Supplementary Material for *'Shape of U: The Non-monotonic Relationship Between Object-Location Memory and Expectedness'*

by Quent, Greve & Henson

S1. Objects Used

Kitchen Objects				**Non-Kitchen Objects**	
1 microwave	5 bowl of fruits	9 bread		13 towels	17 hat
2 kitchen roll	6 tea pot	10 glass jug		14 toy	18 helmet
3 saucepan	7 knife	11 mug		15 books	19 calendar
4 toaster	8 mixer	12 dishes		16 umbrella	20 fan

Table S1: List of kitchen objects and non-kitchen objects used.

S2. Normative study

In order to select where to place each object and to get a range of expectancy values, a group of six participants (five females and one male; Mean age 36.83 (SD = 2.14)) were shown screenshots of each object at each location (i.e., 400 trials in total) and asked to rate how expected that object was in that location, from -100 (unexpected) to +100 (expected). They were then asked to also rate the general expectancy of each object anywhere in a kitchen (a further 20 trials). Four additional objects (kitchen: peppers and white pot; non-kitchen: dumbbell and wrench) were used to create eight object/location practice trials, which were shown first to give participants an idea about the task and calibrate their ratings. Responses were given by moving a slider across a scale. Results are shown in Figure S2.

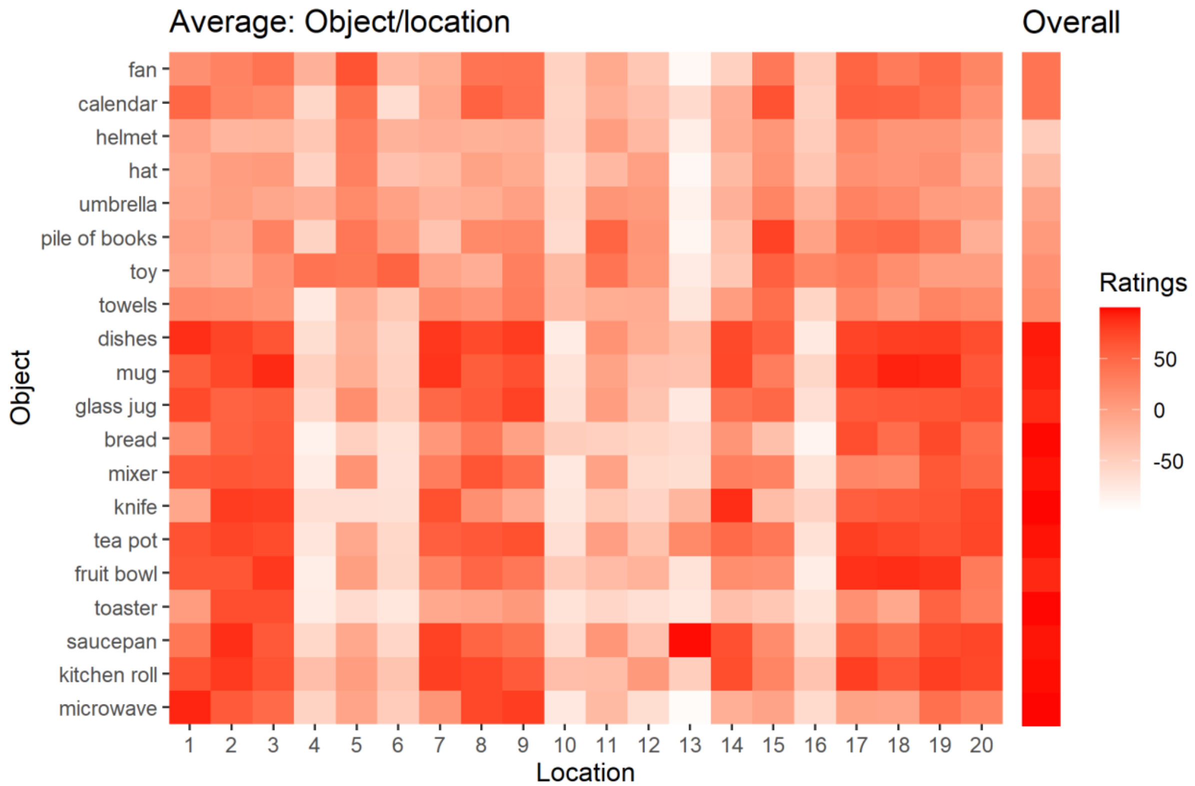

Figure S2: Mean normative rating of object-location for each object at each location: Individual ratings could vary from -100 to 100. The column titled 'overall' (right) shows average general expectancy rating for each object.

