

# Shape of U: The Nonmonotonic Relationship Between Object–Location Memory and Expectedness



Jörn Alexander Quent<sup>1,2</sup>, Andrea Greve<sup>1</sup>, and Richard N. Henson<sup>1,3</sup>

<sup>1</sup>MRC Cognition and Brain Sciences Unit, University of Cambridge; <sup>2</sup>Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University; and <sup>3</sup>Department of Psychiatry, University of Cambridge

Psychological Science  
1–14

© The Author(s) 2022



Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/09567976221109134

[www.psychologicalscience.org/PS](http://www.psychologicalscience.org/PS)



## Abstract

The *schema-linked interactions between medial prefrontal and medial temporal lobe* (SLIMM) model predicts that memory for object locations is a U-shaped function of the expectancy of those locations. Using immersive virtual reality, we presented participants with 20 objects in locations that varied in their congruency with a kitchen schema. Bayes factors across four experiments (137 adults in total) confirmed the (preregistered) prediction of better memory for highly expected and unexpected locations relative to neutral locations. This U shape was found in location recall and in forced-choice recognition in which the foil locations were matched for expectancy, controlling for the bias toward guessing expected locations. A second prediction was that the two ends of the U shape are associated with different expressions of memory: recollection of unexpected locations and familiarity for expected locations. BFs, propagated across experiments, provided evidence against this second prediction; recollection was associated with both ends of the U shape. These findings further constrain theories about the role of schema in episodic memory.

## Keywords

prediction error, schema, virtual reality, memory, open data, open materials, preregistered

Received 11/15/21; Revision accepted 4/12/22

Memory is not a *tabula rasa* onto which new experiences are inscribed; rather, what we encode into memory depends on what we already know. Activated knowledge about the world that is relevant to one's current situation has been called a *schema*, and there is a long literature on how schemas affect memory encoding and retrieval (Alba & Hasher, 1983; Bartlett, 1932; Bransford & Johnson, 1972; Ghosh & Gilboa, 2014; van Kesteren et al., 2012). As abstracted knowledge about recurring situations, schemas enable us to make predictions, such as what to expect when walking into a kitchen (compared to a bathroom, for example).

Numerous studies have shown that memory is better for information that fits a schema (Alba & Hasher, 1983; Anderson, 1981; Craik & Tulving, 1975). This *congruency effect* has been obtained using a broad range of

memoranda (e.g., Bein et al., 2015; Brod & Shing, 2019; van Buuren et al., 2014). At the same time, other studies show the opposite finding, of better memory for unexpected information (e.g., von Restorff, 1933).

However, this raises the question how both expected and unexpected information can both be remembered well. One reason for the advantage for expected information is no doubt the ability to use schema to generate possible occurrences during retrieval—for example, to make a guess when memory fails (e.g., Bayen et al., 2000). Reasons for the advantage for unexpected information

## Corresponding Author:

Jörn Alexander Quent, Fudan University, Institute of Science and Technology for Brain-Inspired Intelligence

Email: [alexander.quent@rub.de](mailto:alexander.quent@rub.de)

likely include improved encoding following a surprise or prediction error (Brod et al., 2018; Greve et al., 2017) and improved retrieval due to increased distinctiveness of unexpected events (see Schmidt & Schmidt, 2017; Worthen & Hunt, 2006). A neuroscientific model called *schema-linked interactions between medial prefrontal and medial temporal lobe* (SLIMM; van Kesteren et al., 2012) offers an integrated account by assuming that different brain systems support memory under these two extremes, resulting in a U-shaped function of memory against expectancy: Unexpected events are hypothesized to trigger encoding of a snapshot of the entire episode (supported by the medial temporal lobes [MTL]), whereas expected events are hypothesized to benefit from rapid consolidation (into neocortex, facilitated by the medial prefrontal cortex). Consistent with the idea of different memory systems operating under different principles, our results in an earlier study showed various manipulations that dissociate the two ends of this U-shape (Greve et al., 2019).

Our previous study required participants to learn simple rules about the arbitrary values of different objects, rules whose consistency varied across three conditions (Greve et al., 2019). However, the conclusions were contingent on the middle (unrelated) condition being equivalent to the two extreme conditions in all ways other than rule congruency. Furthermore, the experimentally learned, artificial rules were extremely impoverished schemas compared with those used in the real world, which are typically derived from years of experience. We therefore sought to replicate the U-shaped function using much more realistic preexperimental schemas and ones that allow a continuous definition of expectancy. To this end, we employed immersive virtual reality (iVR) to “place” participants in a virtual kitchen and test their memory for the locations of objects, as a function of their prior knowledge of how likely objects were to appear in those locations. For example, a kettle might appear on the kitchen counter (expected), a kitchen table (neither strongly expected nor unexpected), or on top of a trash can (unexpected).

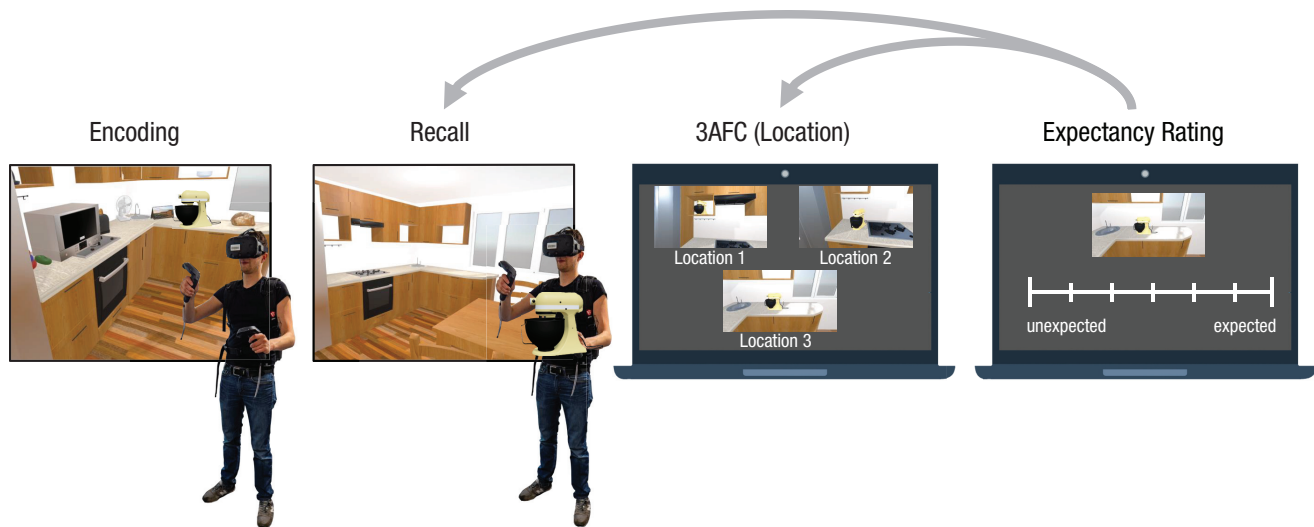
A U-shaped function of memory against expectancy could reconcile previous studies that have examined memory for object locations. For example, Brewer and Treyens (1981) showed that memory for objects was positively correlated with their schema expectancy, whereas other studies reported the opposite: better memory for atypical objects (Lampinen et al., 2001; Pezdek et al., 1989; Prull, 2015). With only two conditions, whether one finds one advantage or the other would depend on the relative position of experimental conditions along the U-shaped expectancy continuum (van Kesteren et al., 2012).

