Unpublished experiment from "Working memory for cross-domain sequences"

Simon Farrell

University of Bristol

Klaus Oberauer

University of Zurich

Author Note

Simon Farrell, School of Experimental Psychology, University of Bristol; Klaus Oberauer, Department of Psychology, University of Zurich.

This research was supported by ESRC grant RES-000-22-1665 awarded to Farrell and Oberauer. We thank Fiona Laver for her assistance in running the experiments.

Correspondence should be addressed to Simon Farrell, School of Experimental Psychology, University of Bristol, 12a Priory Road, Clifton, Bristol, BS8 1TU, UK. Email: Simon.Farrell@bristol.ac.uk. Web: http://eis.bris.ac.uk/~pssaf/

Abstract

This unpublished manuscript presents an initial version of the experiment in Farrell & Oberauer’s paper "Working memory for cross-domain sequences"

Unpublished experiment from "Working memory for cross-domain sequences"

Experiment

This experiment examined serial short-term memory for lists composed of stimuli drawn from three domain: words, visual doodles, and spatial matrices (assumed to tap spatial memory). Lists were either composed of six exemplars from a single class, or two exemplars from each of the three classes. Lists were recalled using a serial reconstruction procedure that allowed positional confusions and extra-list intrusions.

Method

Participants and Design. Twenty members of the campus community at the University of Bristol participated in exchange for course credit or reimbursement. All of them participated in all cells of the stimulus class (words, spatial, doodles) × list type (pure vs mixed) design.

Materials. Three sets of stimuli were constructed to serve as verbal, visual and spatial memoranda.

The verbal set consisted of 40 single-syllable words selected from the phonologically dissimilar set of Coltheart (1993).

The visual set consisted of 40 filled complex filled shapes (or ‘doodles’), an example of which is shown in Figure 1. To minimize the likelihood of participants naming the stimuli, the 40 stimuli were selected from a larger set of 100 shapes constructed by the experimenter, using the freehand tool in a graphics software program. The larger set of 100 pictures was printed in a booklet, and 10 pilot participants were asked to attempt to provide a verbal label to each picture in the booklet based on its visual appearance; if no label came to mind, the participants were to leave a blank. Based on these participants’ responses, the 40 least nameable doodles were selected; 13 of these were not named by any participant, a further 14 were named by only a single participant, and the remaining 13 were named by two participants, but with different labels for each participant.

The spatial set consisted of 20 spatial matrices often used in tasks such as the spatial updating task (Oberauer & Bialkova, 2011); an example is shown in Figure 1. Each matrix consisted of a 5 × 5 array of squares, with a filled circle appearing in one of the squares. From the 25 possible configurations stimuli with a dot in the corner or in the centre were excluded, as these were expected to be more amenable to naming; this left the 20 stimuli used in the experiment.

The stimuli were stored in PNG format (greyspace) at a size of 180 × 180 pixels.

Procedure. A Windows-based computer running the Psychophysics Toolbox for MATLAB (Brainard, 1997; Pelli, 1997) displayed all stimuli and recorded all responses. The display background was white, with stimuli being presented in black. The display was a 17” CRT monitor set to a display resolution of 1280 × 1024.

Each trial began with the presentation of a fixation cross in the middle of the screen for 800 ms, followed by a blank screen for approximately 200 ms. Following this, a sequence of six stimuli was displayed in the middle of the screen. On pure lists the six stimuli were all drawn from the same class, while on mixed list trials each list consisted of two stimuli drawn from each class. The stimuli were presented one by one in the middle of the screen, each stimulus appearing for 1 s, and successive stimuli being separated by a blank screen of 500 ms.

Following presentation of six stimuli, a reconstruction screen appeared. This screen contained the list items and 3 lures (items that were not presented on the list). On pure lists the lures were three further items from the same stimulus class as the memoranda; on mixed lists, exactly one lure came from each of the three stimulus classes. The list stimuli and lures were displayed in a 3 × 3 grid in a psuedo-random order, and participants’ task was to click on the list items in the order of their original presentation. Clicks that did not land within the 180 × 180 pixel square corresponding to a stimulus were ignored. Once an item had been clicked, it was removed from the display to indicate that it had already been recalled (these squares could still be clicked, but this happened very rarely in the experiment). When six stimuli had been selected, a blank screen was displayed for 500 ms. Following this, participants were given feedback on their performance in a message displaying number correct out of 6; this was displayed for 1 second before moving to the next trial.