S3. Selecting object-locations based on normative study

S3.1. Pilot Experiment

The ratings were ranked from 1 to 400 within each participant (since different participants used different ranges of values), and the ranks were then averaged over participants. After that an algorithm was run with about 10 million iterations in which each object was randomly assigned to the 20 locations along with two foil locations that were randomly chosen (the foil locations were used in the 3AFC test of memory, and were chosen to have expectancies close to that of the target location). When the solution was not valid (e.g., the same location was selected for foil 1 and foil 2), this step was repeated until a valid solution was returned. For each valid solution, the sum of squared differences between the ranks of the targets and the intended uniform spread ("SS of targets") and the sum of squared differences between the ranks of the targets and both foils ("SS of foils") were calculated. In a two-step process, the number of solutions was first reduced to only include the 0.0001th percentile of the "SS of foils" distribution, to prioritize adequate foils that are similar in their expectancy to the target. Second, the remaining solutions were sorted by their "SS of targets" values and then checked for problems (e.g., an object not being visible). The first valid set of 20 object-location pairs that was found by this algorithm is shown in Table S3.

S3.2. Experiment 1

With regard to the environment, the only change from the Pilot Experiment was that objects were now reshuffled to other locations, so that kitchen and non-kitchen objects were more evenly distributed across (normative) expectancy ratings. To address the problem that 'kitchen' objects tended to be in extreme positions, only the "SS of targets" for kitchen objects was used; otherwise the algorithm remained unchanged. This ensured that kitchen objects now also occupied middle locations, while the spread of non-kitchen objects was still adequate. We will call this Set 2, which is distinct from the Set 1 used in the Pilot Experiment.

S3.3. Experiment 2a and 2b

There were no differences in the selection process for Experiment 2 (compared to Experiment 1) other than the fact that five sets were selected (i.e., 5 random winning iterations). Here we call these Sets 3-7, though in the code accompanying this paper, they are numbered by iteration in the simulation as 111, 246, 388, 498 and 848 respectively.

Object		**Locations**						
#	Name	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7
1	microwave	14	19	13	13	11	10	19
2	kitchen roll	8	17	11	16	4	19	2
3	saucepan	3	2	15	18	15	12	13
4	toaster	12	4	20	1	8	1	9
5	bowl of fruits	13	11	6	16	18	16	1
6	tea pot	18	5	5	7	2	15	4
7	knife	20	13	10	14	6	13	14
8	mixer	9	18	12	3	14	14	17
9	bread	4	15	19	11	16	4	16
10	glass jug	15	7	8	4	7	8	6
11	mug	17	3	3	19	20	3	11
12	dishes	6	10	9	12	10	20	10
13	towels	7	14	17	10	5	18	5
14	toy	11	12	2	6	1	9	15
15	books	2	1	4	8	19	7	7

16	umbrella	1	8	18	17	12	5	8
17	hat	16	6	1	2	17	11	20
18	helmet	5	9	16	20	3	6	12
19	calendar	10	20	14	5	13	17	3
20	fan	19	16	7	9	9	2	18

Table S3: This table lists the seven object-location sets which we used across the different experiments. Each set contained 20 objects (left column) which were shown across 20 different locations as indexed in the right columns which defined the different sets.