### Statement of Relevance

It is important to understand the factors that affect what we remember. We often remember surprising events that are not expected on the basis of our prior knowledge. However, we also remember events that conform to our expectations. A recent theory explains these paradoxical effects of prior knowledge (or “schemas”) by assuming that different memory systems support the encoding of expected and unexpected information. This theory predicts a U-shaped function of memory against expectancy. We confirmed this prediction by testing memory for the location of objects in a virtual reality kitchen, in which their expected locations varied continuously, on the basis of individual ratings. However, a secondary prediction of this theory—that the two ends of the U-shape would be associated with different phenomenological experiences (“recollection” versus “familiarity”)—was not confirmed. This work integrates the psychological literatures on schema, novelty, and surprise but indicates that further theorization is needed.

Probably the study most similar to the present one was reported by Lew and Howe (2017). They used photos of familiar room types, in which objects either stayed in the same location or shifted to a different location. Recognition of objects was better at unexpected locations, whereas recall of objects was better at expected locations. The authors speculated that schemas act differently on item and associative (location) memory, so unexpected locations attract attention but also activate schema-congruent bindings that interfere with memory (see also Bower et al., 1979). However, an alternative theory is that their recall, but not recognition, was influenced by guessing of room-congruent locations (e.g., Bayen et al., 2000). We controlled for this by using three-alternative forced choice (3AFC) recognition, as well as recall, in which an object was shown at one of three locations (one correct and two foils). Importantly, the locations were chosen to be equally expected on the basis of normative data, thus controlling for any bias toward guessing congruent locations when memory fails.

A second prediction of the SLIMM model concerns the type of memory associated with each end of the U shape. The episodic snapshot of unexpected events, encoded by the MTL system, is hypothesized to include contextual information that is incidental to the schema. Retaining such incidental information is advantageous



**Fig. 1.** Schematic overview of the paradigm, which differed across experiments only in object–location pairings and precise memory tests. The first step was encoding (in immersive virtual reality [iVR]): Participants explored a virtual kitchen (45 s) and were instructed to count and memorize 20 object locations. The second step was recall (in iVR): All 20 objects were removed, and participants were given one object at a time to be placed where previously seen. The third step was a three-alternative forced choice (3AFC; outside iVR): Each of the 20 objects was presented on a computer screen but in three alternative locations, where the two foil locations were expected approximately equally as the correct location (on the basis of normative ratings). Experiments 1, 2a, and 2b collected additional remember and familiar judgments. The fourth step was individual expectancy ratings (outside iVR), used for analysis of recall and 3AFC data, which were collected for all 20 objects for each of the 3AFC locations as well as their general expectancy in a kitchen context.

(if it keeps recurring) for updating the schema, so that the same prediction error is avoided in future. By contrast, incidental information is assumed to be lost when the cortical system encodes expected events, because this information is not part of the schema. The SLIMM model therefore predicts that memory of unexpected information will be accompanied by recollection of its episodic context, whereas memory for expected information will be associated with a feeling of familiarity, in the absence of contextual details.

We tested these two hypotheses across four experiments. The first experiment was a pilot to obtain a basic U-shape; the subsequent experiments were preregistered and added a variant of the remember/familiar procedure (Taylor & Henson, 2012; Tulving, 1985) to test the SLIMM model's second hypothesis that unexpected locations will be recollected whereas expected locations will seem familiar. Note that, although the spatial location of an object is often associated with recollection (Mayes et al., 2007), our instructions (see the Method section) emphasized that remember responses should be given only for other types of context (see the Discussion section for further theoretical consideration).

## Pilot Experiment

Memory was tested for 20 objects at different locations within a virtual kitchen, chosen from pilot ratings to

range from highly unexpected (−100) to highly expected (+100). Participants spent 45 s counting these objects, before all objects were removed, and participants replaced them one at a time at their remembered location (Fig. 1). This was followed by the 3AFC test outside iVR, using stills on a computer screen. To allow for individual differences in schema, we collected expectancy ratings in a final debriefing phase.

## Method

**Participants.** Participants were 16 adult Cambridge community members from the volunteer panel of the MRC Cognition and Brain Sciences Unit (<https://www.mrc-cbu.cam.ac.uk/take-part/>), all of whom reported normal or corrected-to-normal visual acuity, provided informed consent, and received monetary compensation for participation, as approved by a local ethics committee (CPREC 2020.018). There were eight males and eight females, and their mean age was 26.38 years ( $SD = 3.52$ ).

**Materials.** Stimuli comprised 12 kitchen objects and 8 nonkitchen objects (see Table S1 in the Supplemental Material available online) inspired by Lew and Howe (2017). A normative study was run to obtain 400 expectancy ratings in total; one for each of the 20 objects at each of the 20 possible locations in the kitchen (see Section S2 in the Supplemental Material). From these, a single assignment of each object to one of the 20 locations was chosen so as to

maximize the range of expectancies (see Section S3 in the Supplemental Material). The size of the virtual kitchen was 5.15 by 4.40 virtual meters (vm), where 1 vm corresponds to approximately 1 m in the real world (material available at <https://osf.io/4sw2t/>).

**Procedure.** The basic paradigm for all experiments is illustrated in Figure 1 (for a video of the VR task, see <https://vimeo.com/645994321>). During the encoding phase, participants were asked to navigate freely through a virtual kitchen for 45 s, with the instruction to count and memorize the locations of 20 objects that were scattered across the room. Following encoding, participants entered a blank room for approximately 2.5 min to practice how to place new objects (simple cubes) using the iVR hand controls, a skill that was needed in the subsequent recall phase. This also acted as a distractor phase to minimise short-term memory/rehearsal. In the recall phase, participants reentered the kitchen (now without the 20 original objects), were given one object, and were asked to place it at its previously seen location. Once placed, the object disappeared, and the process was repeated for the remaining 19 objects. Participants were encouraged to guess if they were unsure but could skip an object if they did not remember the object at all (a miss). For recall accuracy, we calculated the Euclidean distance between each of the 20 canonical object locations and the location where the object was placed by the participant. If the correct location was the closest of these 20, recall was scored as correct; otherwise it was scored as incorrect (equivalent results were found when using the continuous Euclidean metric of distance from the correct location, as shown in the Supplemental Material; Section S7.6 and Table S7.7).

Recall was followed by a 3AFC recognition test, performed on a computer outside iVR. Each trial showed one studied object in three alternative locations, one of which was correct. Importantly, the target and two foil locations were matched in expectancy according to the normative ratings, so using prior knowledge to guess the location could not help performance. Participants indicated which they thought was the studied location, followed by a rating of their confidence on a 3-point scale (1 = *did not see the object*, 2 = *guess the object was there*, 3 = *know the object was there*). Preliminary analysis of confidence did not add any new information (see Table S7.1 in the Supplemental Material), so we combined trials that were given a rating of 2 or 3 (i.e., excluding rare trials in which objects were forgotten). In the final phase, participants provided *expectancy ratings* for how likely they thought it would be to find each of the 20 objects in each of the three locations tested in 3AFC, together with an additional rating of the *general expectancy* of an object appearing anywhere in a kitchen at all. Ratings were collected

with a sliding scale from unexpected (−100) to expected (+100). These ratings were analogous to the normative ratings but allowed for potential individual differences in expectancy.

**Statistical analysis.** Statistical analysis was performed in R (R Core Team, 2018) using Bayesian multilevel models with *brms* (Bürkner, 2018, version 2.16.3) based on Stan (Carpenter et al., 2017). All analyses, scripts, and data are available at <https://osf.io/4sw2t/>.

Memory for individual trials was modeled as a function of a participant's object–location expectancy rating. Memory was a binary outcome (correct/incorrect), fitted using logistic regression models with the Bernoulli linking function. A full model was fitted first, with random slopes and intercepts for both objects and participants. Bayes factors (BFs) using marginal likelihoods from bridge sampling (Gronau et al., 2017) were then used to compare the full model with the model with random intercepts only, which was in turn compared with the model without random intercepts.

