effect was not evident when the hippocampal ROI was analysed independently. This suggests that some subtle neural signals might only emerge when the contributions of multiple brain regions to behaviour are considered.

More generally, a causal inference approach to modelling (e.g. [83]) conceptualizes brain activity to cause behavioural outcomes, rather than *vice versa*. Traditionally, neuroimaging studies, particularly those focused on brain mapping, have used encoding models with brain activity as the dependent variable. However, behaviour is likely driven by coordinated activity and interactions across multiple brain regions, a complexity that is rarely addressed explicitly. Although multi-voxel pattern analysis (MVPA) achieves decoding across many voxels, the interpretations usually remain limited to the level of an individual ROI, rather than broader network dynamics.

5. Conclusion

Overall, while two distinct fMRI paradigms provided behavioural support for SLIMM's predicted U-shaped relationship between expectancy and memory, they did not shed much insight on the underlying neural mechanisms. Contrary to SLIMM's hypothesis, we did not find a differential encoding-related activity of the hippocampus and MPFC for unexpected versus expected information, respectively. More generally, the lack of hippocampal activity in our findings may reflect its sensitivity to task demands, the specific nature of expectancy violations and/or the limitations of standard-resolution imaging. Beyond encoding-related activity, future research should also examine post-encoding and retrieval-related activity, as these processes may play a critical role in shaping the effects of expectancy on memory.

Ethics. Ethics approval for the project was provided by the University of Cambridge Ethics Committee CPREC 2020.018. Informed consent was obtained from all participants and they were paid £40 for participating in the study.

Data accessibility. Data to reproduce the main results in the manuscript and R code for the analyses are hosted at [84]. Raw imaging data can be provided upon request. Supplementary material is available online [85].

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors' contributions. P.P.R.: data curation, formal analysis, investigation, methodology, project administration, software, validation, visualization, writing—original draft, writing—review and editing; K.G.: conceptualization, data curation, formal analysis, investigation, project administration, software; M.M.: conceptualization, data curation, investigation, software; A.J.: conceptualization, data curation, methodology, resources, software; A.G.: conceptualization, investigation, methodology, supervision, validation, writing—review and editing; R.N.H.: conceptualization, funding acquisition, methodology, resources, supervision, validation, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed herein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. Funding for the project was provided by the UK Medical Research Council (SUAG/046/G101400). K.G. is supported by a Basque Government Fellowship (POS_2023_2_0034) and a Marie Skłodowska-Curie Fellowship funded by the UK Research and Innovation Engineering and Physical Sciences Research Council (EP/Y016815/1).

Acknowledgements. We should like to acknowledge Chuyi Zhang and Isabel Turner for their help with the data collection.

