The Production Effect Becomes Spatial

Jean Saint-Aubin

Université de Moncton

Marie Poirier

City, University of London

James M. Yearsley

City, University of London

Dominic Guitard

Cardiff University

Authors' Note

This research was supported by a discovery grant from the Natural Sciences and Engineering Research Council of Canada to Jean Saint-Aubin. Correspondence concerning this article should be addressed to Jean Saint-Aubin, École de psychologie, Université de Moncton, Moncton, New Brunswick, E1A 3E9, Canada.

Contact: Jean.Saint-Aubin@umoncton.ca

Abstract

In the verbal domain, it is well established that words read aloud are better remembered than their silently read counterparts. It has been hypothesised that this production effect stems from the addition of distinctive features, with the caveat that the processing that generates added features interferes with rehearsal. Here, we tested the idea that a similar trade-off is found in the visuo-spatial domain. In all experiments, a short series of single dots sequentially appeared at various locations on a screen. Participants either produced the items by clicking on them at presentation, watched the items appear quietly, or produced an irrelevant click after each item to better even out rehearsal opportunities between produced and control conditions. In Experiment 1, the dots appeared within a visible grid and an order reconstruction task was used. Experiment 2 also called upon reconstruction, but with the grid removed. In Experiments 3, a recall task was used. Results show that producing items hindered performance compared to the control condition. Conversely, production improved performance compared to the control condition where rehearsal was hindered. This is the first demonstration of a visuo-spatial production effect. The key findings were successfully modeled by the Revised Feature Model (RFM).

The Production Effect Becomes Spatial

Imagine yourself teaching in a lecture theater. During a brief pause, a student asks a question. You think of a useful answer and turn away from the group to write something on the board, to support your explanation. Turning back toward the room to finish your answer, you look at the student who asked the question, and then return to the thread of your lecture. This example nicely illustrates the complexity and flexibility of our cognitive functioning. Among other things, it exemplifies the interaction between prior knowledge and current thought processes. This is exemplified in background knowledge about the lecture topic, memory for the general plan of what is to be covered, and the use of appropriate language and vocabulary, etc. Moreover, the lecturer must maintain recent verbal and conceptual content (the question of the student, where the lecture was interrupted) while quickly compiling a strategy to answer said question and remembering the steps needed to provide the answer. Furthermore, spatial information must be at least temporarily remembered, including the location of the student in the lecture theater, for instance.

Prior knowledge and experience are obviously critical in thinking and action– as is short-term or primary memory. The latter is thought to maintain immediate aims, currently relevant information, etc. —in short, the materials of immediate planning, thought, and action (Baddeley, 1986; Barrouillet & Camos, 2015; Cowan, 1999; Engle, 2018; Oberauer, 2009). Moreover, research on primary memory has often included an important role for some form of rehearsal (e.g., Atkinson & Shiffrin, 1968, Baddeley, 1986; Camos, 2015; Murray, 1967; but see, Souza & Oberauer, 2018), a process that is embedded in the predictions tested here.

We call upon a view of primary memory where, distinctiveness, interference, long-term memory and rehearsal are all important. This view’s main ideas are embodied in a computational model known as the Revised Feature Model (RFM; Saint-Aubin et al., 2021, 2023). The RFM owes a lot to the proposals of Nairne (1988, 1990), of Nosofsky (2011) and the work of others on rehearsal (Bhatarah et al., 2009; Grenfell-Essam et al., 2013; Murray, 1967).

Here, we tested a series of predictions derived from the RFM; importantly said tests involve stimuli and tasks within the visuo-spatial domain, as opposed to the typical verbal items. According to the RFM, basic encoding and retrieval processes do not change with the specific characteristics of the processed material (Poirier et al., 2019). In other words, many processes would be invariant across domains, even though item features in different domains could vary considerably and rely on different brain areas for their development (e.g., visual processing areas, language-specific processing; see Poirier et al., 2019). The current paper is a straightforward test of RFM predictions regarding a visuo-spatial task, including hypotheses about how visuo-spatial rehearsal interacts in predictable ways with encoding operations.

We assumed that the features that are encoded, processed, and retrieved in visuo-spatial tasks are domain-specific. We also assumed that rehearsal for visuo-spatial materials recruits different systems than does rehearsal of verbal materials, although obeying some of the same rules. We based the latter view on prior research indicating that visual control mechanisms, called upon in identifying locations, are important in visuo-spatial rehearsal (e.g., Awh & Jonides, 2001; Tremblay et al., 2006) whereas in the verbal case, it is thought that the mechanisms controlling actual speech output are at least partially involved in subvocal rehearsal (e.g., Page & Norris, 2009).

In the RFM, encoding of episodic and primary memory information relies on building unique feature combinations – i.e., each event or item is represented by a series of ordered features whose arrangement is largely unique. As an analogy think of spoken words – they are composed by a small number of phonemes (44 in English); yet in combination, can produce large vocabularies of unique utterances. In the RFM, the features that represent events are the product of current processing. The latter include modality-dependent information generated by relevant perceptual information processing, e.g. the quality of someone voice, whether something appears on a screen or is handed over and touched, etc. Another type of feature is generated by knowledge-dependent operations such as categorisation, meaning identification, valence judgments, etc.

Another important mechanism within the RFM is redundancy-based retroactive interference. If the features of the item being encoded are identical to the features of prior items, then retroactive overwriting can occur and the redundant features of previous items are lost. That said, overwriting is reversible: interference can be offset by rehearsal. Rehearsal is thought of as a rapid retrieval exercise, where prior knowledge is relied upon to reconstruct degraded representations. Finally, final retrieval involves both primary and secondary memory and relies on a modern version of the time-honoured Luce choice rule: In the RFM, degraded representations in primary memory are used as cues to identify a retrieval candidate from competing candidates in secondary memory. The main components of the model are schematically represented in Figure 1. We will return to the specifics of the RFM later in the paper, when describing the modelling of the data we report.