Across the mixed trials, each possible ordering of stimulus classes was used, resulting in 90 trials for that condition. Ninety pure lists were also presented (30 for each stimulus class) to give 180 trials in total. For each trial stimuli were pseudo-randomly selected from the stimulus set and allocated to list positions according to the stimulus structure for that trial. The order of trials was psuedo-randomized before beginning the experiment. Each participant completed two sessions of approximately 45 minutes, with 90 trials completed per session. Participants were given self-paced breaks every 15 trials; these breaks were terminated by pressing the space bar. Six demonstration trials initiated the first session; the first three of these were pure lists (words, doodles, and spatial matrices, respectively), and the second three were mixed lists.

Results

The top-left panel of Figure 2 shows the effects of list structure and stimulus class on mean recall accuracy (scored using a correct-in-position criterion). A 2 (list structure: pure vs mixed) × 3 (stimulus class: words, spatial matrices, and doodles) repeated-measures ANOVA was conducted on overall accuracy . The ANOVA revealed a significant main effect of stimulus class, $F(2,38) = 56.48, p < .001, \eta_p^2 = .75$, and a significant interaction, $F(2,38) = 17.73, p < .001, \eta_p^2 = .48$; the main effect of list structure was not significant, $F(1,19) < 1$. Post-hoc t-tests revealed a significant disruptive effect of mixed lists on memory for words, $t(19) = 4.25, p < .001$, and a significant enhancement for doodles, $t(19) = 2.35, p = .003$; the effect for the spatial matrices was marginal under Bonferroni correction, $t(19) = 3.37, p = .03$.

The remaining analyses broke the overall accuracy patterns down by analyzing different types of possible errors.[1] The top-right panel of Figure 2 shows proportion of responses that were extra-list intrusions (i.e., selection of lures) in the different conditions. An ANOVA

[1] Note that the ANOVAs reported here are technically redundant, since the different error types must necessarily add up to 1-accuracy. This is not a major issue here, as we are using the ANOVAs on errors to determine what is driving any differences in response accuracies.

revealed significant effects of stimulus class, $F(2,38) = 61.60$, $p < .001$, $\eta_p^2 = .76$, and list structure, $F(1,19) = 4.73$, $p = .042$, $\eta_p^2 = .20$, qualified by a significant interaction, $F(2,38) = 12.50$, $p < .001$, $\eta_p^2 = .40$. Post-hoc tests revealed a significant increase in intrusions between pure and mixed lists for words, $t(19) = 3.75$, $p = .001$, and a significant decrease in intrusions for spatial matrices, $t(19) = 3.02$, $p = .007$, and doodles, $t(19) = 3.32$, $p = .004$.

The bottom-left panel of Figure 2 plots the effects of the manipulated variables on the mean proportion of responses that were confusions between list items from the same class. In the case of pure lists (where all items were from the same class), this mean was calculated in a manner reflecting the constraints imposed in mixed lists. Specifically, for each pure list trial, all possible pairs of items were constructed, with each pair corresponding to a possible pair in the mixed list condition. A score for that trial was then calculated by averaging across the bootstrapped scores. An ANOVA revealed a significant effect of stimulus class, $F(2,38) = 16.14$, $p < .001$, $\eta_p^2 = .46$, and a significant main effect of list structure, with mixed lists producing more confusions between items from the same class, $F(1,19) = 25.44$, $p < .001$, $\eta_p^2 = .57$. The interaction was not significant, $F(2,38) < 1$.

Finally, the bottom-right panel of Figure 2 plots the mean proportion of responses that were confusions between list items from different stimulus classes. In the case of pure lists such errors are not possible, such that the plotted means show control performance calculated using a method similar to that for examining within-class confusions outlined above. Specifically, for each pure list trial each possible mixed list structure was imposed on the list, and any confusions between items that were classed as between-class intrusions according to the mixed list template were scored as such. An ANOVA on the factorial means revealed a significant effect of stimulus class, $F(2,38) = 40.83$, $p < .001$, $\eta_p^2 = .68$, and list structure, $F(1,19) = 8.44$, $p = .009$, $\eta_p^2 = .31$, the latter qualified by a significant interaction, $F(2,38) = 20.34$, $p < .001$, $\eta_p^2 = .52$. Post-hoc tests revealed a significant increase in between-class confusions between pure and mixed lists for words, $t(19) = 3.02$, $p = .007$, and a significant decrease in such confusions for spatial matrices, $t(19) = 3.86$, $p = .001$, and doodles,

$t(19) = 4.47$, $p < .001$.