References

1. Bird CM. 2020 How do we remember events? *Curr. Opin. Behav. Sci*. **32**, 120–125. (doi:10.1016/j.cobeha.2020.01.020)
2. Clewett D, Davachi L. 2017 The ebb and flow of experience determines the temporal structure of memory. *Curr. Opin. Behav. Sci*. **17**, 186–193. (doi:10.1016/j.cobeha.2017.08.013)
3. Niv Y, Schoenbaum G. 2008 Dialogues on prediction errors. *Trends Cogn. Sci*. **12**, 265–272. (doi:10.1016/j.tics.2008.03.006)
4. Nolden S, Turan G, Güler B, Günseli E. 2024 Prediction error and event segmentation in episodic memory. *Neurosci. Biobehav. Rev*. **157**, 105533. (doi:10.1016/j.neubiorev.2024.105533)
5. Greve A, Cooper E, Kaula A, Anderson MC, Henson R. 2017 Does prediction error drive one-shot declarative learning? *J. Mem. Lang*. **94**, 149–165. (doi:10.1016/j.jml.2016.11.001)
6. Antony JW, Hartshorne TH, Pomeroy K, Gureckis TM, Hasson U, McDougle SD, Norman KA. 2021 Behavioral, physiological, and neural signatures of surprise during naturalistic sports viewing. *Neuron* **109**, 377–390. (doi:10.1016/j.neuron.2020.10.029)
7. Barron HC, Auksztulewicz R, Friston K. 2020 Prediction and memory: a predictive coding account. *Prog. Neurobiol*. **192**, 101821. (doi:10.1016/j.pneurobio.2020.101821)
8. Ben-Yakov A, Smith V, Henson R. 2022 The limited reach of surprise: evidence against effects of surprise on memory for preceding elements of an event. *Psychon. Bull. Rev.* **29**, 1053–1064. (doi:10.3758/s13423-021-01954-5)
9. Chen J, Cook PA, Wagner AD. 2015 Prediction strength modulates responses in human area CA1 to sequence violations. *J. Neurophysiol*. **114**, 1227–1238. (doi:10.1152/jn.00149.2015)
10. Liedtke N, Boeltzig M, Siestrup S, Schubotz RI. 2026 Shaping memory from the start: initial prediction errors during first encoding. *Neuroimage* **325**, 121660. (doi:10.1016/j.neuroimage.2025.121660)
11. Lisman JE, Grace AA. 2005 The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron* **46**, 703–713. (doi:10.1016/j.neuron.2005.05.002)
12. Ranganath C, Rainer G. 2003 Neural mechanisms for detecting and remembering novel events. *Nat. Rev. Neurosci*. **4**, 193–202. (doi:10.1038/nrn1052)
13. Bein O, Trzewik M, Maril A. 2019 The role of prior knowledge in incremental associative learning: an empirical and computational approach. *J. Mem. Lang*. **107**, 1–24. (doi:10.1016/j.jml.2019.03.006)
14. Bellana B, Mansour R, Ladyka-Wojcik N, Grady CL, Moscovitch M. 2021 The influence of prior knowledge on the formation of detailed and durable memories. *J. Memory Lang.* **121**, 104264. (doi:10.1016/j.jml.2021.104264)
15. Bird CM, Davies RA, Ward J, Burgess N. 2011 Effects of pre-experimental knowledge on recognition memory. *Learn. Mem*. **18**, 11–14. (doi:10.1101/lm.1952111)
