Supplementary Materials for *Order is ordinal in serial memory*

Simon D. Lilburn and Gordon D. Logan

1 Details of data analysis

Each of the analyses in the paper was based on a Bayesian mixed-effects model. We ran each test separately for mean response time (using a standard linear model) and for accuracy (using a beta–binomial GLM predicting the number of correct trials given the total number of trials). In each model, the participant was treated as a random effect on the intercept.

Uniform spacing experiments The mixed-effects models for Experiments 1 and 3 estimated a main effect on performance (mean RT or accuracy) based on whether the first (memory) array had a "wide" spacing, whether the two arrays were incompatible, or whether the comparison was perceptual. The data were first split into cells based on these conditions.

The model for these experiments was:

$$Y_i \sim \mathscr{F}(\mu_i, \theta) \tag{1}$$

$$\mu_i = \beta_0 + \beta_1 W_i + \beta_2 I_i + \beta_3 P_i + u_{[participant]} \tag{2}$$

Here Y_i is the response variable (either accuracy or response time) for the cell i; $\mathscr{F}$ is the distribution function (Gaussian for mean RT, beta–binomial for accuracy); W_i is an indicator for whether the first (stimulus) array spacing was wide; I_i is an indicator for whether the stimulus and probe spacing were incompatible; P_i is an indicator for whether the presentation was simultaneous (requiring perceptual comparison); and $u_{[participant]_i}$ is the random intercept for the participant.

Uniform spacing experiments—Distance effect For Experiments 1 and 3, we also examined the metric influence on the distance effect. To examine the distance effect, we created an additional model which includes a binary predictor for transposition distance:

$$\mu_i = \beta_0 + \beta_1 N_i + \beta_2 W_i + \beta_3 I_i + \beta_4 P_i + u_{[participant]} \tag{3}$$

This has the same structure as the model for correspondence effect above, with addition of an indicator variable, N_i, indicating whether items were spatially near (absolute lag = 1).

Random spacing experiments Experiments 2 and 4 collapsed over different specific stimulus and probe spacing patterns. We estimated the main effects on performance for whether the arrays were unaligned and whether the comparison was perceptual.

The model for these experiments was:

$$Y_i \sim \mathscr{F}(\mu_i, \theta) \tag{4}$$

$$\mu_i = \beta_0 + \beta_1 U_i + \beta_2 P_i + u_{[participant]} \tag{5}$$

Where Y_i is the response variable (either accuracy or response time) for cell i; $\mathscr{F}$ is the distribution function (Gaussian for mean RT, beta–binomial for accuracy); U_i is an indicator for whether the alignment was unaligned; P_i is an indicator for whether the presentation was perceptual; and $u_{[participant]_i}$ is the random intercept for the participant.

Inference configuration For each model, we ran ten independent chains for 30,000 iterations each after discarding 10,000 iterations. All parameter estimates converged with $\hat{R}$ ¡ 1.01; no chains had divergent transitions after warm-up.

Priors For all models, we specified weakly informative priors. For models predicting mean response time, we specified a $\mathscr{N}(0, 500 \text{ ms})$ prior on fixed effects. For models predicting accuracy, we specified a $\mathscr{N}(0, 5)$ prior on fixed effects.

Planned contrasts and model evaluation To evaluate key hypotheses, we conducted planned contrasts using the `emmeans` package. For each analysis, we extracted Maximum A Posteriori (MAP) estimates and 95% highest density intervals (HDI).

Contrast weights Tables 1–3 present the contrast weights used for each analysis. The weights shown are prior to normalization. During analysis, we normalized these weights so that positive weights summed to +1 and negative weights summed to -1.

Uniform spacing experiments contrasts For Experiments 1 and 3 (uniform spacing manipulation), we assessed a series of planned contrasts targeting our theoretical questions. We evaluated the effect of wide versus narrow spacing in memory by comparing performance across these conditions. We tested the memory compatibility effect by comparing compatible versus incompatible arrays in memory trials. Similarly, we examined the effect of compatibility in perceptual trials. Finally, we assessed the overall difference between memory and perceptual modalities by comparing performance collapsed across all spacing and compatibility conditions.