Recently, the RFM has met with considerable success in modelling a complex series of findings related to the production effect (Cyr et al., 2022; Dauphinee et al., 2024; Saint-Aubin et al., 2021). In the production effect, memory for verbal material improves when encoding involves some form of active encoding relative to items processed more passively, by simply reading them silently (see MacLeod & Bodner, 2017, for an overview). Usually, “active” processing is achieved by asking participants to read items aloud, but the production effect has also been observed by asking participants to sing, mouth, spell, draw, write, type, and even imagine typing the items (Fernandes et al. 2018; Forrin et al., 2012; Jamieson & Spear, 2014; MacLeod et al., 2010; Quinlan & Taylor, 2013, 2019). The production effect has been observed with a variety of tasks including recognition, free recall, immediate serial recall, and order reconstruction (e.g., Cyr et al., 2022; Gionet et al., 2022; Jonker et al., 2014; Kelly et al., 2022; Saint-Aubin et al., 2021). The impact of production on recognition performance has been successfully modelled using MINERVA 2 (Jamieson et al., 2016), REM (Kelly et al., 2022) and attention subsetting theory (Caplan & Guitard, this issue). However, these models cannot be applied directly to recall or order reconstruction tasks, or to visual-spatial materials.

According to the RFM, producing an item adds modality dependent features (Cyr et al., 2022; Saint-Aubin et al., 2021). Forrin et al. (2012) suggested that these additional features could consist of those generated by the auditory presentation derived from hearing one's own voice as well as incorporating the motor features involved in the pronunciation of the items. In the RFM, these additional modality dependent features can increase the distinctiveness of the produced items, especially relative to silent items which lack such extra features. However, this benefit comes at a cost: in the RFM, producing the items interferes with rehearsal. This is easy to understand intuitively; if a participant is busy reading aloud, simultaneous verbal rehearsal will be difficult (e.g., Murray, 1967). The RFM also assumes that rehearsal declines with list length; this is based on the observation, in the verbal domain, that early items are typically more rehearsed than later ones (Bhatarah et al., 2009; Rundus, 1971; Tan & Ward, 2000; Ward 2002). It follows that early items should suffer more from production, relative to items appearing later in the list.

The RFM summarised above was able to successfully model production effect findings for immediate serial recall, immediate serial reconstruction, as well as for free recall, delayed recall, and delayed reconstruction. Moreover, the RFM handled the interactions between production, list composition (all produced, all silent, or mixed), and serial positions. Hence, findings from paradigms that are typical of both short- and long-term memory have been successfully modelled. The implication is that a relatively small set of assumptions and processes can account for a diverse range of findings, covering multiple experimental effects as well as a variety of tasks, from paradigms taken from short-term and long-term memory literatures.

In addition to accounting for known effects, the RFM generates new predictions. Although the RFM can account for the production effect, said effect highlights the role of extra features related to the language processing system – there is no evidence that producing extra features would be beneficial in any other domain. However, the RFM clearly predicts that any encoded features that can be relied upon to increase distinctiveness[1] should lead to an advantage at least for some serial positions. In this paper, we set out to test these ideas by examining the case of visuo-spatial primary memory. We also test the RFM predictions related to the interaction between visuo-spatial production and visuo-spatial rehearsal.

Jones et al. (1995) developed a visuo-spatial task in which dots are serially presented at various locations on a screen. At recall, all dots reappear, and participants must click on them in their presentation order. This task is often considered as a spatial analogue to the immediate order reconstruction task (e.g., Couture & Tremblay, 2006). According to the RFM, clicking on

[1] Providing the encoding effort necessary to procure the extra distinctive features is not prohibitive.

the dots during their presentation would increase the number of modality dependent features, but would also hinder spatial rehearsal of the order and position of both the current and prior items.

The role played by rehearsal in serial memory for verbal material is well-established (e.g., Baddeley, 1986). Further, the first list items are typically rehearsed more, with rehearsal frequency decreasing across list positions (Bhatarah et al., 2009; Rundus, 1971; Tan & Ward, 2000; Ward 2002). In the visuo-spatial domain, an emerging body of research suggests the presence of spatial rehearsal. Said rehearsal can be gaze-based or attentional-based (Souza et al., 2020). Gaze-based rehearsal processes have been investigated with the dot task mentioned above (Morey et al., 2017; Tremblay et al., 2006). For instance, Tremblay et al. inserted a 10-second retention interval during which all to-be-remembered dots were visible. They computed the number of pairs for which there was an eye movement from dot n to dot n+1. Results showed that recall performance systematically increased with the number of rehearsals/fixations of dot pairs. Moreover, when rehearsal was blocked by asking participants to alternate fixations between two irrelevant locations, performance dropped dramatically. Guérard et al. (2009a) further showed that this type of spatial suppression abolishes the path length effect thought to rely on spatial rehearsal [e.g., when the overall length of the imaginary path connecting successive dots is lengthened, recall suffers (Parmentier et al., 2006)].

Echoing the work done with verbal rehearsal, in an early study, Geiselman and Bellezza (1977) investigated the distribution of gaze-based rehearsal across serial positions. In their study, eight to-be-remembered words were presented simultaneously on a single line. Their results showed that the number of gaze-based rehearsals decreased across serial positions. Furthermore, the number of gaze rehearsals was a good predictor of immediate recall. Within the dot task, at encoding, fixation durations on each dot systematically decreased across serial positions (Saint-

Aubin et al., 2007; Morey et al., 2017). Overall, these results suggest that spatial, gaze-based rehearsal likely supports performance in a visuo-spatial memory task like the dot task, and there is a clear suggestion that rehearsal decreases across serial positions, as is the case in the verbal domain (Bhatarah et al., 2009; Rundus, 1971; Tan & Ward, 2000; Ward 2002).

Here, we investigated the presence of a production effect for spatial information. More specifically, participants were asked to memorize the order and position of dots appearing at various spatial locations. At recall, all dots reappeared, and participants had to click on them in their presentation order (Exp.1 and 2). In Experiment 3, at recall, only a blank screen was provided, and participants had to click on the location of the dots in their correct order. This task has been found to be functionally equivalent to verbal serial recall (Couture & Tremblay, 2006; Jones et al., 1995), although specifically relying on spatial representations (Guérard & Tremblay, 2008; Guitard & Saint-Aubin, 2015). In the production condition, participants were asked to click on the items as they were presented. According to the RFM, clicking on the items should add relevant modality-dependent features, improving memory performance. However, asking participants to click on the items would constrain their eye movements and inhibit spatial rehearsal.