16. Bransford JD, Johnson MK. 1972 Contextual prerequisites for understanding: some investigations of comprehension and recall. *J. Verbal Learning Verbal Behav*. **11**, 717–726. (doi:10.1016/S0022-5371(72)80006-9)
17. Gilboa A, Marlatte H. 2017 Neurobiology of schemas and schema-mediated memory. *Trends Cogn. Sci*. **21**, 618–631. (doi:10.1016/j.tics.2017.04.013)
18. Klatzky RL, Forrest FH. 1984 Recognizing familiar and unfamiliar faces. *Mem. Cognit*. **12**, 60–70. (doi:10.3758/BF03196998)
19. McAndrews MP, Girard TA, Wilkins LK, McCormick C. 2016 Semantic congruence affects hippocampal response to repetition of visual associations. *Neuropsychologia* **90**, 235–242. (doi:10.1016/j.neuropsychologia.2016.07.026)
20. Ortiz-Tudela J, Milliken B, Botta F, LaPointe M, Lupiañez J. 2017 A cow on the prairie vs. a cow on the street: long-term consequences of semantic conflict on episodic encoding. *Psychol. Res*. **81**, 1264–1275. (doi:10.1007/s00426-016-0805-y)
21. Ortiz-Tudela J, Nolden S, Pupillo F, Ehrlich I, Schommartz I, Turan G, Shing YL. 2023 Not what u expect: effects of prediction errors on item memory. *J. Exp. Psychol. Gen*. **152**, 2160–2176. (doi:10.1037/xge0001367)
22. Raykov PP, Keidel JL, Oakhill J, Bird CM. 2022 The importance of semantic network brain regions in integrating prior knowledge with an ongoing dialogue. *eNeuro* **9**, ENEURO-0116. (doi:10.1523/ENEURO.0116-22.2022)
23. Sommer T. 2016 The emergence of knowledge and how it supports the memory for novel related information. *Cereb. Cortex* **27**, 1906–1921. (doi:10.1093/cercor/bhw031)
24. Wu X, Packard PA, García-Arch J, Bunzeck N, Fuentemilla L. 2023 Contextual incongruency triggers memory reinstatement and the disruption of neural stability. *Neuroimage* **273**, 120114. (doi:10.1016/j.neuroimage.2023.120114)
25. Greve A, Cooper E, Tibon R, Henson RN. 2019 Knowledge is power: prior knowledge aids memory for both congruent and incongruent events, but in different ways. *J. Exp. Psychol. Gen*. **148**, 325–341. (doi:10.1037/xge0000498)
26. Quent JA, Greve A, Henson RN. 2022 Shape of U: the nonmonotonic relationship between object-location memory and expectedness. *Psychol. Sci*. **33**, 2084–2097. (doi:10.1177/09567976221109134)
27. Paller KA, Wagner AD. 2002 Observing the transformation of experience into memory. *Trends Cogn. Sci*. **6**, 93–102. (doi:10.1016/S1364-6613(00)01845-3)
28. Uncapher MR, Wagner AD. 2009 Posterior parietal cortex and episodic encoding: insights from fMRI subsequent memory effects and dual-attention theory. *Neurobiol. Learn. Mem*. **91**, 139–154. (doi:10.1016/j.nlm.2008.10.011)
29. van Kesteren MTR, Ruiter DJ, Fernández G, Henson RN, Fernandez G, Henson RN. 2012 How schema and novelty augment memory formation. *Trends Neurosci*. **35**, 211–219. (doi:10.1016/j.tins.2012.02.001)