Table 1: Contrast weights for uniform spacing experiments (Experiments 1 and 3)

Contrast	Cell							
	Memory				Perceptual			
	Compatible		Incompatible		Compatible		Incompatible	
	Narrow	Wide	Narrow	Wide	Narrow	Wide	Narrow	Wide
Memory correspondence	-1	-1	1	1	0	0	0	0
Perceptual correspondence	0	0	0	0	-1	-1	1	1
Modality comparison	-1	-1	-1	-1	1	1	1	1

Distance effects contrasts For distance effect analyses, we evaluated contrasts that allowed us to isolate the influence of spatial proximity on performance. We examined the memory-specific distance effect by comparing near versus far conditions in memory trials only. Likewise, we assessed the perceptual-specific distance effect by comparing near versus far conditions in perceptual trials only. To test whether spacing width modulated distance effects, we examined the memory-specific metric distance interaction between spacing width and proximity in memory trials. We similarly tested the perceptual-specific metric distance interaction between spacing width and proximity in perceptual trials.

Table 2: Contrast weights for distance effects in uniform spacing experiments

Contrast	Cell							
	Memory				Perceptual			
	Narrow spacing		Wide spacing		Narrow spacing		Wide spacing	
	Near	Far	Near	Far	Near	Far	Near	Far
Memory distance	-1	1	-1	1	0	0	0	0
Perceptual distance	0	0	0	0	-1	1	-1	1
Memory metric distance	-1	1	1	-1	0	0	0	0
Perceptual metric distance	0	0	0	0	-1	1	1	-1

Random spacing experiments contrasts For Experiments 2 and 4 (the random spacing manipulation), we tested several key contrasts to evaluate our hypotheses. We assessed the memory correspondence effect by comparing aligned versus unaligned conditions in memory trials only. We similarly evaluated the perceptual correspondence effect by comparing aligned versus unaligned conditions in perceptual trials. We also examined the overall difference between memory and perceptual modalities by comparing performance across all memory trials against all perceptual trials, regardless of alignment.

Table 3: Contrast weights for random spacing experiments (Experiments 2 and 4)

Contrast	Cell			
	Memory		Perceptual	
	Same spacing	Different spacing	Same spacing	Different spacing
Memory correspondence	1	-1	0	0
Perceptual correspondence	0	0	1	-1
Modality	-1	-1	1	1

Results reporting For each contrast of interest, we reported the maximum a posteriori (MAP) estimate of the mean difference (MD) and the 95% highest density interval (HDI) around that mean difference. When relevant, we also reported empirical means for conditions with positive and negative contrast weights.

Effects were considered statistically supported when the 95% HDI did not include zero.

2 Replication of Experiment 1

To test the distance function with an equal number of transposition distances, we needed to fix either the number of times each position in the memory list was probed or the number of times each position in the probe list was cued. In the article, we reported experiments where each position in the memory list was probed equally often. Here we report the experiment where each position in the probe list was cued equally often.

2.1 Method

2.1.1 Participants

Thirty-two participants were recruited from Prolific (`https://prolific.co`). Participants were filtered by Prolific to be between 18 and 40 years of age, located within the United States of America, with English as their first language, and with over a 95% approval rate for previous studies.The experimental session lasted an hour for which participants were paid USD $12.

2.1.2 Stimuli, apparatus, and procedure

The setup of this replication was the same as the Experiment 1 reported in the article. Other than the difference in the subjects recruited, the design only changed to fix the number of times each position in the probe array was cued to be equal.

2.2 Results

Prior to analysis, we discarded any trials cancelled where the participant took longer than 4,000 ms to respond. The total number of trials was reduced from 20,480 to 20,383 (0.5% of the dataset).

The analysis approach was the same as that used for the original Experiment 1, with one exception. Due to very high accuracy rates in some conditions, we used binomial models rather than beta-binomial models for the accuracy data to avoid convergence issues. The mean RT data were analyzed using the same Bayesian mixed-effects models as in the main experiments.

Figure 1 shows the results for the memory trials (left-hand panels) and perceptual trials (bottom panels). Figure 1A and C displays mean RTs and error rate for memory trials with the same spacing and those with different spacing. Changing the spacing on memory trials had little to no effect on RTs, *MD* = -1.24 ms, 95% HDI = [-30.56 ms, 28.22 ms], or error rate, *MD* = -0.11%, 95% HDI = [-1.29%, 1.06%].