In Experiment 1a, we investigated the production effect in an immediate order reconstruction task (see Figure 2). A 6 x 6 grid was visible throughout each trial. The dots appeared within the grid. Souza et al. (2020) showed that gaze-based rehearsal behaviours were more frequent and efficient in the presence of a grid compared to a control condition without a grid, leading to better memory performance. Therefore, if producing items interferes with rehearsal, deleterious effects should be more easily observed, relative to a condition where rehearsal is less likely (Cyr et al., 2022; Dauphinee et al., 2024; Saint-Aubin et al., 2021).

Simultaneously, producing the items should add modality dependent features which would be beneficial to retrieval. The observed result should depend on the relative weight of these two competing factors: the deleterious effect of production on rehearsal and the beneficial impact of additional features associated with production.

This interplay was further investigated in Experiment 1B by modifying the control condition to better equate rehearsal opportunities in the control and the production conditions. Four squares were displayed outside the grid, with one square at each external corner. Participants were asked to click on one of these squares each time a dot appeared. The squares were clicked clockwise with one clicked for each dot. This procedure was modelled after the fifth experiment of Saint-Aubin et al. (2021) in which participants were asked to say an irrelevant word after the presentation of each to-be-remembered word. According to the RFM, under these conditions, performance should be higher for produced than control items.

Experiment 1A

Method

Transparency and Openness. In all experiments, we report all manipulations and measures. All data are available in the Open Science Framework repository (https://osf.io/qj4u3/). Study designs and analyses were not preregistered. The research ethics committee of Université de Moncton approved all experiments.

Sample Size. To determine our sample size, we used G*Power 3.1.9.4 (Faul et al., 2007) and the results of Experiment 1 from Cyr et al. (2022) who also used an 8-item list. Specifically, we used the effect size for the interaction between presentation modality (aloud vs. silent) and serial position (1–8) with the free recall procedure (η_p^2 = .17). An a priori interaction for serial position and production as repeated measures was computed with α = .05, power of .95; default parameters were used for the correlation among the repeated measures as well as the non-sphericity correction. The analysis revealed that eight participants were needed. However, we were cautious, because the impact of production on a spatial task is unknown. We therefore overpowered our design and calculated a sensitivity analysis. The results showed that a total of 24 participants with $\alpha = .05$, power of .95, and the default parameters would allow us to detect a small effect (Cohen's $f = 0.19$).

Participants. Twenty-four participants (16 female, 8 male) were recruited through the Prolific platform. Participants had to be between 18 and 30 years old; be from the United States; have English as their first language, normal or corrected-to-normal vision, a Prolific approval rate of at least 90%, and not have reading or writing related disorders, cognitive impairments, or dementia. These selection criteria were used for all experiments. Participants were paid £3.00; they gave their free and informed consent for all experiments. Five participants were excluded and replaced for not following the instructions (e.g., on almost all trials, they failed to click on the dots during presentation in the production condition or they never produced a response).

Materials. The experiment was controlled via PsyToolkit (Stoet, 2010, 2017) and the display is illustrated in Figure 2. All stimuli were presented on a black background. The stimuli were eight dots 30 pixels across displayed at random locations within a 6 x 6 grid of 600 x 600 pixels. White dots were used for the control condition, blue dots were used for the produced condition and for all conditions yellow dots were used at test. Each trial was initiated by participants' clicking on a green square 40 x 40 pixels in a middle of blank screen.

Design. A 2 x 8 repeated measure design was used with production (control vs. clicked), and serial position (1 to 8) as factors. After two practice trials, there were two, 20-trial blocks, counterbalanced across participants. In the control block, participants did not click on the dots during presentation, while in the produced block, they had to click on each dot as it appeared. For each trial, the location of the dots was randomly drawn (without replacement) from the pool of 36 positions.

Procedure. Participants were tested in one online experimental session lasting approximately 20 minutes. Each trial began when the participants clicked on the green square presented at the center of the computer screen. Immediately after, the 6 x 6 grid appeared and remained visible throughout the trial. After 500 ms, the dots were presented at a rate of one dot every second (1000 ms on, 0 ms off). In the control block, participants were instructed not to click on these dots, while in the production block, participants had to click on the dots as they appeared; their responses were recorded to ensure instructions were followed. After the last dot, the empty grid remained on screen for 1000 ms, before all the dots reappeared simultaneously at their presented locations. Participants were instructed to reconstruct the presentation order by clicking on the dots in their original sequence from the first to the last. Once clicked, dots disappeared; the process continued until all the dots were selected.

Results

Correct responses were analyzed as a function of production (control, produced) and serial position (1 to 8) via a repeated-measure ANOVA. As shown in Figure 3, performance was better in the control condition (*M* = .51, *SD* = .15) relative to the produced condition (*M* = .41, *SD* = .18). This production cost is present on all serial positions; importantly, as predicted by the RFM, it appears slightly larger for the first serial positions.

The repeated-measure ANOVA confirmed these observations. The analysis revealed a main effect of production condition, $F(1, 23) = 15.25$, $p < .001$, $\eta_p^2 = .40$, a main effect of serial position, $F(7, 161) = 78.09$, $p < .001$, $\eta_p^2 = .77$, and an interaction between these factors, *F*(7,

161) = 2.06, p = .026, η_p^2 = .09. The latter interaction was further investigated via Post hoc Tukey's honestly significant difference (HSD) tests. They revealed that the detrimental effect of production was observed for all positions (all ps < .042), except positions 5 (p = .06) and 7 (p = .11).

These results show a large reversed production effect with a detrimental impact of clicking on memory for item order. According to the RFM, this negative impact occurs because production disrupts rehearsal (Cyr et al., 2021; Saint-Aubin et al., 2021) and this disruption is not offset sufficiently by the added modality-dependent features. If rehearsal opportunities were better equated between the production and the control conditions, the positive impact of production-related added features should more easily be observed.

Experiment 1B

Method

Participants. Twenty-four participants (20 female, 4 male) who did not take part in Experiment 1A were recruited through the Prolific platform; the same selection criteria used previously were applied. Eight participants were excluded and replaced for not following the instructions (e.g., not clicking when they had to).