S4. Individual Expectancy Ratings

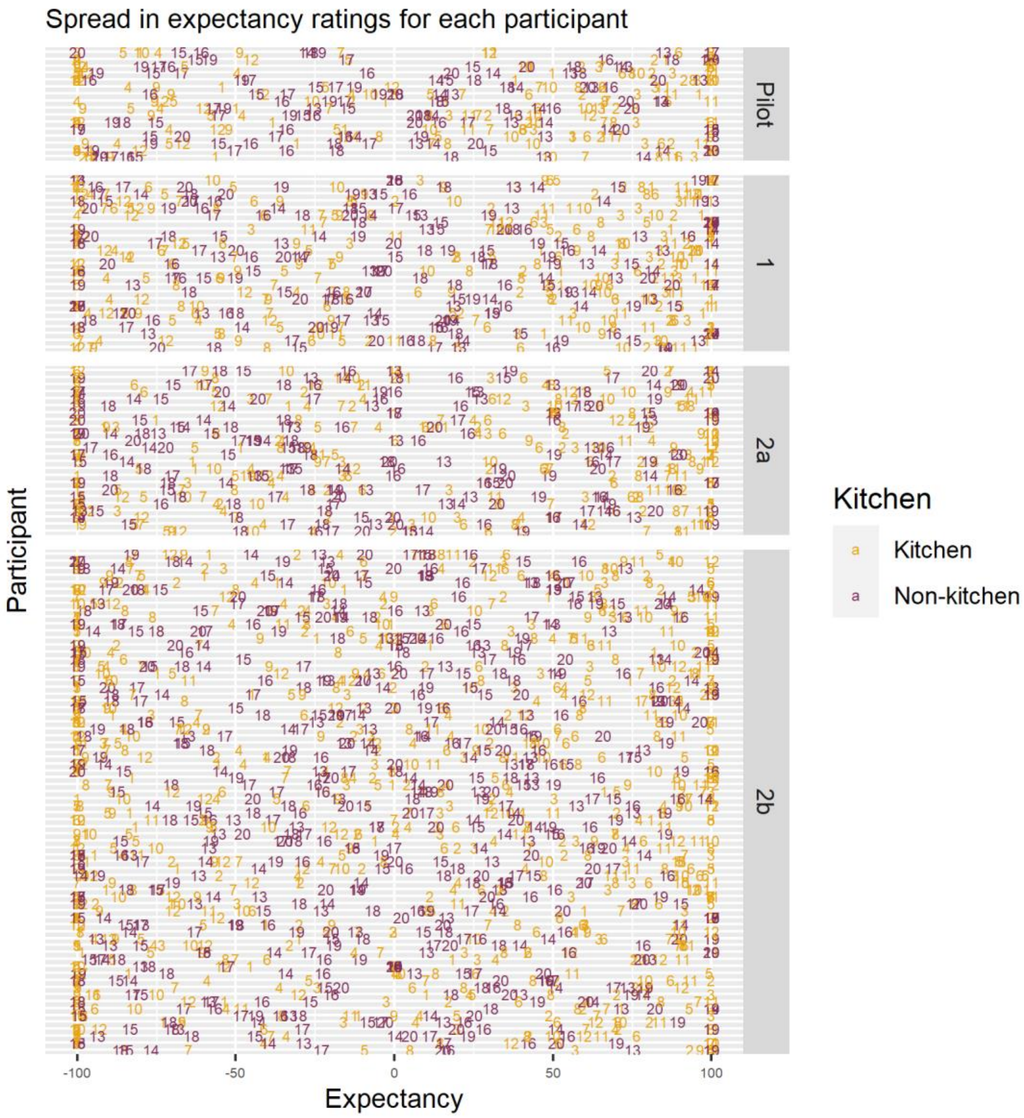

Figure S4: This figure illustrates the spread of all object-location expectancy ratings (x-axis) collected from each participant (y-axis) tested across all four experiments (i.e. panel labelled Pilot, 1, 2a and 2b). Numbers represent the objects (see Table S1) and are coloured orange if they are generally expected in the kitchen and purple if they are not.

S5. Priors for analysis propagating priors as posteriors

For the analyses of the Pilot Experiment, we wanted each polynomial term to have a SD of 0.5, to match the scale of the priors used. This was achieved by first dividing the raw expectancy values (which ranged from -100 to 100) by 100, so that they ranged from -1 to 1, and then mean-correcting them. These values, their squares and their cubes were then scaled separately to have zero mean and SD of 0.5. These scaling factors were stored and applied to the corresponding polynomial terms of Experiments 1-2, so that the posteriors from each experiment were scaled appropriately to function as priors for subsequent experiments (i.e., to ensure the scaling was identical across experiments). These priors are shown below.