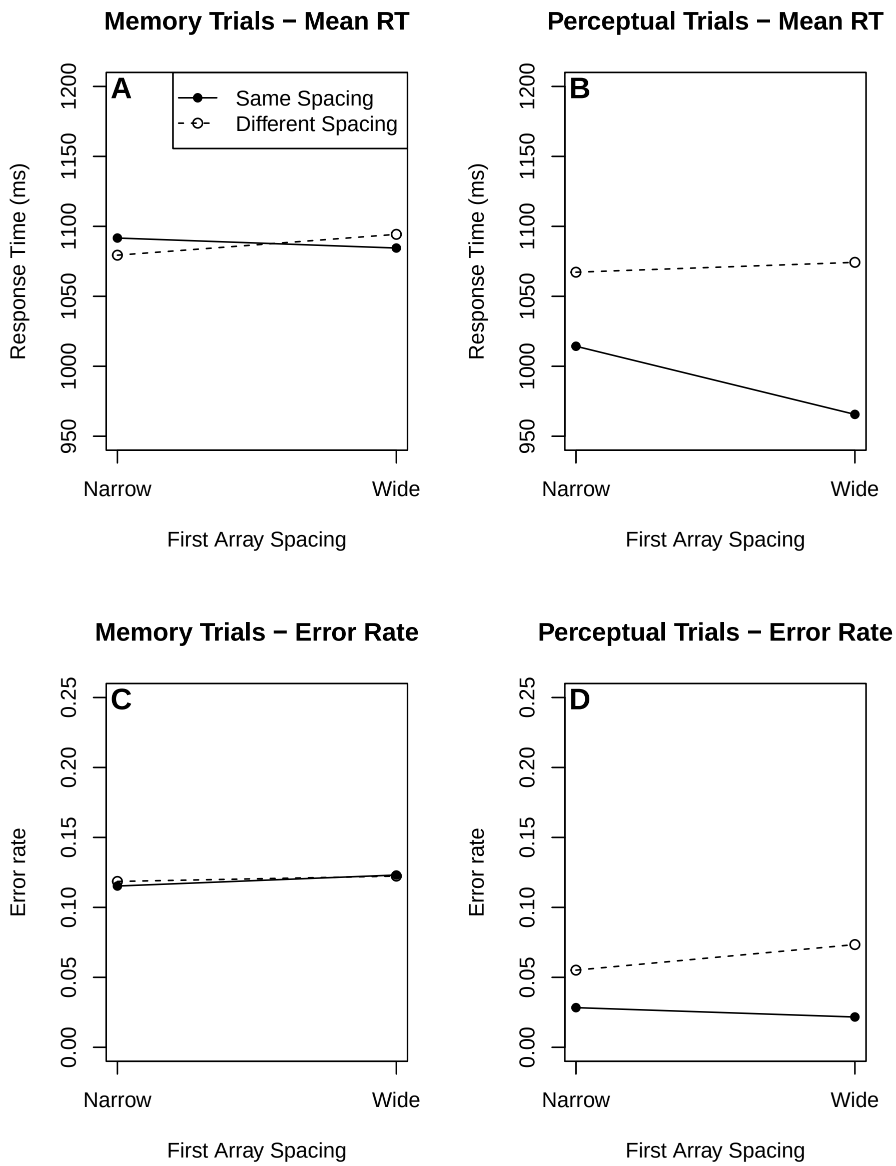

Figure 1: Main effects of metric correspondence in the replication of Experiment 1. Panels **A** and **B** give mean response times for memory and perceptual matching, respectively, plotted against the first-array spacing (narrow vs. wide). Filled circles/solid lines represent *same spacing* trials; open circles/dashed lines represent *different spacing* trials. Panels **C** and **D** show the corresponding error rates.

By contrast, Figure 1B and D shows that changing the spacing on perceptual trials had strong effects on both RTs and error rate: Perceptual matching RTs were slower, *MD* = 80.80 ms, 95% HDI = [50.80 ms, 109.98 ms], and less accurate, *MD* = -3.39%, 95% HDI = [-4.47%, -2.48%], when the spacing between the stimulus array and probe array did not match.

Figure 2 shows the distance effect in both memory trials (panel A and C) and perceptual trials (panel B and D). We found a substantial distance effect overall (*MD* = 58.59 ms, 95% HDI = [41.05 ms, 76.38 ms] for RT; *MD* = -4.45%, 95% HDI = [-5.85%, -3.27%] for accuracy). The memory-specific distance effect was robust, with near memory trials being substantially slower (*MD* = 73.59 ms, 95% HDI = [48.40 ms, 98.65 ms]) and less accurate (*MD* = -6.07%, 95% HDI = [-8.18%, -4.32%]) than far memory trials. Similarly, perceptual trials showed a distance effect in both RT (*MD* = 43.60 ms, 95% HDI = [19.15 ms, 68.93 ms]) and accuracy (*MD* = -2.75%, 95% HDI = [-3.99%, -1.75%]).