Materials, Design, Procedure. The materials, design, and procedure were identical to Experiment 1A, except for the following. While the grid was present, 4 gray squares of 25 x 25 pixels were shown outside of the 4 grid corners (see Figure 2). As each dot was presented, one of the squares, starting from the top left, changed from gray to red continuing clockwise, at a rate of one change every second (1000 ms, 0 ms off). In the control-square condition, participants had to click on the red squares while they were simultaneously presented with the dots. In the produced condition, participants had to ignore the red squares and click on the dots while they were presented.

Results

As shown in Figure 3, the performance was superior in the produced condition (*M* = .37, *SD* = .17), relative to the control-square condition (*M* = .28, *SD* = .13). The production benefit of clicking on the dots as they appeared was found for almost all serial positions. The ANOVA showed the presence of a main effect of production condition, $F(1, 23) = 17.67$, $p < .001$, $\eta_p^2 = .43$, a main effect of serial position, $F(7, 161) = 23.65$, $p < .001$, $\eta_p^2 = .51$, and an interaction, $F(7, 161) = 5.06$, $p < .001$, $\eta_p^2 = .18$. Post hoc Tukey's HSD tests confirmed that the benefit of clicking on the items relative to clicking on irrelevant information was observed for all serial positions (all *p*s < .015), except the 6th ($p = .86$).

Discussion

Overall, results of Experiments 1A and 1B are compatible with the idea that producing visuo-spatial items by clicking on them is associated with both a benefit and a cost. The predictions based on the RFM were supported by the reported findings: increasing relevant modality-dependent features through production supports recall. However, clicking on the items would hinder performance by interfering with rehearsal. Interactions with serial positions are expected because of the diminishing role of rehearsal across positions. In the present case, the cost related to production appears to outweigh the distinctiveness benefits of production, given that Experiment 1A showed worse performance in the production condition. This hypothesis is supported by the presence of a large positive production effect when the control condition involved clicking on irrelevant locations (Experiment 1B). Overall, performance in the control condition of Experiment 1A, where rehearsal was not impeded, was 51%; that dropped to an overall performance of 28% in Experiment 1B where rehearsal opportunities would have been limited. This 23% decrement is of the same magnitude as the decrement observed in previous studies blocking visuo-spatial rehearsal by requiring irrelevant eye movements (Guérard et al., 2009a; Tremblay et al., 2006). To further establish the role of rehearsal, a between-experiment ANOVA was computed with production (production vs. control) and experiment (Experiment 1A vs. Experiment 1B) as factors. The analysis revealed the expected interaction, $F(1, 46) = 32.35$, $p < .001$, $\eta_p^2 = .41$.

Experiment 2A: Basic Effect

In Experiment 1, a grid was used to promote visuo-spatial rehearsal (Souza et al., 2020). However, the presence of the grid may also have promoted verbal recoding. Labels could be used based on the limited number of grid squares; alternatively, with some effort, positions could be converted into a set of coordinates as in the battleship board game. In the example provided in Figure 2, the dots can be represented by the following coordinates: A6, D1, C5, etc. To reduce the probability of verbal recoding, it has been suggested that the dots be presented on a blank screen without place holders (Jones et al., 1995; Guérard & Tremblay, 2008). Therefore, the design of Experiment 1 was replicated in Experiment 2 without the grid.

Method

Participants. Another sample of twenty-four different participants (14 female, 10 male) who did not take part in either of the previous experiments was recruited through the Prolific platform with the same selection criteria as in Experiment 1. Eleven participants were excluded and replaced for not following the instructions (e.g., not clicking).

Materials, Design, Procedure. The materials, design, and procedure of Experiment 1 were replicated, except for the following changes. The 6 x 6 grid was replaced by a 600 x 600 square with a grey border in which the dots were presented at random locations. Unlike the previous experiment, due to programming constraints, 40 random sequences of 8 locations were created with the rule that any two dots had to be separated by at least the distance of one dot (30 pixels) and all dots had to be presented within the square. Half of the lists were allocated to the control condition and the other half to the experimental condition. The order of the dots within a list was identical for all participants, but the lists were randomized within each block. Lastly, the lists were counterbalanced across participants to ensure that the lists were used equally often in the control and the production condition.

Results

As in Experiment 1, Figure 4 shows that the proportion of correct responses was superior in the control condition (*M* = .47, *SD* = .13) relative to the produced condition in which participants had to click on the dots while they were presented at encoding (*M* = .36, *SD* = .15). The ANOVA revealed a main effect of production condition, $F(1, 23) = 19.17$, $p < .001$, $\eta_p^2 = .45$, a main effect of serial position, $F(7, 161) = 37.19$, $p < .001$, $\eta_p^2 = .62$, and a two-way interaction, $F(7, 161) = 3.89$, $p < .001$, $\eta_p^2 = .14$. Post hoc Tukey's HSD tests indicated that the detrimental effect was observed for early serial positions 1 to 3 and position 6 (all *p*s < .002), but not the other positions (i.e., 4, 5, 7, 8; all *p*s > .06).

Discussion

Results of Experiment 2A nicely reproduced those observed in Experiment 1A. The main difference was the more pronounced interaction between production and serial position. This larger interaction fits well with the hypothesis that in some cases there might have been a verbal recoding of the spatial information in Experiment 1 (Guérard et al., 2009a; Jones et al., 1995). With verbal recoding, the detrimental effect of producing the items on visuospatial rehearsal would be attenuated. Since rehearsal is more prevalent for earlier serial positions, disrupting rehearsal should have more effect on early positions; this is more clearly the case here than in Experiment 1.

Experiment 2B:

Experiment 2B was identical to Experiment 1B except that, in the control condition, participants had to click on the four squares located outside the grid corners, as in Experiment 1B. The experiment served two purposes. First, this within-study conceptual replication of Experiment 1B is important for establishing that the spatial production effect observed in Experiment 1B is reproducible. Second, it is important to demonstrate a spatial production effect when it is not possible to verbally recode the items.

Method

Participants. A novel sample of twenty-four different participants (19 female, 5 male) was recruited via Prolific. None of the participants had taken part in the previous experiments and the eligibility criteria were identical to the previous experiments. As before, some participants (8 in this experiment) were excluded and replaced for not following the instructions (e.g., not clicking).

Materials, Design, Procedure. The materials, the design, and the procedure were identical to those of Experiment 2A, except that in the control condition, participants had to click on the four squares located just outside the corners of the square in which the dots were presented. Participants clicked on the squares in step with dot presentation.