		Pilot			**Exp. 1**			**Exp. 2a**			**Exp. 2b**		
Measure	Predictor	df	μ	σ	df	μ	σ	df	μ	σ	df	μ	σ
Recall	Intercept	7	0	10	14.41	-0.47	0.35	27.67	-056	0.27	73.74	-0.2	0.19
	Linear	7	0	1	72.36	0.05	0.39	123.47	0.39	0.25	121.99	0.8	0.17
	Quadratic	7	0	1	75.24	1.05	0.35	113.52	0.71	0.22	139.21	0.5	0.16
3AFC	Intercept	7	0	10	14.22	1.08	0.33	39.73	1.09	0.23	53.93	1.21	0.17
	Linear	7	0	1	80.92	0.28	0.4	91.81	0.75	0.25	129.49	0.6	0.18
	Quadratic	7	0	1	97.21	0.38	0.32	116.5	0.54	0.2	104.89	0.49	0.16
Recollection (independence/ redundancy)		Intercept			7	0	10	17.11	-0.09	0.26	39.58	0.01	0.17
		Linear			7	0	1	104.61	0.29	0.26	107.12	0.02	0.17
		Quadratic			7	0	1	83.67	0.71	0.24	84.47	0.63	0.16
Recollection (exclusivity)		Intercept			7	0	10	17.48	0.75	0.33	35.84	0.86	0.21
		Linear			7	0	1	82.28	0.32	0.33	101.41	0.02	0.22
		Quadratic			7	0	1	78.38	0.72	0.31	118.84	1.67	0.21
Familiarity (independence, exclusivity)		Intercept			7	0	10	14.96	00.02	0.19	26.98	0.02	0.14
		Linear			7	0	1	74.79	0.09	0.29	96.65	-0.04	0.21
		Quadratic			7	0	1	69.62	0.06	0.28	102.72	0.05	0.2
Familiarity		Intercept			7	0	10	18.26	1.22	0.23	35.52	1.28	0.16

(redundancy)	Linear		7	0	1		92.83	0.18	0.28		98.59	-0.05	0.19
	Quadratic		7	0	1		83.05	0.5	0.25		75.85	0.46	0.18

Table S5: List of priors for models reported in the main text.

S6. R/F/G/N-task

S6.1 instructions given to participants

As explained in the main text, we also asked participants to describe their subjective experience of memory with an R/F/G/N instructions. The following instruction was given to participants verbally:

"After you have placed the object, I will ask you how you made that decision. If you don't remember seeing the object at all, tell me "no memory". If you remember the object, but have no idea where it was, you still have to place the object where you think it was, so what you have to do is guess, which is what you then say. If you didn't initially remember where it was, but once you had placed somewhere, that location looked familiar, say "familiar". So it is a vague feeling of familiarity but you don't have any specific source to base your decision on. Finally, if you remembered where the object was immediately that you saw it, then say "remember". For instance that could be because you remember what you thought when you saw the object, you remember hearing something while you saw the object (e.g., the instructor talked to you) or you remembered the order in which you saw the objects etc."

S6.2 R/F/G/N distribution across experiments

Experiment	R/F/G/N	# of responses	Percent %
1	remember	220	44
	familiar	114	23
	guess	113	23
	no memory	53	11
2a	remember	230	48
	familiar	110	23
	guess	108	22

	no memory	32	7
2b	remember	677	47
	familiar	347	24
	guess	278	19
	no memory	138	10

Table S6.2: Distribution of R/F/G/N-responses in 3AFC tasks across the experiments.

S7. Supplementary analyses

All frequentist models (if not specified differently) were run using the R package lmerTest (Kuznetsova et al., 2017) and included random intercepts for objects and participants. The binomial family function was used to predict binary outcomes, while the gamma family function was used for Euclidean distance (placement error). To make the parameter estimates comparable, the same scaling was used as for the analyses reported in the main text.

S7.1. Pilot Experiment

We also performed a frequentist analysis with the probability of giving a high confidence response in the 3AFC task (analogous to the analysis of recollection in Experiments 1, 2a and 2b). This showed a significant quadratic relationship between object-location expectancy and high confidence responses (Table S7.1).