30. van Kesteren MTR, Beul SF, Takashima A, Henson RN, Ruiter DJ, Fernández G. 2013 Differential roles for medial prefrontal and medial temporal cortices in schema-dependent encoding: from congruent to incongruent. *Neuropsychologia* **51**, 2352–2359. (doi:10.1016/j.neuropsychologia.2013.05.027)
31. Brod G, Lindenberger U, Wagner AD, Shing YL. 2016 Knowledge acquisition during exam preparation improves memory and modulates memory formation. *J. Neurosci*. **36**, 8103–8111. (doi:10.1523/JNEUROSCI.0045-16.2016)
32. Kumaran D, Maguire EA. 2006 An unexpected sequence of events: mismatch detection in the human hippocampus. *PLoS Biol*. **4**, e424. (doi:10.1371/journal.pbio.0040424)
33. Hasselmo ME, Wyble BP, Wallenstein GV. 1996 Encoding and retrieval of episodic memories: role of cholinergic and GABAergic modulation in the hippocampus. *Hippocampus* **6**, 693–708. (doi:10.1002/(SICI)1098-1063(1996)6:6<693::AID-HIPO12>3.0.CO;2-W)
34. Lörincz A, Buzsáki G. 2000 Two-phase computational model training long-term memories in the entorhinal-hippocampal region. *Ann. N. Y. Acad. Sci*. **911**, 83–111. (doi:10.1111/j.1749-6632.2000.tb06721.x)
35. Vinogradova OS. 2001 Hippocampus as comparator: role of the two input and two output systems of the hippocampus in selection and registration of information. *Hippocampus* **11**, 578–598. (doi:10.1002/hipo.1073)
36. Spens E, Burgess N. 2024 A generative model of memory construction and consolidation. *Nat. Hum. Behav.* **8**, 526–543. (doi:10.1038/s41562-023-01799-z)
37. Tse D, Takeuchi T, Kakeyama M, Kajii Y, Okuno H, Tohyama C, Bito H, Morris RGM. 2011 Schema-dependent gene activation and memory encoding in neocortex. *Science* **333**, 891–895. (doi:10.1126/science.1205274)
38. Ciaramelli E, Ghetti S, Frattarelli M, Làdavas E. 2006 When true memory availability promotes false memory: evidence from confabulating patients. *Neuropsychologia* **44**, 1866–1877. (doi:10.1016/j.neuropsychologia.2006.02.008)
39. Ghosh VE, Moscovitch M, Melo Colella B, Gilboa A, Colella BM, Gilboa A. 2014 Schema representation in patients with ventromedial PFC lesions. *J. Neurosci*. **34**, 12057–12070. (doi:10.1523/JNEUROSCI.0740-14.2014)
40. McCormick C, Ciaramelli E, De Luca F, Maguire EA. 2018 Comparing and contrasting the cognitive effects of hippocampal and ventromedial prefrontal cortex damage: a review of human lesion studies. *Neuroscience* **374**, 295–318. (doi:10.1016/j.neuroscience.2017.07.066)
41. Melo B, Winocur G, Moscovitch M. 1999 False recall and false recognition: an examination of the effects of selective and combined lesions to the medial temporal lobe/diencephalon and frontal lobe structures. *Cogn. Neuropsychol*. **16**, 343–359. (doi:10.1080/026432999380825)
42. Spalding KN, Jones SH, Duff MC, Tranel D, Warren DE. 2015 Investigating the neural correlates of schemas: ventromedial prefrontal cortex is necessary for normal schematic influence on memory. *J. Neurosci*. **35**, 15746–15751. (doi:10.1523/JNEUROSCI.2767-15.2015)
43. Warren DE, Jones SH, Duff MC, Tranel D. 2014 False recall is reduced by damage to the ventromedial prefrontal cortex: implications for understanding the neural correlates of schematic memory. *J. Neurosci*. **34**, 7677–7682. (doi:10.1523/JNEUROSCI.0119-14.2014)
44. Baldassano C, Hasson U, Norman KA. 2018 Representation of real-world event schemas during narrative perception. *J. Neurosci*. **38**, 9689–9699. (doi:10.1523/JNEUROSCI.0251-18.2018)
45. Hsieh LT, Ranganath C. 2015 Cortical and subcortical contributions to sequence retrieval: schematic coding of temporal context in the neocortical recollection network. *Neuroimage* **121**, 78–90. (doi:10.1016/j.neuroimage.2015.07.040)