Results

As shown in Figure 4, despite the methodological change (no grid), the results echo those of Experiment 1B. Participants were better when they clicked on the dots when they were presented, that is in the produced condition (*M* = .33, *SD* = .13), relative to when they clicked on the irrelevant squares in the control-square condition (*M* = .25, *SD* = .10). The beneficial effect of clicking on the dots while they were presented can be seen for all serial positions except the last two (7 and 8).

The results from the ANOVA confirmed these observations. Once again, there was a main effect of production condition, $F(1, 23) = 24.41$, $p < .001$, $\eta_p^2 = .51$, a main effect of serial position, $F(7, 161) = 18.44$, $p < .001$, $\eta_p^2 = .44$, and a two-way interaction, $F(7, 161) = 6.92$, $p < .001$, $\eta_p^2 = .23$. In line with the visual inspection, the post hoc Tukey's HSD tests revealed that the beneficial effect of production was observed on the first 6 serial positions (all *p*s < .02), but not on the last two positions (all *p*s > .09).

Discussion

Removing the grid in Experiment 2 produced results that replicated the findings with a grid in Experiment 1. More specifically, as expected by the RFM, a detrimental effect of production when rehearsal opportunities were not equated was observed in Experiment 2A and a beneficial effect of production when rehearsal opportunities were better equated was observed in Experiment 2B. This was further tested by computing a between-experiments ANOVA with production (production vs. control) and experiment (Experiment 2A vs. Experiment 2B) as factors. As predicted by the RFM, the interaction between both factors was significant, $F(1, 46) = 42.29$, $p < .001$, $\eta_p^2 = .48$.

Experiment 3A

In this last series of experiments, we examined whether the pattern of results observed in previous studies would be obtained with a recall task where participants had to remember the actual locations of the presented items. In verbal recall, the production effect is present with both reconstruction and recall, and the effect is larger with recall than with reconstruction (Cyr et al., 2022; Saint-Aubin et al., 2021). In the spatial domain, the boundary conditions of the phenomenon remain unknown. Here we attempted to fill this gap. Experiment 3A was identical to Experiment 2A except that at the point of recall the dots did not reappear. Participants were asked to click on the positions of the presented items, as they remembered them, and in the original order of presentation.

Method

Participants. Twenty-four participants (15 female, 9 male) were recruited via Prolific based on the same inclusion criteria as the previous experiments. In addition, none of the participants had taken part in any of the previous experiments. In this experiment, six participants were excluded and replaced for not following the instructions (e.g., not clicking).

Materials, Design, Procedure. The materials, the design, and the procedure were identical to those used in Experiment 2A, except that list length was reduced to sequences of six dots to account for the increased difficulty of the task. In addition, as mentioned, dots were not re-presented at the point of retrieval: Participants had to click on the blank screen to recall the location of the dots. They were asked to reproduce the presentation order of the items in their clicking responses. When the participants clicked on a location, a yellow dot appeared. This was repeated until the participants had clicked six times.

Results

As previously used in the field, we scored performance with the best fit solution. With the best fit solution, the distance between a given response and the location of all presented dots is computed. The response is then associated to the dot with the smallest distance (Guérard et al., 2009b; Postma & DeHaan, 1996). As shown in Figure 5, the results were consistent with previous experiments: There was a large detrimental effect of production by clicking (*M* =.51, *SD* = .20) relative to the control condition (*M* = .63, *SD* = .18) despite the methodological changes. The production cost was observed across all serial positions.

A 2 x 6 ANOVA was conducted with production condition and serial position as factors. The results from the analysis revealed a main effect of production condition, $F(1, 23) = 31.36$, $p < .001$, $\eta_p^2 = .57$, a main effect of serial position, $F(5, 115) = 27.74$, $p < .001$, $\eta_p^2 = .54$, but the interaction did not reach significance, $F < 1$.

Experiment 3B

Experiment 3B was identical to Experiment 3A except that, as in Experiments 1B and 2B, the control condition involved clicking on four irrelevant squares while the dots were presented. A beneficial production effect was expected because rehearsal opportunities would be more similar across conditions.

Method

Participants. A last group of twenty-four participants (15 female, 9 male) was recruited via Prolific based previous inclusion criteria with the additional constraint that none of the participants had participated in the previous experiments. In this experiment, eight participants were excluded and replaced for not following the instructions (e.g., not clicking).

Materials, Design, Procedure. The materials, the design, and the procedure were identical to Experiment 3A, except that participants in the control condition had to click on irrelevant squares during the presentation of the dots, as in Experiments 1B and 2B.

Results

The results are illustrated in Figure 5. As expected according to RFM, participants were better when they clicked on the dots (*M* = .54, *SD* = .22) during presentation relative to when they clicked on irrelevant squares (*M* = .35, *SD* = .17). The production benefit was observed across serial positions.

An ANOVA confirmed these observations. In line with previous experiments, the analysis revealed the presence of a main effect of production condition, $F(1, 23) = 42.26$, $p < .001$, $\eta_p^2 = .64$, a marginally significant main effect of serial position, $F(5, 115) = 2.28$, $p = .051$, $\eta_p^2 = .09$, and a two-way interaction, $F(5, 115) = 3.80$, $p < .01$, $\eta_p^2 = .14$. The post hoc Tukey's HSD tests confirmed that the production benefit was observed across all positions (all *p*s < .01).

Discussion

Once again, despite the methodological changes and another performance measure (the best fit solution), the results in Experiments 3A and 3B were consistent with the expectations derived from the RFM. In effect, the between-experiments ANOVA revealed the expected two-way interaction between production (production vs. control) and experiment (Experiment 3A vs. Experiment 3B), $F(1, 46) = 72.97$, $p < .001$, $\eta_p^2 = .61$. More specifically, when rehearsal opportunity was not equated across conditions in Experiment 3A, production had a negative impact on response accuracy. Conversely, when rehearsal opportunity was better equated in Experiment 3B, production had a positive effect on recall performance. Overall, the results provide unambiguous evidence to support the robustness of the short-term spatial production effect across all experiments with 3 different methodologies.