The linear and quadratic terms based on individually defined expectancy ratings (see Section S4 in the Supplemental Material) were scaled to have a standard deviation of 0.5 (see Section S5 in the Supplemental Material for further details), and the prior for each regression coefficient was based on a Student's *t* distribution, with hyperparameters of  $df = 7$ ,  $\mu = 0$ ,  $\sigma = 1$ , except for the intercept, which had hyperparameters of  $df = 7$ ,  $\mu = 0$ ,  $\sigma = 10$  (for justification, see <https://jaquent.github.io/post/the-priors-that-i-use-for-logistic-regression-now/>). These generic, weakly informative shrinkage priors are chosen to regularize unexpectedly large effects (Gelman et al., 2008). Eight Markov chain Monte Carlo chains were run, with 2,000 warm-up and 16,000 regular iterations and a total of 112,000 post-warm-up samples for each main model. All models converged with an  $\hat{R}$  of 1.

Evidence for or against our hypotheses was quantified by BFs for the linear and quadratic component of a second-order polynomial expansion of expectancy. A perfectly symmetrical U shape would have a positive quadratic coefficient and a zero linear coefficient (see Experiment 2b for a more stringent test based on opposite signs of interrupted linear regression). We also tested whether a cubic component was needed, but BFs provided no evidence for this, even when pooling data across experiments (see Section S7.8 in the Supplemental Material).

The BF for each coefficient was estimated by the Savage-Dickey ratio (Wagenmakers et al., 2010). The test for the quadratic term was order-restricted (one-tailed), in line with our hypothesis; all other tests were not order restricted. For order-restricted tests, we compared the density of the truncated and renormalized



prior distributions at zero with the logspine nonparametric density estimate of the truncated and renormalized posterior distributions of our parameters at zero (based on the 112,000 post-warm-up samples). For unrestricted comparisons, BFs were just density ratios at zero of prior/posterior ( $BF_{10}$ ) or posterior/prior ( $BF_{01}$ ). The Savage-Dickey ratio function used can be found in this GitHub repository: <https://github.com/JAQuent/assortedRFunction>. Evidence from BFs was categorized as inconclusive ( $BF_{10} > 1/3 - BF_{10} < 3$ ), anecdotal ( $BF_{10} > 3 - BF_{10} < 6$ ), moderate ( $BF_{10} > 6 - BF_{10} < 10$ ), or strong ( $BF_{10} > 10$ ). In addition to BFs, we report 95% credible intervals (CIs) for parameter estimates.

Finally, when testing the means across trials, BFs were derived from Bayesian *t*-tests (Morey & Rouder, 2018) with the package *BayesFactor* (version 0.9.12-4.2) and the default scale parameter of  $\sqrt{2}/2$ .

## Results

Objects that were reported as “not seen” were excluded from further analysis. A mean across participants and recall/3AFC tasks of 2.84 ( $SD = 1.96$ ) objects were excluded.

For the recall data (Fig. 2a; 256 trials), a model with random intercepts for participants and objects was used, because this was favored relative to one that also included random slopes ( $BF = 3.16 \times 10^5$ ) and relative to one that did not have random intercepts ( $BF = 50.2$ ). This suggests that the shape of the relationship between memory and expectancy was the same across participants and objects, but the baseline level of memory (intercept) varied across participants and objects. In the favored model, there was little evidence for or against a linear effect,  $BF_{10} = 0.383$ ,  $\beta = 0.0545$ , 95% CI =  $[-0.708, 0.846]$ , but more importantly there was strong evidence for the predicted positive quadratic effect,  $BF_{10} = 62.7$ ,  $\beta = 1.05$ , 95% CI =  $[0.363, 1.76]$ .

It is possible that the advantage in retrieving expected locations reflects a bias toward guessing such locations (based on preexperimental knowledge rather than the study phase). To test this, we compared the mean expectancy of incorrectly versus correctly recalled locations. The mean expectancy of incorrectly recalled locations was +40.7 ( $SD = 14.4$ ), which was greater than zero and clearly favored expected locations. Importantly, this expectancy was also greater than that for correctly recalled locations, +2.90 ( $SD = 10.6$ ),  $BF_{10} = 2.39 \times 10^4$ ,  $d = 2.19$ , representing a mix of expected and unexpected locations. Hence, participants were biased to report expected locations when unsure.

The 3AFC foils were designed to control for this expectancy bias. Similar to recall, model comparison showed that the 3AFC data were better fitted by a model

with random intercepts but not random slopes ( $BF = 289$ ) and compared with a model with no random intercepts ( $BF = 29.9$ ; 272 trials). This model again showed little evidence for or against a linear effect,  $BF_{10} = 0.477$ ,  $\beta = 0.28$ , 95% CI =  $[-0.49, 1.10]$ . However, in this case, the evidence for a quadratic effect was also inconclusive,  $BF_{10} = 1.08$ ,  $\beta = 0.377$ , 95% CI =  $[-0.264, 1.02]$  (Fig. 2b).

## Discussion

This pilot experiment confirmed the predicted U-shape function for recall of object locations as a function of the expectancy of those locations. However, this U shape was not replicated in 3AFC recognition in which foils were matched for expectancy, suggesting that this U shape might in part reflect retrieval-related processes, such as a bias toward guessing expected locations. Nonetheless, as evidence from the 3AFC data was moot, we collected more data in Experiment 1 (preregistered).

## Experiment 1

This experiment was powered to have a better chance of detecting a quadratic component in the 3AFC task and hence ruling out a guessing account of the expectancy advantage in recall. Furthermore, a potential confound in the pilot experiment was that kitchen objects tended to be in highly congruent or incongruent locations, whereas nonkitchen objects primarily occupied neutral locations. Although separate intercepts were allowed for each object, which should capture average differences between kitchen and nonkitchen objects, it is possible that the type of object interacted with expectancy in a more complex way. In Experiment 1 therefore, object locations were reselected to cover the range of expectancy values more uniformly for both kitchen and nonkitchen objects.

The only other important difference in Experiment 1 was that participants indicated the quality of their memory by using remember/familiar/guess judgments (Gardiner et al., 2002; Rajaram, 1993; Taylor & Henson, 2012). This was to test the second hypothesis of the SLIMM model, that is, that the two ends of the U shape are associated with different types of memory.

## Method

This experiment was preregistered on OSF (<https://osf.io/s9er3/>).

**Participants.** We collected data from 25 new participants from the same community as in the pilot experiment. These had a mean age of 24.52 years ( $SD = 2.83$ ), with 18 females, six males, and one nonbinary. Sample