46. Masís-Obando R, Norman KA, Baldassano C. 2022 Schema representations in distinct brain networks support narrative memory during encoding and retrieval. *eLife* **11**, e70445. (doi:10.7554/eLife.70445)
47. Raykov PP, Keidel JL, Oakhill J, Bird CM. 2021 Activation of person knowledge in medial prefrontal cortex during the encoding of new lifelike events. *Cereb. Cortex* **31**, 3494–3505. (doi:10.1093/cercor/bhab027)
48. Reagh ZM, Ranganath C. 2023 Flexible reuse of cortico-hippocampal representations during encoding and recall of naturalistic events. *Nat. Commun*. **14**, 1279. (doi:10.1038/s41467-023-36805-5)
49. van Kesteren MTR, Fernández G, Norris DG, Hermans EJ. 2010 Persistent schema-dependent hippocampal-neocortical connectivity during memory encoding and postencoding rest in humans. *Proc. Natl Acad. Sci. USA* **107**, 7550–7555. (doi:10.1073/pnas.0914892107)
50. van Kesteren MTR, Rijpkema M, Ruiter DJ, Morris RGM, Fernández G. 2014 Building on prior knowledge: schema-dependent encoding processes relate to academic performance. *J. Cogn. Neurosci*. **26**, 2250–2261. (doi:10.1162/jocn_a_00630)
51. Bein O, Reggev N, Maril A. 2014 Prior knowledge influences on hippocampus and medial prefrontal cortex interactions in subsequent memory. *Neuropsychologia* **64**, 320–330. (doi:10.1016/j.neuropsychologia.2014.09.046)
52. Wagner IC, van Buuren M, Kroes MCW, Gutteling TP, van der Linden M, Morris RG, Fernández G. 2015 Schematic memory components converge within angular gyrus during retrieval. *eLife* **4**, e09668. (doi:10.7554/eLife.09668)
53. Audrain S, McAndrews MP. 2022 Schemas provide a scaffold for neocortical integration of new memories over time. *Nat. Commun*. **13**, 5795. (doi:10.1038/s41467-022-33517-0)
54. Liu ZX, Grady C, Moscovitch M. 2016 Effects of prior-knowledge on brain activation and connectivity during associative memory encoding. *Cereb. Cortex* **27**, 1991–2009. (doi:10.1093/cercor/bhw047)
55. Raykov PP, Keidel JL, Oakhill J, Bird CM. 2020 The brain regions supporting schema-related processing of people's identities. *Cogn. Neuropsychol*. **37**, 8–24. (doi:10.1080/02643294.2019.1685958)
56. Rajaram S. 1993 Remembering and knowing: two means of access to the personal past. *Mem. Cogn*. **21**, 89–102. (doi:10.3758/BF03211168)
57. Esteban O *et al*. 2019 fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nat. Methods* **16**, 111–116. (doi:10.1038/s41592-018-0235-4)
58. Gorgolewski KJ *et al*. 2018 *Nipype*.
59. Tian Y, Margulies DS, Breakspear M, Zalesky A. 2020 Topographic organization of the human subcortex unveiled with functional connectivity gradients. *Nat. Neurosci*. **23**, 1421–1432. (doi:10.1038/s41593-020-00711-6)
60. Winkler AM, Ridgway GR, Webster MA, Smith SM, Nichols TE. 2014 Permutation inference for the general linear model. *Neuroimage* **92**, 381–397. (doi:10.1016/j.neuroimage.2014.01.060)
61. Morey R, Rouder J. 2026 *BayesFactor: computation of Bayes factors for common designs*. See https://github.com/richarddmorey/bayesfactor.
62. Abdulrahman H, Henson RN. 2016 Effect of trial-to-trial variability on optimal event-related fMRI design: implications for beta-series correlation and multi-voxel pattern analysis. *Neuroimage* **125**, 756–766. (doi:10.1016/j.neuroimage.2015.11.009)
63. Mumford JA, Turner BO, Ashby FG, Poldrack RA. 2012 Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *Neuroimage* **59**, 2636–2643. (doi:10.1016/j.neuroimage.2011.08.076)
64. Quent JA, Henson RN, Greve A. 2021 A predictive account of how novelty influences declarative memory. *Neurobiol. Learn. Mem*. **179**, 107382. (doi:10.1016/j.nlm.2021.107382)