Computational Modelling

Our central claim is that the RFM can accommodate the key experimental findings of a spatial production effect that we have observed in the experiments reported above. Our aim is therefore to show that the RFM can produce satisfactory fits to the data and accommodate our results, while relying on a small number of principled and psychologically relevant parameter adjustments. Below we give a summary of the essential elements of the RFM, further details can be found in Saint-Aubin et al. (2021) and Cyr et al. (2022).

In the RFM, items are represented by vectors of features taking values 1-3, or 0 for a feature which has been overwritten. The representation will include modality independent features, in this case basic information about the spatial location, and modality dependent features, which depend on how the stimuli were presented.

On presentation, a representation of the item is stored in secondary memory, and a copy of that representation is also stored in primary memory as a cue, which will subsequently be used to try to retrieve the item. As items are presented, there is a process of retroactive interference where, if feature i of item n matches feature i of previously presented item m , then this feature of item m is overwritten with probability $e^{-\lambda(n-m-1)}$. This means there is complete overwriting of any shared features by the immediately subsequent item, but presentation of an item can also, with a smaller probability, overwrite features of items further back in the list. Moreover, each item is tagged with positional information, which can drift over time as originally proposed by Estes (1989, 1990); the drift parameter is θ which was constant across all simulations for this series (see the Appendix for parameter values). In the RFM, this noisy positional information is called upon to determine which item is lined-up for the next retrieval attempt, i.e. which of the traces in PM will be selected as the following cue for the retrieval mechanism.

Overwriting degrades the representations of items in primary memory, but this can be partially restored by rehearsal. After presentation of each item, there is a rehearsal process which functions to restore any overwritten feature with a probability given by $re^{-\frac{(m-1)^2}{9}}$ where m is the most recently presented item and r is a constant which encodes the rehearsal strength or effectiveness. The factor of 9 means rehearsal tends to be most effective for the first four items in a list, in line with Bhatarah, et al. (2009).

After list presentation, there is a final process of overwriting and rehearsal of modality independent features only, and the resulting set of cues can then be used for recall. Recall in the RFM is similarity based, where the similarity between an item n and a cue m is given by $s(n,m) = e^{-aP_{nm}}$ where P_{nm} is the proportion of mismatching features between the cue and the item. In line with previous work (Cyr et al., 2022), we assume that reconstruction is functionally similar to serial recall. The probability of recalling item n as having appeared in position m is then given by a softmax rule, with temperature parameter τ,

$$p(n,m) = \frac{e^{\frac{s(n,m)}{\tau}}}{\sum_k e^{\frac{s(n,k)}{\tau}}}$$

The important parameters in the model are therefore the numbers of modality dependent and independent features, the distance scaling parameter, the overwriting and rehearsal strengths, and the temperature parameter. Our central results are based on the fact that increasing the number of modality dependent features improves recall, particularly at the end of the list, whereas increasing rehearsal also improves recall, particularly at the start of the list.

As with previous work, our strategy is to try to fit as many conditions simultaneously as possible, to provide the most severe test of the model. We therefore group the data from Experiments 1a+1b, 2a+2b, and 3a+3b and treat these as if they arise from three experiments, each with four conditions. For all three experiments, we fix the overwriting and temperature parameters for all four conditions, we allow the rehearsal parameter to vary between the control, produced, and control-square conditions. In other words, the rehearsal parameter was the same in Experiment A and B. Furthermore, we allow the distance scaling parameter to vary between the conditions that formed Experiments A and B to allow for any overall difficulty increase or variation in participant quality. For all conditions, we set the number of modality independent features to 20; the control and control-square had an additional 5 modality dependent features, and the produced condition had 10 modality dependent features. Experiments 2 and 3 were treated identically except that we assumed fewer modality independent features in the control and control-square conditions (2 instead of 5), reflecting the absence of the grid as a reference point.

Model fitting details can be found in the Appendix, and in Figures 6-8. Generally, fits are good, capturing the patterns in the data well with little systematic variation. Estimates for the best fitting parameters are also included in the Appendix, but the key finding is systematically lower rehearsal rates in the Control-Square vs. Produced vs. Control conditions, for Experiment 1 (0.40, 0.65, 0.98), Experiment 2 (0.12, 0.41, 0.73), and Experiment 3 (.08, 0.65, 0.95) confirming our hypothesis that production and irrelevant clicking suppress rehearsal.

In summary, the RFM can capture the observed pattern in the data from all experiments, both qualitatively, and with good quantitative agreement. It does this by assuming more modality dependent features, and suppressed rehearsal, for produced items.

General Discussion

In this paper, we focused on memory for recent visuo-spatial events. Our capacity to encode and use this type of information underpins numerous everyday activities, including things like orienting relative to other people or objects, planning routes, controlling movement, and building our knowledge of the physical world. Such activities encompass the mundane, e.g., going to the kitchen for a cup of tea as well as the life-preserving, i.e. remembering the position of a hidden driveway indicated on a recent road sign. Better understanding these small feats of processing and memory will deliver benefits in applied areas such as learning, training, and sports, as well as improve our basic knowledge of cognitive functioning. It may also contribute to the development of useful tools which, for example, could support failing memory in everyday life.

With the aim of contributing to these advances, we examined a series of predictions derived from the RFM (Saint-Aubin et al., 2021, 2023). The model has already been successful in accounting for memory performance in the verbal domain, both in paradigms associated with memory over the short-term (immediate serial recall and immediate order reconstruction) and over the long-term (e.g., delayed free recall; Cyr et al., 2022; Poirier et al., 2019). Here, we relied on the RFM to predict new empirical effects in the visuo-spatial domain. More specifically, we tested predictions relating to non-verbal production by calling upon simple visuo-spatial tasks, requiring memory for visual events and their order. In Experiments 1 and 2, the task was to remember the *order* of appearance of a sequence of dots briefly appearing on a computer screen (Jones et al., 1995). In Experiment 3, memory for the order and the *actual* locations of items was required. As outlined above, we assumed feature-based encoding, and assumed that pointing at the to-be-remembered positions provided extra, retrieval-relevant features; rehearsal was also called upon to explain our findings, as was the idea that retrieval is based on relative distinctiveness. We believe these demonstrations to be important because they suggest a combination of processes that are domain general while also acknowledging domain-specific dimensions (see also, Poirier et al., 2019).