	β	SE	Z	P	Sig
Intercept	-0.55	0.28	-1.98	0.048	*
Linear	0.02	0.36	0.07	0.9449	
Quadratic	1.58	0.8	1.98	0.0472	*

Table S7.1: Results of frequentist model predicting high confidence 3AFC responses in Experiment 1.

S7.2. Experiment 1

Like in the Pilot Experiment, the average expectancy for incorrectly-placed objects was +34.49 (16.59) and so clearly higher than for correctly-placed objects, which was -4.54 (14.58), $BF_{10} = 1.74 \times 10^5$, $d = 1.51$.

Remember/familiar judgements were initially analysed in line with pre-registered analysis of the mean expectancy rating for remember and familiar judgments (https://osf.io/kcr2q), but further simulations showed that this trial-averaged analysis is biased by boundary effects on expectancy values (see https://jaquent.github.io/post/a-u-shape-that-appears-as-a-linear-correlation-when-averaged/). So instead, we fitted a model to individual trials, where the binary outcome was whether a response was given a remember or familiar judgment, and the trials that were included depended on one of three possible relationships between the theoretical concepts of recollection/familiarity and the response labels of "remember" and "familiar" (i.e., independent, redundant and exclusive; see main paper).

For the recollection estimate under exclusivity (333 trials), there was no linear effect, $BF_{10} = 0.495$, β = 0.318 (95 % CI [-0.323, 0.986]), and anecdotal evidence for a quadratic effect, $BF_{10} = 4.98$, β = 0.724 (95 % CI [0.124, 1.34]). For familiarity scored under the redundancy assumption (447 trials), there was anecdotal evidence against a linear effect, BF_{10} = 0. 327, β = 0.178 (95 % CI [-0.365, 0.737]), but inconclusive evidence for a quadratic effect, $BF_{10} = 1.61$, β = 0.495 (95 % CI [-0.00322, 1.01]).

S7.3. Experiment 2a

Similar to previous experiments, the average expectancy for incorrectly-placed objects M = +37.39 (SD = 15.69) was clearly higher than for correctly-placed objects, M = +1.08 (SD = 14.18), $BF_{10} = 6.63 \times 10^5$, d = 1.7.

For recollection estimated under exclusivity (338 trials), we found no linear effect, BF_{10} = 0.346, PB_{10} = 0.206, β = 0.0175 (95 % CI [-0.419, 0.451]), and very strong evidence for a quadratic effect, BF_{10} = 2.85, PB_{10} = 36.7, β = 0.667 (95 % CI [0.255, 1.08]). For familiarity scored under the redundancy assumption (448 trials), there was no linear effect, BF_{10} = 0.372, PB_{10} = 0.189, β = -0.0524 (95 % CI [-0.432, 0.326]), but moderate evidence for quadratic effect, BF_{10} = 0.979, PB_{10} = 4.6, β = 0.461 (95 % CI [0.111, 0.813]), which can be traced back to the contribution of recollection.

S7.4. Experiment 2b

Again mirroring the results from the previous experiments, the average expectancy for incorrectly-placed objects, M = +23.44 (SD = 22.33), was clearly higher than for correctly-placed objects, M = -9.06 (SD = 25.32), BF_{10} = 1.46 x 10^{10}, *d* = 1.09.

For recollection estimate under exclusivity (955 trials), we found a strong linear effect, BF_{10} = 109, PB_{10} = 4.72, β = -0.387 (95 % CI [-0.677, -0.101]), and moderate evidence for a quadratic effect, BF_{10} = 0.626, PB_{10} = 41.8, β = 0.467 (95 % CI [0.19, 0.744]). For familiarity scored under the redundancy assumption (1302 trials), there was inconclusive evidence regarding a linear effect, BF_{10} = 3.26, PB_{10} = 0.957, β = -0.255 (95 % CI [-0.502, -0.0107]), but moderate evidence for quadratic effect, BF_{10} = 0.438, PB_{10} = 6.08, β = 0.335 (95 % CI [0.0983, 0.569]).