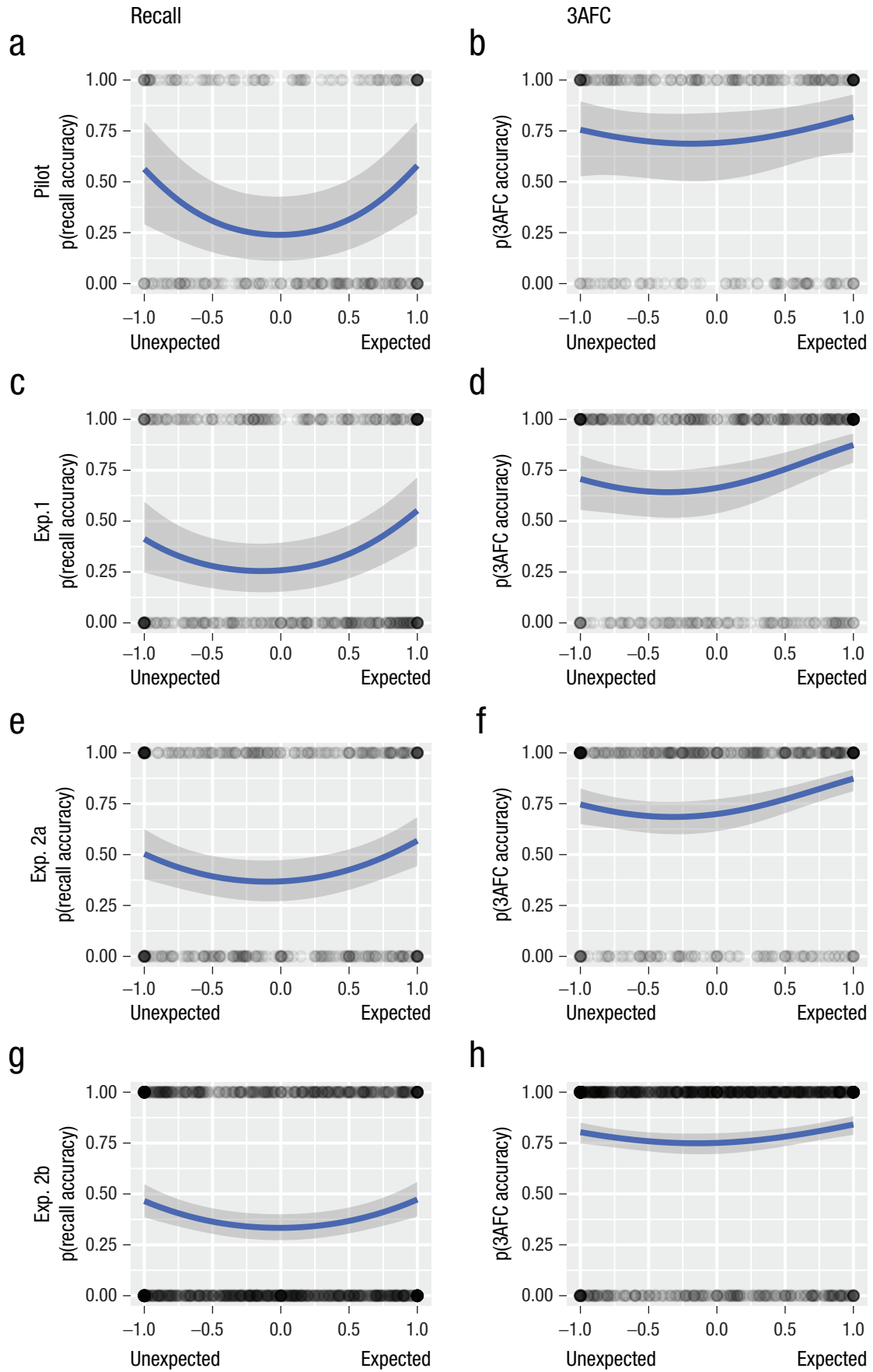


Fig. 2. (continued on next page)

**Fig. 2.** Accuracy results: recall (left column) and recognition (right column) against individual expectancy ratings for the pilot experiment and Experiments 1, 2a, and 2b (rows). The blue line represents the predicted second-order polynomial relationship, using evidence propagated across experiments; the shaded area around the blue line represents the 95% credible intervals of the prediction. Expectancy ratings originally ranged from  $-100$  to  $+100$  but were scaled to have a mean of zero and a standard deviation close to 0.5 (see the Method section of Experiment 1). The dots illustrate the density of individual trials for which location memory was correct (top) or incorrect (bottom). 3AFC = three-alternative forced choice.

size was determined on the basis of a frequentist power analysis to achieve 80% power (<https://osf.io/gr98d/>)<sup>1</sup> based on the quadratic effect size ( $\beta = 0.51$  on unit scale) for 3AFC (<https://osf.io/s9er3>).

**Materials.** The only change from the Pilot was that objects were reshuffled to other locations, so that kitchen and nonkitchen objects were more evenly distributed across (normative) expectancy ratings (see Section S3.2 in the Supplemental Material).

**Procedure.** The only procedural change was how participants categorized their memory responses for both recall and 3AFC (replacing the previous categories based on confidence only): If they did not remember seeing the object at all, they were told to indicate “no memory.” As in Experiment 1, these no-memory responses were excluded from further analysis. If they remembered the object itself, but had little idea where it was, they were to indicate “guess.” If they did not initially remember where it was, but the location just looked familiar once they had placed it (in recall) or compared the three choices (in 3AFC), they were to indicate “familiar.” Finally, if they immediately remembered where the object was when they saw it (because, for instance, they remembered what they thought when they saw the object), then they were to indicate “remember.” The precise instructions are given in Section S6.1 in the Supplemental Material.

**Statistical analysis.** Statistical analysis was identical to the pilot experiment, with BFs based on the same zero-centered priors. However, in order to propagate evidence across experiments, we also used the posterior distributions of the pilot experiment as prior distributions by estimating the family-specific parameters of the Student's  $t$  distribution (see Section S5 in the Supplemental Material). We used the marginal distribution, ignoring any correlation between parameters, and we used the same factors to scale expectancy ratings as in the pilot experiment, to ensure comparable expectancy ratings across experiments and hence correct priors (so the standard deviation was approximately 0.5). The BF from this second model allowed us to update the posterior belief in favor of our hypotheses ( $PB_{10}$ ), given the data from both experiments. See Section S7.8 in the Supplemental Material for confirmation that combining BFs gave very similar

results to estimating the BF on pooled data (i.e., little bias was introduced by ignoring the conditional dependencies between parameters). See also Section S7.6 in the Supplemental Material for evidence for similar conclusions when using pooled frequentist analysis.

Remember and familiar judgments were initially analyzed in line with preregistered analysis of the mean expectancy rating for remember and familiar judgments, but further simulation showed that this trial-averaged analysis is biased by boundary effects on expectancy values (see Section S7.2 in the Supplemental Material). Therefore, we analyzed them using the same single-trial logistic regression model as for overall accuracy. To estimate the probability of recollection, we used an outcome of 1 for remember judgments and an outcome of 0 otherwise. There is debate over the best way to estimate familiarity—that is, whether familiarity and recollection are redundant, independent, or exclusive (Knowlton & Squire, 1995). To model redundancy, we estimated familiarity with an outcome of 1 for remember or familiar responses and 0 otherwise; to model independence, we used an outcome of 1 for familiar responses, but only trials that were not given a remember response were included; to model exclusivity, we estimated familiarity as for independence and recollection by an outcome of 1 for remember responses, but only trials that were not given a familiar response were included. In the main text, we report results from the most popular independence model (Yonelinas & Jacoby, 1995), but the results for the redundant and exclusive models are shown in Sections S7.2 through S7.4 in the Supplemental Material and did not affect the conclusions.

## Results

The numbers of remember, familiar, guess, and no-memory responses are shown in Table S6.2 in the Supplemental Material, and individual expectancy ratings for kitchen and nonkitchen objects, which were more evenly spread across expectancy, as intended, are shown in Figure S4.

For recall (Fig. 2c; 418 trials), there was again little evidence for or against a linear term,  $BF_{10} = 1.52$ , even when combined with the pilot experiment,  $PB_{10} = 0.80$ ,  $\beta = 0.385$ , 95% CI =  $[-0.0936, 0.877]$ . There was only suggestive evidence for a positive quadratic

term,  $BF_{10} = 2.00$ , but this strengthened the evidence when combined with the pilot experiment,  $PB_{10} = 84.6$ ,  $\beta = 0.715$ , 95% CI = [0.272, 1.16].