Discussion

The results of Experiment 1 showed some heterogeneous effects of mixing stimulus classes on recall performance. Mixing words with other items reduced accuracy, and increased all types of errors (intrusions, within-class confusions, and between-class confusions) for words. For spatial matrices and doodles, intrusions and between-class confusions were reduced on mixed lists, with a specific increase in within-class confusions.

The results of Experiment 1 do not relate straightforwardly to any of the predictions derived in the Introduction, as mixing stimulus classes did not have a systematic effect across the individual classes. One potential explanation for the differences between words and the visual and spatial stimuli is that participants were strategically focussing on remembering the spatial and visual stimuli because the words were relatively easy to remember. To account for this possibility, a second experiment was conducted in which non-words were used as verbal stimuli, to bring recall performance for that stimulus class down.

In a further unreported simulation, we also confirmed that a reverse pattern in words—where words on mixed lists are remembered less well than on pure lists—can be obtained from SEM. This was accomplished by lowering the parameter scaling the confusability of words (λ) from 0.6 to 0.2 (to reflect the overall superior recall of words, as shown in Figure 2), and multiplying the encoding rate for words by a factor of 0.93 (i.e., Equation B3 in Henson, 1998, was multiplied by a value of 0.93). Accordingly, the full set of results is entirely consistent with a global model, with inconsistencies between stimulus classes being attributable to differences in confusability, and the strategic response of participants to task difficulty by reducing the attention paid to words.

—

References

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436.

Coltheart, V. (1993). Effects of phonological similarity and concurrent irrelevant articulation on short-term memory recall of repeated and novel word lists. *Memory & Cognition, 21,* 539–545.

Henson, R. N. A. (1998). Short-term memory for serial order: The Start-End Model. *Cognitive Psychology, 36,* 73–137.

Oberauer, K. & Bialkova, S. (2011). Serial and parallel processes in working memory after practice. *Journal of Experimental Psychology: Human Perception and Performance, 37,* 606–614.

Pelli, D. G. (1997). The Video Toolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442.

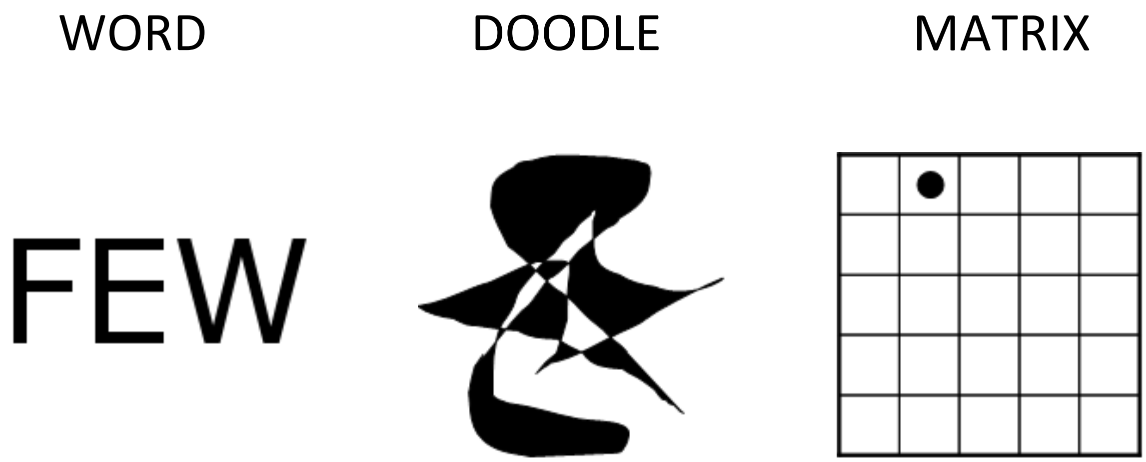

Figure 1. Example stimuli from unpublished experiment. The stimulus classes, from left to right, are verbal (words), visual (doodles), and spatial (spatial matrices).