65. Ben-Yakov A, Eshel N, Dudai Y. 2013 Hippocampal immediate poststimulus activity in the encoding of consecutive naturalistic episodes. *J. Exp. Psychol. Gen*. **142**, 1255–1263. (doi:10.1037/a0033558)
66. Halpern DJ, Tubridy S, Davachi L, Gureckis TM. 2023 Identifying causal subsequent memory effects. *Proc. Natl Acad. Sci. USA* **120**, e2120288120. (doi:10.1073/pnas.2120288120)
67. Kim H. 2011 Neural activity that predicts subsequent memory and forgetting: a meta-analysis of 74 fMRI studies. *Neuroimage* **54**, 2446–2461. (doi:10.1016/j.neuroimage.2010.09.045)
68. Tang L, Yu Q, Homayouni R, Canada KL, Yin Q, Damoiseaux JS, Ofen N. 2021 Reliability of subsequent memory effects in children and adults: the good, the bad, and the hopeful. *Dev. Cogn. Neurosci*. **52**, 101037. (doi:10.1016/j.dcn.2021.101037)
69. Höltje G, Bader R, Meßmer JA, Zogaj D, Mecklinger A. 2025 Unexpected words that become your best memories: how sentential constraint and word expectedness affect memory retrieval. *Front. Hum. Neurosci*. **19**, 1645907. (doi:10.3389/fnhum.2025.1645907)
70. Sinclair AH, Barense MD. 2018 Surprise and destabilize: prediction error influences episodic memory reconsolidation. *Learn. Mem*. **25**, 369–381. (doi:10.1101/lm.046912.117)
71. Tompary A, Zhou WX, Davachi L. 2020 Schematic memories develop quickly, but are not expressed unless necessary. *Scient. Rep*. **10**, 16968. (doi:10.1038/s41598-020-73952-x)
72. van Kesteren MTR, Rijpkema M, Ruiter DJ, Fernández G. 2013 Consolidation differentially modulates schema effects on memory for items and associations. *PLoS One* **8**, e56155. (doi:10.1371/journal.pone.0056155)
73. Yacoby A, Reggev N, Maril A. 2023 Lack of source memory as a potential marker of early assimilation of novel items into current knowledge. *Neuropsychologia* **185**, 108569. (doi:10.1016/j.neuropsychologia.2023.108569)
74. Bartlett FC. 1932 *Remembering: a study in experimental and social psychology*. Cambridge, UK: Cambridge University Press.
75. Bromis K, Raykov P, Wickens L, Roseboom W, Bird CM. 2021 The neural representation of events is dominated by elements that are most reliably present. *J. Cogn. Neurosci.* **34**, 517–531. (doi:10.1162/jocn_a_01802)
76. Raykov PP, Varga D, Bird CM. 2023 False memories for ending of events. *J. Exp. Psychol. Gen*. **152**, 3459–3475. (doi:10.1037/xge0001462)
77. Zeng T, Tompary A, Schapiro AC, Thompson-Schill SL. 2021 Tracking the relation between gist and item memory over the course of long-term memory consolidation. *eLife* **10**, e65588. (doi:10.7554/eLife.65588)
78. Aitken F, Kok P. 2022 Hippocampal representations switch from errors to predictions during acquisition of predictive associations. *Nat. Commun*. **13**, 3294. (doi:10.1038/s41467-022-31040-w)
79. Varga DK, Raykov PP, Jefferies E, Ben-Yakov A, Bird CM. 2025 Hippocampal mismatch signals are based on episodic memories and not schematic knowledge. *Proc. Natl Acad. Sci. USA* **122**, e2503535122. (doi:10.1073/pnas.2503535122)
80. Lu Q, Hasson U, Norman KA. 2022 A neural network model of when to retrieve and encode episodic memories. *eLife* **11**, e74445. (doi:10.7554/eLife.74445)
81. Huang S *et al*. 2024 Cortico-hippocampal interactions underlie schema-supported memory encoding in older adults. *BioRxiv*.
82. Yacoby A, Greve A, Schaverien E, Shtoots L, Henson R, Levy D. 2025 The Effects of Prior Knowledge on Episodic Memory Depend on Retrieval Goals and Orientation. *PsychArchives* (doi:10.23668/psycharchives.16493)
83. Chen G, Cai Z, Taylor PA. 2024 Through the lens of causal inference: decisions and pitfalls of covariate selection. *Apert. Neuro* **4**, 10-52294. (doi:10.52294/001c.124817)
84. Raykov P. 2025 *Predictions and declarative memory encoding: two fMRI paradigms provide slim pickings for SLIMM*. See https://osf.io/ut4m9/overview.
85. Raykov P, Gurunandan K, Moniri M, Jain A, Greve A, Henson R. 2026 Supplementary material from:Predictions and declarative memory encoding: two fMRI paradigms provide slim pickings for SLIMM. Figshare (doi:10.6084/m9.figshare.c.8547874)