As in the pilot experiment, there was a bias for incorrect recall to be drawn to expected locations (see Section S7.2 in the Supplemental Material). Focusing on the 3AFC, therefore (Fig. 2d; 447 trials), there was evidence for a linear term,  $BF_{10} = 37.1$ , even when combined with evidence from the pilot experiment,  $PB_{10} = 19.9$ ,  $\beta = 0.746$ , 95% CI = [0.254, 1.25], but more importantly there was also evidence for a positive quadratic term,  $BF_{10} = 7.66$ , which strengthened the evidence from the pilot experiment,  $PB_{10} = 12.2$ ,  $\beta = 0.538$ , 95% CI = [0.137, 0.947]. The positive linear term produced an asymmetry in the U shape toward expected locations.

Assuming that recollection and familiarity are independent, the estimate of recollection from 3AFC responses (Fig. 3a; 447 trials) showed no evidence for a linear term,  $BF_{10} = 0.452$ ,  $\beta = 0.288$ , 95% CI = [-0.22, 0.812], but strong evidence for a quadratic term,  $BF_{10} = 24.9$ ,  $\beta = 0.715$ , 95% CI = [0.252, 1.19]. By comparison, the estimate of familiarity (Fig. 3b; 227 trials) showed no evidence for a linear term,  $BF_{10} = 0.292$ ,  $\beta = 0.0932$ , 95% CI = [-0.471, 0.666], nor for a quadratic term,  $BF_{10} = 0.278$ ,  $\beta = 0.0581$ , 95% CI = [-0.495, 0.621]. See Section S7.2 in the Supplemental Material for similar results under different assumptions about the relationship between recollection and familiarity.

## Discussion

Experiment 1 replicated the U shape in recall, using a more even distribution of kitchen and nonkitchen objects across expectancy values, and now also found a U shape in 3AFC, which rules out a contribution from guessing expected locations. This confirmed the SLIMM model's first hypothesis. However, the results did not support the SLIMM model's second hypothesis: rather than finding that the advantage for unexpected locations was accompanied by recollection and the advantage for expected locations was accompanied by familiarity, we observed that both ends of the U shape were accompanied by increased recollection, whereas familiarity showed little effect of expectancy.

However, even though the set of 20 object–location pairings changed from the pilot experiment, the same pairings were used for all participants in Experiment 1, still leaving the possibility that the U shape was a quirk of specific pairings. Therefore, we attempted to replicate the U shape across five new sets of object–location pairings, counterbalanced across participants.

## Experiment 2a

### Method

Experiments 2a and 2b used the same design, differing only in that the participants in Experiment 2b also performed an unrelated task prior to the iVR experience. That design was the same as in Experiment 1 except that a larger range of object–location pairings was used. Experiment 2a was preregistered at <https://osf.io/kcr2q/>.

**Participants.** We recruited 25 participants from the same community population as in previous experiments, but data from one were lost because of experimenter error. Hence a total of 24 participants (11 females and 13 males, mean age = 25 years,  $SD = 3.71$  years) were analyzed, with the caveat that the five stimulus sets were not perfectly counterbalanced across participants.

**Materials.** The method was identical to that of Experiment 2, except that five new sets of 20 object–location pairings were based on the normative data (Sets 3–7; see Table S3 in the Supplemental Material), each chosen to maximize the range of expectancy values (materials are available at <https://osf.io/4sw2t/>).

### Results

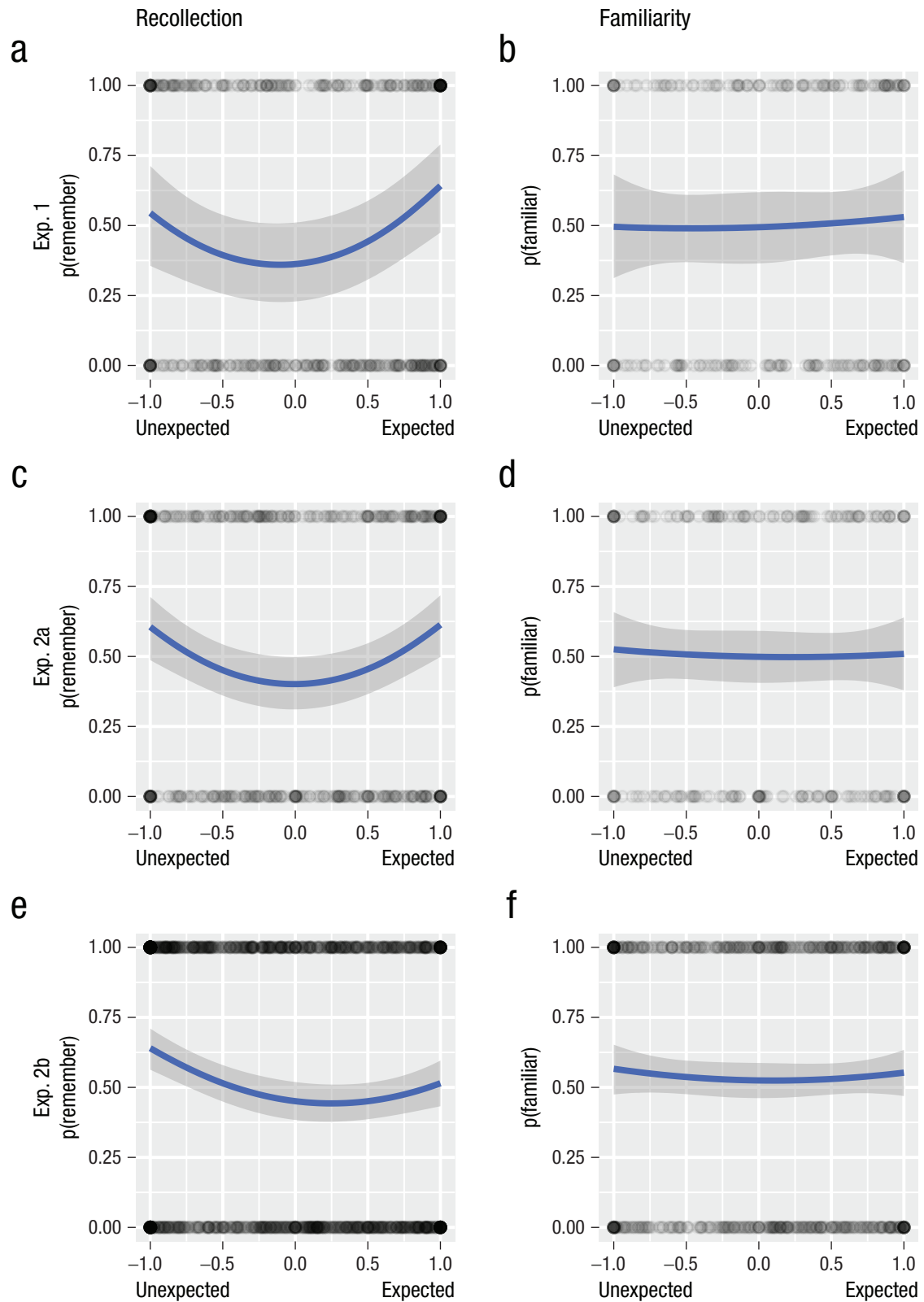
The numbers of remember, familiar, guess, and no-memory responses are shown in Table S6.2 in the Supplemental Material.

For recall (Fig. 2e; 410 trials), the results showed anecdotal evidence against a linear term,  $BF_{10} = 0.211$ ,  $PB_{10} = 0.28$ ,  $\beta = 0.179$ , 95% CI = [-0.153, 0.511]. There was little new evidence for a quadratic term,  $BF_{10} = 0.69$ , though the combined evidence remained very strong,  $PB_{10} = 48.9$ ,  $\beta = 0.497$ , 95% CI = [0.181, 0.812]. As in the previous experiments, there was a bias for incorrect recall to be drawn toward expected locations (see Section S7.3 in the Supplemental Material).