S7.5. Modelling guess responses

Following the suggestion of one of the reviewers, we also modelled “guess” responses. Ideally we would analyse guesses in the same way as we analysed remember and familiar responses, i.e., by estimating their probability conditional on responses of the other category (i.e., under independence or exclusive scoring). However, if we excluded trials with remember or familiar responses, the majority of the remaining trials were guesses (i.e., there were very few “no memory” trials), meaning that estimated conditional probabilities were too close to ceiling to analyse. Instead, we estimated the unconditional probability (i.e., guesses as a proportion of all trials). This showed an inverted U-shape, with most guesses for middling expectancies (see Figure below). Indeed, there was little evidence that guesses increased linearly with expectancy (2420 trials), BF_{10} = 0.69, β = 0.224 (95 % CI [-0.008, 0.455]), but there was moderate evidence for a negative quadratic relationship, BF_{10} = 6.36, β = -0.324 (95 % CI [-0.551, -0.101]). However, this is difficult to interpret because it could simply reflect the dependency of these probabilities on remember responses, which we know from the main analyses showed the opposite, U-shaped function (i.e., more remember responses for the two extremes necessarily means fewer guesses can occur there).

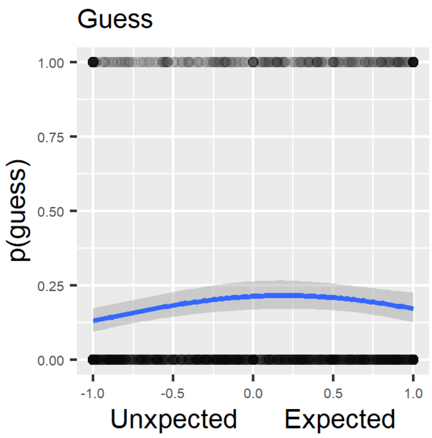

Figure S7.4: The blue line represents the predicted 2nd-order polynomial relationship for the guess model, using evidence propagated across experiments; the shaded area round the blue line represents the 95% credible intervals of the prediction.

S7.6. Continuous (Euclidean) distance error

Given that the definition of the "correct" location during recall was sometimes ambiguous, and to provide a more continuous measure of accuracy (that does not require logistic regression), we also analysed the Euclidean error between the recalled and correct location. According to the original hypothesis, this error should be an inverted-U shape, i.e., with less error towards the two extremes. The results are included in the pooled analyses across experiments shown in Table S7.7 below.

S7.7. Pooled frequentist analysis

To compare with the propagated Bayesian analysis, we performed a final frequentist analysis after pooling data across all four experiments, for recall accuracy, Euclidean error and recognition accuracy (Table S7.7).

Measure		β	SE	z	p	sign
Recall	Intercept	-0.67	0.16	-4.15	< .0001	***
	Linear	0.07	0.1	0.67	.5004	
	Quadratic	1.01	0.26	3.93	.0001	***
Euclidean distance (placement error)	Intercept	0.7	0.04	17.61	< .0001	***
	Linear	0.12	0.03	4.21	< .0001	***
	Quadratic	0.28	0.07	4.24	< .0001	***
3AFC	Intercept	1.01	0.15	6.51	< .0001	***
	Linear	0.18	0.11	1.66	.0977	
	Quadratic	0.85	0.27	3.14	.0017	**
Remember	Intercept	-0.2	0.16	-1.25	.2105	
	Linear	-0.33	0.1	-3.29	.001	*
	Quadratic	1.01	0.25	3.96	.0001	***

Remember (exclusivity)	Intercept	0.61	0.2	3.12	.0018	**
	Linear	-0.32	0.13	-2.4	.0164	*
	Quadratic	1.29	0.33	3.86	.0001	***
Familiarity (independence)	Intercept	-0.57	0.12	-4.91	< .0001	***
	Linear	-0.17	0.13	-1.35	.1767	
	Quadratic	0.61	0.31	1.97	.0488	*
Familiarity (redundancy)	Intercept	1.16	0.15	7.9	< .0001	***
	Linear	-0.22	0.12	-1.8	.0713	
	Quadratic	0.89	0.29	3.03	.0025	**

Table S7.7: Frequentist analysis pooled across all experiments. * p < .05, ** p < .001, *** p < .0001

S7.8. Pooled Bayesian analysis, including additional cubic expansion

To check that the propagation of Bayesian evidence was itself accurate, despite ignoring any posterior dependencies between parameter values, we checked that the BF from the pooled data was similar to the final PB after Experiment 2b. In addition, we added a cubic expansion to ask whether the relationship was S-shaped rather than U-shaped.