The metric influence on the distance effect was assessed by comparing trials with narrow spacing and those with wide spacing. We found no evidence for a metric influence on the distance effect in either memory trials (*MD* = 8.72 ms, 95% HDI = [-26.38 ms, 43.74 ms] for RT; *MD* = -0.24%, 95% HDI = [-2.38%, 1.79%] for accuracy) or perceptual trials (*MD* = -24.95 ms, 95% HDI = [-60.15 ms, 10.11 ms] for RT; *MD* = 2.22%, 95% HDI = [0.72%, 3.95%] for accuracy), as the 95% HDI included zero in all cases except for a slight effect in perceptual trials for accuracy.

These results replicate the findings from Experiment 1. The lack of a correspondence effect in memory matching supports the conclusion that items in memory are accessed and retrieved without using metric relationships. The presence of a correspondence effect in perceptual matching demonstrates that perceptual representations of order are highly dependent on the metric properties of the arrays being compared. While we observed distance effects in both modalities, the metric influence on these distance effects was not substantial, further supporting an ordinal view of memory.

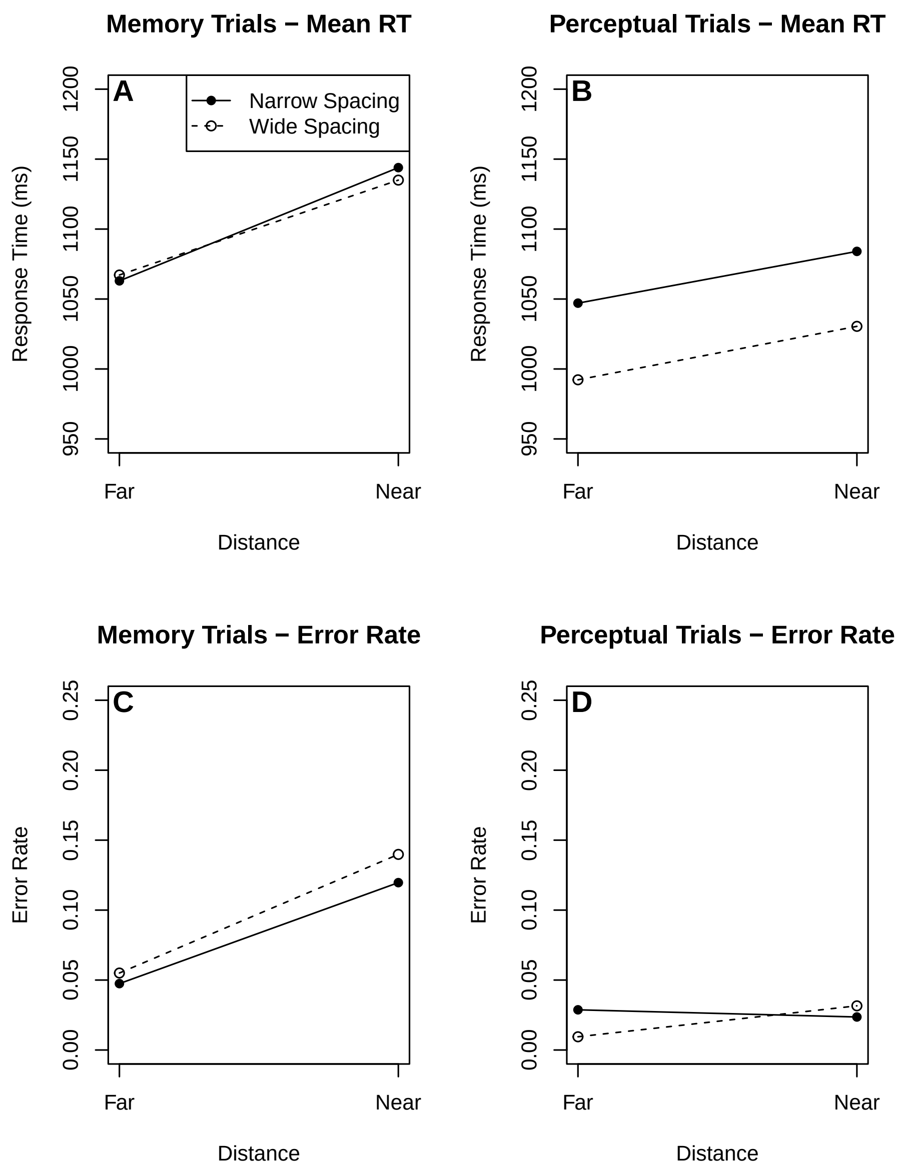

Figure 2: Distance effects in the Experiment 1 replication. Panels **A** (memory) and **B** (perceptual) plot mean response times for *near* vs. *far* probe positions (absolute transposition distance of 1 vs. absolute transposition distance of 2) when the first array was either narrowly or widely spaced. Panels **C** and **D** present the corresponding error rates.