For 3AFC (Fig. 2f; 448 trials), there was little new evidence for or against a linear term,  $BF_{10} = 0.85$ , though the combined evidence for a (positive) linear component remained strong,  $PB_{10} = 43$ ,  $\beta = 0.597$ , 95% CI = [0.242, 0.952]. There was also little support for a quadratic term,  $BF_{10} = 1.45$ , but the combined evidence remained very strong,  $PB_{10} = 32.3$ ,  $\beta = 0.486$ , 95% CI = [0.179, 0.793].

Recollection estimates from 3AFC (Fig. 3c, 448 trials) showed moderate evidence against a linear term,  $BF_{10} = 0.279$ , when combined with the previous experiments,  $PB_{10} = 0.162$ ,  $\beta = 0.0234$ , 95% CI = [-0.315, 0.362], but





**Fig. 3.** Estimates for recollection (left column) and familiarity (right column), assuming they are independent processes (see the Method section) from the three-alternative forced choice (3AFC) task for Experiments 1, 2a and 2b (rows). The blue line represents the predicted second-order polynomial relationship using evidence propagated across experiments; the shaded area around the blue line represents the 95% credible interval of the prediction. Expectancy ratings originally ranged from  $-100$  to  $+100$  but were scaled to have a mean of zero and a standard deviation of  $\approx 0.5$  (see the Method section).

more importantly, continued evidence for a quadratic term,  $BF_{10} = 4.71$ , which was extremely strong when combined with the evidence so far,  $PB_{10} = 403$ ,  $\beta = 0.634$ , 95% CI = [0.312, 0.959]. For familiarity (Fig. 3d; 218 trials), on the other hand, there was continued anecdotal evidence against a linear component,  $BF_{10} = 0.341$ ,  $PB_{10} = 0.207$ ,  $\beta = -0.0413$ , 95% CI = [-0.462, 0.375], and continued inconclusive evidence regarding the quadratic term,  $BF_{10} = 0.286$ ,  $PB_{10} = 0.205$ ,  $\beta = 0.054$ , 95% CI = [-0.347, 0.456]. See Section S7.3 in the Supplemental Material for results under different assumptions about the relationship between recollection and familiarity.

## Discussion

Experiment 2a confirmed the U-shape function for recall and recognition as a function of object–location expectancy, at least when evidence was propagated from the previous experiments, and confirmed that both ends of this U shape were associated with recollection. However, separate evidence for a U shape in 3AFC was not strong, possibly reflecting weaker effects for some of the new stimulus sets. There was also continued evidence (combined across all experiments) for an accompanying positive linear effect for 3AFC (see Fig. 2f). We therefore collected more data with these stimulus sets, also to enable more stringent tests of a U shape: namely, *interrupted regression*, which tests whether both ends of a U shape are independently reliable (see the Method section). Experiment 2b was identical to Experiment 2a except that participants had previously studied a list of unrelated words as part of an experiment testing novel iVR experiences on preceding information (Quent & Henson, 2022).

## Experiment 2b

### Method

**Participants.** A total of 72 participants (50 females, 21 males, and one nonbinary; mean age = 26.12 years,  $SD = 6.53$  years) were recruited from the same community population as in previous experiments, approximately counterbalanced across the five stimulus sets. Prior to this experiment, the participants made judgments about a list of 288 words. The words were not related to kitchens or the objects used in the iVR phase, and participants were told the words were not relevant to the iVR phase (see Quent & Henson, 2022, for more details). Experiment 2b was also preregistered on OSF (<https://osf.io/b9dqg/>).<sup>2</sup>

**Statistical analysis.** Interrupted regression (e.g., Simonsohn, 2018) is detailed in the Supplemental Material

(Section S8). To maximize power for this analysis, we pooled across all experiments.

Following the suggestion of a reviewer, we also analyzed guess responses as a function of expectancy. Combined across Experiments 1, 2a, and 2b, these showed a negative quadratic relationship with expectancy. However, this could simply be a consequence of remember responses showing the opposite pattern (i.e., fewer opportunities for guesses at extremes), so we refrain from discussing this further (see Fig. S7.5 in the Supplemental Material for results and further discussion).

## Results

The numbers of remember, familiar, guess, and no-memory responses are shown in Table S6.2 in the Supplemental Material.

For recall (Fig. 2g; 1,245 trials), the results showed strong evidence against a linear term,  $BF_{10} = 0.173$ , consistent with evidence combined across experiments,  $PB_{10} = 0.105$ ,  $\beta = 0.0203$ , 95% CI = [-0.197, 0.236]. By contrast, there was additional evidence for a quadratic term,  $BF_{10} = 3.95$ , resulting in extreme evidence after combining across experiments,  $PB_{10} = 387$ ,  $\beta = 0.413$ , 95% CI = [0.204, 0.623]. As in previous experiments, there was a bias for incorrect recall to be drawn toward expected locations (see Section S7.4 in the Supplemental Material).

For 3AFC (Fig. 2h; 1,302 trials), the evidence against a linear term,  $BF_{10} = 0.226$ , meant that the combined evidence remained inconclusive,  $PB_{10} = 0.41$ ,  $\beta = 0.182$ , 95% CI = [-0.0598, 0.421]. There was little new evidence for a quadratic term,  $BF_{10} = 0.362$ , but the combined evidence remained strong,  $PB_{10} = 13.2$ ,  $\beta = 0.319$ , 95% CI = [0.098, 0.54].

For recollection estimates from 3AFC under independent scoring (Fig. 3e; 1,302 trials), there was strong evidence for a linear term,  $BF_{10} = 1,300$ , which held even when combined with previous experiments,  $PB_{10} = 27$ ,  $\beta = -0.358$ , 95% CI = [-0.574, -0.147]. There was no additional evidence for a quadratic term,  $BF_{10} = 0.381$ , but the combined evidence remained strong,  $PB_{10} = 135$ ,  $\beta = 0.387$ , 95% CI = [0.179, 0.592]. For familiarity (Fig. 3f; 625 trials), the evidence remained against a linear component,  $BF_{10} = 0.18$ ,  $PB_{10} = 0.139$ ,  $\beta = -0.0356$ , 95% CI = [-0.306, 0.239], and (anecdotally) against the quadratic component,  $BF_{10} = 0.227$ ,  $PB_{10} = 0.175$ ,  $\beta = 0.1$ , 95% CI = [-0.159, 0.359]. See Section S7.4 in the Supplemental Material for results under different assumptions about the relationship between recollection and familiarity.

Finally, for the interrupted regression, which combined data across all experiments, recall showed the strongest effect with a breakpoint of +0.06 (+9.24 on

the original scale ranging from  $-100$  to  $100$ ), where the estimated leftward slope (toward negative expectancy values) was  $\beta = -0.778$ , 95% CI =  $[-1.30, -0.284]$ ,  $BF_{10} = 54.9$ , and the rightward slope was  $\beta = 0.687$ , 95% CI =  $[0.0642, 1.34]$ ,  $BF_{10} = 6.27$  (2,329 trials). For 3AFC, the best breakpoint of  $-0.26$  ( $-34.3$  in the original scale) had a leftward slope of  $\beta = -0.835$ , 95% CI =  $[-1.77, 0.116]$ ,  $BF_{10} = 3.93$ , and a rightward slope of  $\beta = 0.505$ , 95% CI =  $[0.105, 0.918]$ ,  $BF_{10} = 8.48$  (2,469 trials). Note that the BF for the leftward slope fell below the specified criterion of 6 (see the Method section). For recollection estimates of 3AFC, the best breakpoint of  $+0.22$  ( $+30$  in the original scale) had a leftward slope of  $\beta = -0.841$ , 95% CI =  $[-1.23, -0.454]$ ,  $BF_{10} = 620$ , and a rightward slope of  $\beta = 1.22$ , 95% CI =  $[0.37, 2.10]$ ,  $BF_{10} = 40.4$  (2,197 trials). These results provide strong support for both sides of the U shape being reliable in recall, and 3AFC, at least for recollection.