For recall, the pooled estimate of the linear effect, β = -0.269 (95 % CI [-0.795, 0.253]) had a BF_{10} = 0.428, which is comparable to PB_{10} = 0.105 reported for Experiment 2b above; while the pooled estimate of the quadratic effect, β = 0.408 (95 % CI [0.205, 0.612]), had a BF_{10} = 471, again is comparable to the above PB_{10} = 387. Importantly, there was no evidence in favour of a cubic effect, β = 0.347 (95 % CI [-0.17, 0.87]), BF_{10} = 0.586.

For 3AFC, the pooled estimate of the linear effect, β = -0.0495 (95 % CI [-0.606, 0.503]), had a BF_{10} = 0.276, comparable to the above PB_{10} = 0.41, while the pooled estimate of the quadratic effect, β = 0.342 (95 % CI [0.131, 0.556]), had a BF_{10} = 30, comparable to the above PB_{10} = 13.2. Again there was no evidence in favour of a cubic effect, β = 0.238 (95 % CI [-0.314, 0.792]), BF_{10} = 0.388.

Finally, for recollection, the pooled estimate of the linear effect, β = -0.923 (95 % CI [-1.46, -0.393]), had a BF_{10} = 81, comparable to the above PB_{10} = 27, while the pooled estimate of the quadratic effect, β = 0.414 (95 % CI [0.214, 0.615]), had a BF_{10} = 305, comparable to the above PB_{10} = 135. In contrast to 3AFC and recall, there was moderate evidence in favour of a cubic effect, β = 0.634 (95 % CI [0.108, 1.17]), BF_{10} = 4.22; however this was much smaller than the extreme evidence for a quadratic effect.

Overall, while there were small differences due to numerical sampling and due to ignoring the covariance between parameters (see main text), pooling the Bayes Factors across experiments (PB) gave similar results to calculating the Bayes Factors on pooled data.

S8. Interrupted regressions

The idea behind interrupted regression is that evidence for a quadratic component alone does not imply a U-shape. To circumvent this, two lines are fit to the data that intersect at a certain point, and a U-shape is inferred if both slopes are different enough from zero and have opposite signs. As pre-registered (see https://osf.io/b9dqg), we tested this by combining data across all experiments. To determine the "breaking-point" (bp) where the two lines intersect, we fit the model:

y ~ xlow + xhigh + high + set + (1 | participant number) + (1 | object number)

where, if x = object-location expectancy, then:

xlow = x – bp if x <= bp and 0 otherwise,

xhigh = x – bp if x > bp and 0 otherwise, and

high = 1 if x > bp and 0 otherwise (to capture any difference in intercept of the two lines)

We then tested whether the 95% CIs for both the "leftward" (xlow) slope and "rightward" (xhigh) slopes included zero. To optimise "bp", we tested 10 equally-spaced breaking points within the middle 80% range of the expectancy ratings. Simulations showed that the false positive rate remained under 5% for testing these 10 breaking points when accepting a $BF_{10} > 6$ as evidence (see https://jaquent.github.io/post/finding-a-u-shape-with-bayesian-interrupted-regression/). Results are shown in Tables S8.1-S9.3. below. Note that the breaking point position refers to the scaled expectancy values (i.e., after scaling to have a SD of 0.5 for the Bayesian statistics).

S8.1. Recall

Breaking point #	Breaking point position	β_1	2.5%	97.5%	BF_{10}	BF_{10} (1-tailed)	β_2	2.5%	97.5%	BF_{10}	BF_{10} (1-tailed)
1	-0.734	-0.0117	-2.36	2.29	0.994	0.985	0.066 8	-0.154	0.291	0.132	0.188
2	-0.576	0.268	-1.63	2.3	0.931	0.736	0.263	-0.0122	0.536	0.796	1.54
3	-0.417	-0.0109	-1.3	1.28	0.628	0.639	0.521	0.187	0.856	17.1	34.2
4	-0.258	-0.849	-1.8	0.0812	2.41	4.64	0.529	0.147	0.917	7.06	14.1
5	-0.0995	-0.613	-1.23	0.00584	2.08	4.05	0.725	0.23	1.22	13.6	27.1
6	0.0592	-0.778	-1.3	-0.284	27.5	54.9	0.687	0.0642	1.34	3.18	6.27
7	0.218	-0.383	-0.776	0.0137	1.16	2.25	1.35	0.462	2.27	35.9	71.7
8	0.377	-0.35	-0.667	-0.0316	1.71	3.37	1.78	0.364	3.27	16.2	32.2
9	0.535	-0.307	-0.581	-0.0372	1.67	3.3	1.15	-0.724	3.45	1.65	2.89