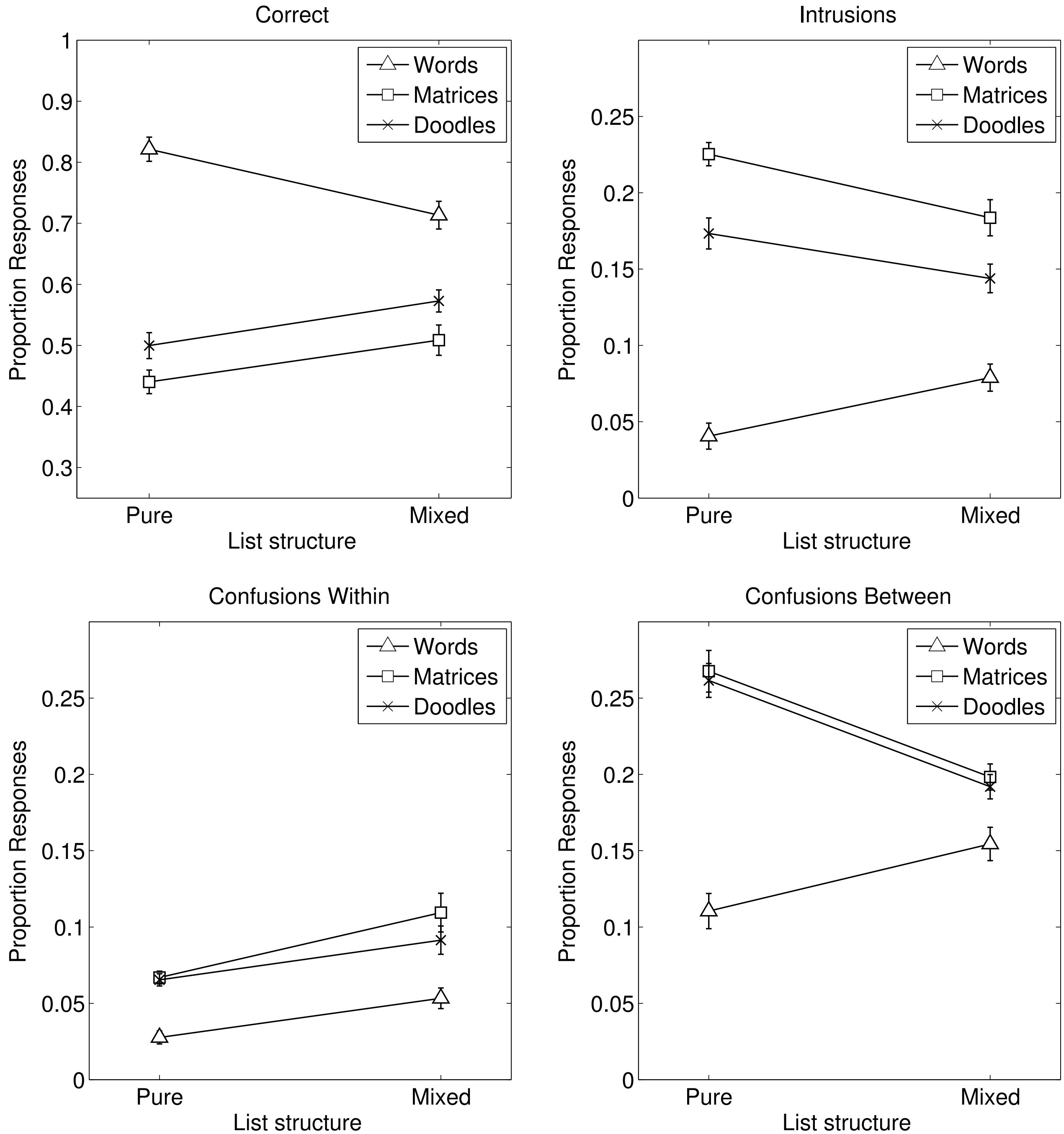

Figure 2. Effect of list structure (pure vs mixed) and stimulus class on proportions of different response types in Experiment 1. Top left: Proportion correct responses; Top right: Proportion Intrusions; Bottom left: proportion within-class confusions; Bottom right: proportion between-class confusions.