## Discussion

Combined with the evidence from the previous experiments, overall memory in Experiment 2b continued to show a U shape, with strong evidence for a positive quadratic component but evidence against a linear component. Importantly, this was now confirmed, at least for recall and for recollection estimates during 3AFC, by a more stringent test of interrupted regression. Interestingly, when restricting to recollection, evidence for a negative linear component also emerged. This suggests an additional bias toward better recollection of unexpected locations.

## General Discussion

We confirmed the first prediction of the SLIMM model that memory is a U-shaped function of the expectancy of an event, with better memory for highly expected or highly unexpected object locations. This replicates our findings from a paradigm in which participants learned a schema during the experiment (a rule about the relative value of two types of object; Greve et al., 2019), but importantly extends to more realistic, preexperimental schema (object locations within a kitchen) and to continuous, participant-specific measures of expectancy.

This U-shape function can reconcile previous studies that have reported a memory advantage for either unexpected or expected object locations (Brewer & Treyens, 1981; Lampinen et al., 2001; Lew & Howe, 2017; Pezdek et al., 1989; Prull, 2015) and integrate the broader literature that has tended to focus separately on the effects of schema congruency versus the effects of schema incongruency (surprising or novel information; van Kesteren et al., 2012).

Our paradigm also demonstrates the value of iVR, in which memory could be tested quickly (after less than a minute of encoding) and could benefit from participants' preexperimental knowledge (i.e., without the need for extensive training of new, artificial schema). Indeed, iVR enables one to test memory in more naturalistic situations while simultaneously providing the experimental control needed to measure memory (e.g., quantifying the error in recalled locations). Being "present" and actively exploring a virtual room is not only more like real life but also may result in stronger effects (e.g., of expectancy) than does showing participants photographs of rooms in which objects are in different places (e.g., Lew & Howe, 2017).

However, we did not confirm the second prediction of the SLIMM model, namely that one side of the U shape (high unexpectedness) is associated with recollection of contextual details whereas the other side (high expectancy) is associated with a feeling of familiarity. Our Bayesian analysis provided evidence against any expectancy effect on familiarity and evidence that recollection was higher for expected as well as unexpected locations. One possibility is that the SLIMM model's conception of the relationship between schema congruency and recollection/familiarity is incorrect. Another possibility is that memory for the location of objects (rather than the objects themselves) inherently requires the same associative mechanisms that support recollection (e.g., binding information across domains; Mayes et al., 2007). In other words, the specific measure of memory we used may have required retrieval of contextual information (supported by the SLIMM model's MTL system) for all levels of expectancy. We tried to prevent this by modifying our remember/familiar instructions so that remembering the location of an object was not sufficient for a remember response (requiring instead retrieval of other context, such as internal thoughts at encoding; see Section S6.1 in the Supplemental Material for precise instructions). However, it is possible that participants did not appreciate this distinction, or even that they simply remapped "remember" and "familiar" to different levels of confidence (Haaf et al., 2021). One way to explore this in future would be to test memory for the perceptual details of objects, rather than their location, where those details are incidental to the schema (e.g., silver or black color of a kettle). The SLIMM model predicts that such incidental perceptual details will be remembered when objects are highly unexpected, but not when highly expected, which is consistent with other findings in the literature that suggest unexpected stimuli are more likely to be recollected, whereas expected stimuli are more likely deemed familiar (Kafkas & Montaldi, 2018).

We chose iVR because of its ability to create rich and lifelike experiments in the laboratory. The paradigm therefore offers greater potential generalizability, at least compared with more typical memory experiments on lists of words. However, the results may not generalize to all cultures given that we tested a relative narrow demographic from the UK Cambridge community. In this context, Draschkow (2022) recently made an interesting proposal to use VR as a way of diversifying participation beyond the typical Western, educated, industrialized, rich, and democratic (WEIRD) sample used in laboratory experiments.

Our study raises interesting further questions. One question is whether the U shape applies to all types of expectancy or only to predictions deriving from preexisting knowledge (i.e., schema). For example, would the same U shape emerge for events that are expected or unexpected given an episodic context, such as a temporal sequence of items (cf. von Restorff, 1933)? Another question is whether the U shape varies across development or aging. For example, one might expect the advantage for schema-congruent events to increase during childhood (as knowledge increases), and the advantage for schema-incongruent events to decline in old age (if the MTL is particularly affected by aging). One limitation of the present study is that we tested only immediate memory. It is possible that the U shape changes with increased retention intervals, such as following consolidation processes that may occur overnight. Either way, our results reinforce the importance of a schema in shaping the encoding of new memories—either by facilitating encoding of expected information or by highlighting unexpected information for better encoding—and hence unify two topics (effect of congruency on memory and effect of surprise on memory) that have tended to be studied separately.

## Transparency

*Action Editor:* Sachiko Kinoshita

*Editor:* Patricia J. Bauer

### Author Contributions

All authors developed the study concept, study design and theoretical background. Testing and data collection were performed by J. A. Quent. J. A. Quent analyzed and interpreted the data under the supervision of R. N. Henson and A. Greve. All authors drafted the manuscript and approved the final version of the manuscript for submission.

### Declaration of Conflicting Interests

The authors declared no conflicts of interest with respect to the authorship or the publication of this article.

### Funding

This work was supported by the UK Medical Research Council (SUAG/046 G101400). J. A. Quent's PhD studentship was funded by the Gates Cambridge Trust.

## Open Practices

All data, code, and materials have been made publicly available via OSF and can be accessed at <https://osf.io/4sw2t/> and <https://zenodo.org/record/6614685>. The design and analysis plans for the experiments were pre-registered on OSF (Experiment 1: <https://osf.io/s9er3/>, Experiment 2a: <https://osf.io/kcr2q/>, Experiment 2b: <https://osf.io/b9dqq/>). The task and the fitted models can also be found here <https://zenodo.org/record/6614757>. This article has received the badges for Open Data, Open Materials, and Preregistration. More information about the Open Practices badges can be found at <http://www.psychologicalscience.org/publications/badges>.



## ORCID iD

Jörn Alexander Quent  <https://orcid.org/0000-0002-1800-5219>

## Supplemental Material

Additional supporting information can be found at <http://journals.sagepub.com/doi/10.1177/09567976221109134>

## Notes

1. Note that there is a typo in the OSF registration: The needed sample size was 25, not 24.
2. In that preregistration, we stated that we would test 96 participants, whereas in fact 96 was the maximum for a Bayesian sequential design, which happened to terminate before we reached 96 (i.e., at 72 participants). Note that the stopping criterion for that design concerned a separate hypothesis (Quent & Henson, 2022) that was independent of the U shape tested here.