The results of all experiments are highly coherent. Overall, production effects are large with an average Cohen's *f* of 1.14. This value is much larger than what is found in the verbal domain with pure lists and an immediate serial recall or an order reconstruction task (e.g., Grenfell-Essam et al., 2017; Saint-Aubin et al., 2021). The coherence of the observed results across three experiments contributes to establishing the production effect in the visuo-spatial domain, and suggests it is a useful phenomenon for assessing memory models (Simons, 2014).

Experiment 1a established that production — in this case clicking on a dot's position as it appeared on the screen — had a negative effect on performance when compared to a condition in which items were simply presented with no pointing. This detrimental effect was replicated in Experiment 2a in which the grid was removed from the screen and in Experiment 3a with a recall task. As mentioned in the introduction, according to the RFM, producing the items should add relevant modality dependent features. With these additional features, the items would be more distinctive and therefore more likely to be properly recalled (Cyr et al., 2022; Dauphinee et al., 2024; Saint-Aubin et al., 2021). However, the addition of said features comes at a cost: It blocks rehearsal. In visual-spatial short-term ordered recall tasks, it has been suggested that items are maintained through rehearsal based on eye movements (e.g., Baddeley, 1986; Guérard et al., 2009a; Morey et al., 2018; Tremblay et al., 2006). In this context, rehearsal has been seen as a way of refreshing the activation of an item in memory (e.g., Baddeley, 1986) or as a strategy for transforming a series of items into a sequence of movements, which would support recall (Logie, 1995). Irrespective of how visual-spatial rehearsal is modeled, the need to drag the mouse from one location to another and to click on the item likely disrupts rehearsal. This deleterious by-product of production would partly or totally offset the benefit of producing the item.

When comparing the verbal and the spatial production effects, one may wonder about the absence of a crossover interaction between production and serial positions. In free recall, order reconstruction and immediate serial recall of verbal materials, compared to control items, produced items are better recalled at the end of the list and less well recalled at the beginning of the list (see Dauphinee et al., 2024; Fawcett et al., 2022 and Gionet et al., 2022, for reviews). As discussed above, in the RFM, one difference between visually and aurally presented verbal materials is the smaller number of modality dependent features associated with visual items. Back in 1990, Nairne assumed that visual items would have 2 modality dependent features, while aurally presented items would have 20 (see Nairne (1990) for a full discussion of this assumption – the basic idea is that verbal material is basically encoded in phonemic and semantic fashion, making visual details less relevant / less likely to be encoded). In line with this, in prior work with verbal materials, we assumed control visual items had 2 modality dependent features and orally produced items had 20 (Cyr et al., 2022; Dauphinee et al., 2024; Saint-Aubin et al., 2021). Here, for visuo-spatial materials, we assumed control and control-square items would have 2 and 5 modality-dependent features respectively, while produced items would have 10. The justification for going from 2 modality-dependent features in studies with verbal materials to 5 for visuo-spatial items is simply that we assumed that visual features would have more relevance to a visuo-spatial task. The choice of 10 modality-dependent features for the produced items, relative to the 20 modality-dependent features for verbal items pronounced aloud, is based on the value that produced the best fits in the first experiment. Although post-hoc, this is justifiable based on the idea that point and click is not as rich in distinctive features as what is produced by well-learned, highly distinctive, auditory patterns such as those generated by reading words aloud. As shown in the modelling above, this increase in the number of modality dependent features for produced items is not sufficient to overcome the disadvantage to rehearsal that comes with production.

In Experiment 1B, we manipulated rehearsal opportunities by introducing a control-square condition where participants had to click on irrelevant locations; the premise was that this activity would constrain eye movements, which in turn inhibits visuo-spatial rehearsal (Guérard et al., 2009a; Tremblay et al., 2006). The prediction was that this would remove or reverse the advantage of the control condition because rehearsal would be better equated when comparing the production and control-square conditions. However, we accept that rehearsal may not be perfectly equated because the requirement to click on irrelevant locations further away from the dots may be more disruptive than clicking on the to-be-remembered items. The assumption was that the advantage procured by the distinctive features generated by production could come to the fore if rehearsal was not providing a significant advantage to the control-square condition. This view was supported by the findings of Experiment 1B, as in the control-square condition, performance dropped below what is seen with production. The effect was replicated in Experiment 2B, with the grid removed, and in Experiment 3B with a recall task. In all experiments, the predictions of the RFM were well supported. Moreover, the model provided a good fit for the observed data in Experiments 1A+1B, 2A+2B, and 3A+3B – even though all the conditions within these experiments were modelled simultaneously, something that is a more stringent test than modelling each experiment's data separately.

Taken together what do the reported finding imply? This paper was centred on testing the idea that visually orienting to a relevant location, planning a movement and executing it would lead to a richer encoding of the studied locations, one that would include features generated by the activities just described. The inclusion or addition of these features matters as they will be somewhat different for each item and because most of them will be absent in the encoding of items within the control conditions. These extra relevant features are thought to increase the relative distinctiveness of each item – that is, they make each item a little more unique relative to other candidates competing for retrieval. It is not the number of features that matters – it is the relative distinctiveness they provide that makes a difference.

The other important idea that heavily contributed to the predictions we put forward relates to the cost associated to the generation of the additional features. Based on prior research, we assumed that the visual control mechanisms involved in identifying locations are also involved in visuo-spatial rehearsal (e.g., Awh & Jonides, 2001; Tremblay et al., 2006). Therefore, the production of an item by pointing and clicking on it would disrupt rehearsal because the same resources are involved in both activities, i.e. rehearsal and point-and-click. It follows that although clicking generates relevant features for the to-be-remembered item, it also interferes with rehearsal of the presented information. To test this idea further, we implemented an additional control condition which required pointing and clicking on an irrelevant position, after the presentation of each item. As mentioned above, the idea here was that if rehearsal is hindered in this manner, the advantage of the extra features obtained in the production condition should be more easily observed. The results supported this prediction.