10	0.694	-0.189	-0.413	0.0344	0.443	0.845	0.015 3	-2.31	2.35	1.05	1.05

Table S8.1: A U-shape was found for breaking point 6 for recall as the CI and the BFs indicated that both slopes have opposite signs and are different from zero.

S8.2. 3AFC

Breaking point #	Breaking point position	β_1	2.5%	97.5%	BF_{10}	BF_{10} (1-tailed)	β_2	2.5%	97.5%	BF_{10}	BF_{10} (1-tailed)
1	-0.734	0.0077	-2.31	2.3	1.05	1.04	0.272	0.0303	0.514	1.42	2.81
2	-0.575	-0.629	-2.92	1.34	1.05	1.52	0.397	0.117	0.678	5.88	11.7
3	-0.417	-0.537	-1.91	0.803	0.863	1.34	0.549	0.212	0.893	29.2	58.4
4	-0.258	-0.835	-1.77	0.116	2.06	3.93	0.505	0.105	0.918	4.28	8.48
5	-0.0995	-0.513	-1.16	0.137	1.08	2.02	0.554	0.0277	1.08	2.38	4.66
6	0.059	-0.313	-0.82	0.21	0.491	0.868	0.752	0.111	1.42	4.36	8.63
7	0.218	-0.006	-0.42	0.406	0.202	0.206	1.64	0.686	2.61	117	233
8	0.376	-0.126	-0.46	0.211	0.217	0.334	2.24	0.706	3.95	44.3	88.4

9	0.535	-0.191	-0.47	0.0887	0.307	0.556	1.1	-0.777	3.57	1.5	2.57
10	0.693	-0.0904	-0.33	0.144	0.156	0.243	0.0618	-2.28	2.5	0.991	1.04

Table S8.2: No U-shape was detected for 3AFC as the CI and the BFs (order restricted) indicated that there is enough evidence that both slopes have opposite signs and are different from zero.

S8.3. Recollection

Breaking point #	Breaking point position	β_1	2.5%	97.5%	BF_{10}	BF_{10} (1-tailed)	β_2	2.5%	97.5%	BF_{10}	BF_{10} (1-tailed)
1	-0.732	0.0129	-2.3	2.38	1	0.994	-0.345	-0.545	-0.147	20.4	0.0136
2	-0.573	0.0084	-1.8	1.88	0.871	0.859	-0.152	-0.419	0.115	0.237	0.0643
3	-0.415	-0.088	-1.41	1.19	0.643	0.71	0.0757	-0.232	0.39	0.172	0.236
4	-0.257	-0.771	-1.7	0.168	1.75	3.33	0.127	-0.242	0.49	0.229	0.347
5	-0.0989	-0.725	-1.35	-0.108	4.25	8.4	0.481	0.0024	0.949	1.67	3.26
6	0.0593	-0.854	-1.36	-0.348	46	92	0.843	0.213	1.48	12.2	24.4
7	0.217	-0.841	-1.23	-0.454	310	620	1.22	0.37	2.1	20.3	40.4
8	0.376	-0.816	-1.15	-0.495	4340	8690	1.67	0.325	3.15	12.9	25.6
9	0.534	-0.722	-0.99	-0.451	3570	7140	1.98	-0.132	4.77	4.39	8.45

10	0.692	-0.616	-0.84	-0.389	3410	6810	0.137	-2.22	2.6	0.972	1.07

Table S8.3: A U-shape was found for breaking points 5 - 9 in recollection as the CI and the BFs indicated that both slopes have opposite signs and are different from zero.

S9. References

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13). https://doi.org/10.18637/jss.v082.i13