## References

- Alba, J. W., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin*, 93(2), 203–231. <https://doi.org/10.1037/0033-2909.93.2.203>
- Anderson, J. R. (1981). Effects of prior knowledge on memory for new information. *Memory & Cognition*, 9(3), 237–246.
- Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge University Press.
- Bayen, U. J., Nakamura, G. V., Dupuis, S. E., & Yang, C. L. (2000). The use of schematic knowledge about sources in source monitoring. *Memory & Cognition*, 28(3), 480–500. <https://doi.org/10.3758/BF03198562>
- Bein, O., Livneh, N., Reggev, N., Gilead, M., Goshen-Gottstein, Y., & Maril, A. (2015). Delineating the effect of semantic congruency on episodic memory: The role of integration and relatedness. *PLOS ONE*, 10(2), Article e0115624. <https://doi.org/10.1371/journal.pone.0115624>
- Bower, G. H., Black, J. B., & Turner, T. J. (1979). Scripts in memory for text. *Cognitive Psychology*, 11(2), 177–220. [https://doi.org/10.1016/0010-0285\(79\)90009-4](https://doi.org/10.1016/0010-0285(79)90009-4)
- Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of



- comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, 11(6), 717–726.
- Brewer, W. F., & Treyens, J. C. (1981). Role of schemata in memory for places. *Cognitive Psychology*, 13(2), 207–230.
- Brod, G., Hasselhorn, M., & Bunge, S. A. (2018). When generating a prediction boosts learning: The element of surprise. *Learning and Instruction*, 55, 22–31.
- Brod, G., & Shing, Y. L. (2019). A boon and a bane: Comparing the effects of prior knowledge on memory across the lifespan. *Developmental Psychology*, 55(6), 1326–1337.
- Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, 10(1), 395–411. <https://doi.org/10.32614/RJ-2018-017>
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1). <https://doi.org/10.18637/jss.v076.i01>
- Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 104(3), 268–294.
- Draschkow, D. (2022). Remote virtual reality as a tool for increasing external validity. *Nature Reviews Psychology*, 1, 433–434. <https://doi.org/10.1038/s44159-022-00082-8>
- Gardiner, J. M., Ramponi, C., & Richardson-Klavehn, A. (2002). Recognition memory and decision processes: A meta-analysis of remember, know, and guess responses. *Memory*, 10(2), 83–98. <https://doi.org/10.1080/09658210143000281>
- Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y.-S. (2008). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics*, 2(4), 1360–1383. <https://doi.org/10.1214/08-AOAS191>
- Ghosh, V. E., & Gilboa, A. (2014). What is a memory schema? A historical perspective on current neuroscience literature. *Neuropsychologia*, 53, 104–114. <https://doi.org/10.1016/j.neuropsychologia.2013.11.010>
- Greve, A., Cooper, E., Kaula, A., Anderson, M. C., & Henson, R. (2017). Does prediction error drive one-shot declarative learning? *Journal of Memory and Language*, 94, 149–165. <https://doi.org/10.1016/j.jml.2016.11.001>
- Greve, A., Cooper, E., Tibon, R., & Henson, R. N. (2019). Knowledge is power: Prior knowledge aids memory for both congruent and incongruent events, but in different ways. *Journal of Experimental Psychology: General*, 148(2), 325–341. <https://doi.org/10.1037/xge0000498>
- Gronau, Q. F., Sarafoglou, A., Matzke, D., Ly, A., Boehm, U., Marsman, M., Leslie, D. S., Forster, J. J., Wagenmakers, E.-J., & Steingroever, H. (2017). A tutorial on bridge sampling. *Journal of Mathematical Psychology*, 81, 80–97. <https://doi.org/10.1016/j.jmp.2017.09.005>
- Haaf, J. M., Rhodes, S., Naveh-Benjamin, M., Sun, T., Snyder, H. K., & Rouder, J. N. (2021). Revisiting the remember-know task: Replications of Gardiner and Java (1990). *Memory & Cognition*, 49(1), 46–66. <https://doi.org/10.3758/s13421-020-01073-x>
- Kafkas, A., & Montaldi, D. (2018). Expectation affects learning and modulates memory experience at retrieval. *Cognition*, 180, 123–134. <https://doi.org/10.1016/j.cognition.2018.07.010>
- Knowlton, B. J., & Squire, L. R. (1995). Remembering and knowing: Two different expressions of declarative memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(3), 699–710.
- Lampinen, J. M., Copeland, S., & Neuschatz, J. (2001). Recollections of things schematic: Room schemas revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 1211–1222. <https://doi.org/10.1037/0278-7393.27.5.1211>
- Lew, A. R., & Howe, M. L. (2017). Out of place, out of mind: Schema-driven false memory effects for object-location bindings. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(3), 404–421. <https://doi.org/10.1037/xlm0000317>
- Mayes, A. R., Montaldi, D., & Migo, E. (2007). Associative memory and the medial temporal lobes. *Trends in Cognitive Sciences*, 11(3), 126–135. <https://doi.org/10.1016/j.tics.2006.12.003>
- Morey, R. D., & Rouder, J. N. (2018). *BayesFactor: Computation of Bayes factors for common designs*. <https://cran.r-project.org/package=BayesFactor>
- Pezdek, K., Whetstone, T., Reynolds, K., Askari, N., & Dougherty, T. (1989). Memory for real-world scenes: The role of consistency with schema expectation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(4), 587–595.
- Prull, M. W. (2015). Adult age differences in memory for schema-consistent and schema-inconsistent objects in a real-world setting. *Aging, Neuropsychology, and Cognition*, 22(6), 731–754. <https://doi.org/10.1080/13825585.2015.1037821>
- Quent, J. A., & Henson, R. N. (2022). Novel immersive virtual reality experiences do not produce retroactive memory benefits for unrelated material. *Quarterly Journal of Experimental Psychology*. Advance online publication. <https://doi.org/10.1177/17470218221082491>
- Rajaram, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory & Cognition*, 21(1), 89–102.
- R Core Team. (2018). *R: A language and environment for statistical computing* [Computer software]. <https://www.r-project.org/>
- Schmidt, S. R., & Schmidt, C. R. (2017). Revisiting von Restorff's early isolation effect. *Memory and Cognition*, 45(2), 194–207. <https://doi.org/10.3758/s13421-016-0651-6>
- Simonsohn, U. (2018). Two lines: A valid alternative to the invalid testing of U-shaped relationships with quadratic regressions. *Advances in Methods and Practices in Psychological Science*, 1(4), 538–555. <https://doi.org/10.1177/2515245918805755>
- Taylor, J. R., & Henson, R. N. (2012). Could masked conceptual primes increase recollection? The subtleties of measuring recollection and familiarity in recognition memory. *Neuropsychologia*, 50(13), 3027–3040.
- Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie Canadienne*, 26(1), 1–12. <https://doi.org/10.1037/h0080017>
- van Buuren, M., Kroes, M. C. W., Wagner, I. C., Genzel, L., Morris, R. G. M., & Fernández, G. (2014). Initial investigation of the effects of an experimentally learned schema on spatial associative memory in humans. *Journal of Neuroscience*, 34(50), 16662–16670.

- van Kesteren, M. T. R., Ruiter, D. J., Fernández, G., & Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends in Neurosciences*, 35(4), 211–219. <https://doi.org/10.1016/j.tins.2012.02.001>
- von Restorff, H. (1933). Über die Wirkung von Bereichsbildungen im Spurenfeld [On the effect of spheres formations in the trace field]. *Psychologische Forschung*, 18(1), 299–342. <https://doi.org/10.1007/BF02409636>
- Wagenmakers, E.-J., Lodewyckx, T., Kuriyal, H., & Grasman, R. (2010). Bayesian hypothesis testing for psychologists: A tutorial on the Savage–Dickey method. *Cognitive Psychology*, 60(3), 158–189. <https://doi.org/10.1016/j.cogpsych.2009.12.001>
- Worthen, J., & Hunt, R. R. (2006). *Distinctiveness and memory*. Oxford University Press.
- Yonelinas, A. P., & Jacoby, L. L. (1995). The relation between remembering and knowing as bases for recognition: Effects of size congruency. *Journal of Memory and Language*, 34(5), 622–643. <https://doi.org/10.1006/jmla.1995.1028>