Importantly, because the control condition was always better than the produced condition, an alternative view suggests that the control-square task, involving irrelevant locations, has a general negative impact, drawing spatial attention away from the primary task. Under this interpretation, the production condition is at an advantage when compared to the control-square condition because the latter draws visual attention away from encoding the to-be-remembered items. Our view is aligned with this suggestion, i.e. clicking on irrelevant positions certainly requires attention; however, we argue that this attentional requirement, as well as specific resources (i.e. gaze control), are drawn away from rehearsal. There is relatively strong evidence that the spatial location task called upon here involves visuo-spatial rehearsal, where participants direct their gaze towards the current position and prior ones (in particular the immediately preceding one; Tremblay et al., 2006). Clicking on irrelevant corner positions makes this type of rehearsal much less likely – precisely because attention and the control of gaze are engaged in an incompatible secondary task. Given there is evidence that spatial rehearsal has more impact on the first serial positions, as in the verbal domain, we predicted —a priori— the observed interaction between the production / control-square conditions and serial positions. It is difficult to see how a general attentional interpretation could lead to such a prediction or even contribute to the interpretation of the findings post-hoc. That said, one can entertain the hypothesis that there is *also* an extra dual-task cost that depresses performance more generally, making it more likely that the production condition would be at an advantage (although production also requires attention). The data reported here cannot eliminate this possibility; further research will be required to more clearly disentangle the mechanisms involved.

One important issue to discuss relates to the fact that the current work, together with prior uses of the RFM within the verbal domain, imply that the same architecture and processes can account for both verbal and non-verbal memory performance. This somewhat more general point was the focus of a paper by Poirier et al (2019). They used the Brooks' verbal and visuo-spatial matrix tasks (Brooks, 1967), and compared conditions where the tasks were performed alone, with articulatory suppression, or with a spatial suppression task. The latter secondary tasks are called upon to selectively interfere with rehearsal of verbal and visuo-spatial material, respectively. The results produced the expected double dissociations: Spatial suppression interfered selectively with the spatial version of the task, while verbal suppression had an equivalent selective impact on verbal short-term recall. Poirier et al. fit the RFM to these findings as well as to data from Guérard and Tremblay (2008). The latter study produced a double dissociation while calling upon more typical verbal and visuo-spatial order reconstruction tasks. In both cases, the model performed well; the implication being that double dissociations can easily be obtained without proposing separate memory systems for verbal and visuo-spatial processing. In essence, popular views which propose modularity or separate systems as an interpretation of double dissociations (e.g., Baddeley, 1986) face a competing view: in the RFM, modularity is replaced by the idea that spatial and verbal tasks generate different types of features. Without rehearsing the full argument here, it was proposed that a feature-based view is more parsimonious than a modular view (see Morey, 2018) and offers flexibility in accounting for multiple benchmark effects in the field while being applicable to multiple stimuli and task types.

Future Directions

In the current series, adding modality-dependent features through a production condition was not beneficial enough to counteract the proposed cost to rehearsal that encoding these extra features entails. As a result, a detrimental effect of production was observed, and as predicted, this detrimental effect was more easily observed at early serial positions. When rehearsal was hindered in the control condition, production generated a clear benefit. By contrast, in the verbal domain, the addition of modality dependent features can offset the cost to rehearsal for the recency part of the serial position; this gives rise to a cross-over interaction between production and serial position (Dauphinee et al., 2024; Gionet et al., 2022). In future studies, it would be important to demonstrate that the addition of modality-dependent features in the visuo-spatial domain can lead to the same crossover interaction as seen in the verbal domain. Moreover, using eye movement monitoring techniques, future studies could, for example, directly assess the impact of production on overt rehearsal based on eye movements (e.g., Tremblay et al., 2006).

Conclusion

In this paper, we report a new phenomenon: the spatial production effect. This new empirical phenomenon is important because it was model guided. The interest of the RFM relies on its capacity to highlight the important dimensions and principles underpinning encoding and retrieval. In the case of the RFM, said principles include the impact of extra, distinctive features, relative to surrounding memoranda. Also, modelling production effects with the RFM highlights the role of rehearsal; the latter is viewed as a form of fast covert retrieval of the item, which reinstates some of features affected by retroactive interference. Another important dimension that is underscored in the RFM relates to the cost-benefit trade-off that can exists between rehearsal and the encoding of supplementary features. In conditions where acquiring extra features involves mechanisms that are also recruited by rehearsal, then one will offset the other. In sum, with the RFM, we further demonstrated how a simple set of principles can account for a complex pattern of empirical findings.

References

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* Vol. 2, pp. 89–195). Oxford: Academic Press. doi:https://doi.org/10.1016/S0079-7421(08)60422-3

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences, 5(3)*, 119–126. https://doi.org/10.1016/S1364-6613(00)01593-X

Baddeley, A. D. (1986) *Working memory*. Oxford: Oxford University Press.

Barrouillet, P., & Camos, V. (2015). *Working memory: Loss and reconstruction*. Hove, UK: Psychology Press. Barrouillet, P., De Paepe,

Bhatarah, P., Ward, G., Smith, J., & Hayes, L. (2009). Examining the relationship between free recall and immediate serial recall: Similar patterns of rehearsal and similar effects of word length, presentation rate, and articulatory suppression. *Memory & Cognition, 37(5)*, 689–713. https://doi.org/10.3758/MC.37.5.689

Brooks, L. R. (1967). The suppression of visualization by reading. *The Quarterly Journal of Experimental Psychology, 19*, 289–299. doi: https://doi.org/10.1080/14640746708400105

Camos, V. (2015). Storing verbal information in working memory. *Current Directions in Psychological Science, 24(6)*, 440–445. https://doi.org/10.1177/0963721415606630

Caplan, J. B., & Guitard, D. (submitted). A feature-space theory of the production effect in recognition.

Couture, M., & Tremblay, S. (2006). Exploring the characteristics of the visuospatial Hebb repetition effect. *Memory & Cognition, 34(8)*, 1720–1729. https://doi.org/10.3758/BF03195933

Cowan, N. (1999). An Embedded-Processes Model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge University Press. https://doi.org/10.1017/CBO9781139174909.006

Cyr, V., Poirier, M., Yearsley, J. M., Guitard, D., Harrigan, I., & Saint-Aubin, J. (2022). The production effect over the long term: Modeling distinctiveness using serial positions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 48(12)*, 1797–1820. https://doi.org/10.1037/xlm0001093.supp (Supplemental)

Dauphinee, I., Roy, M., Guitard, D., Yearsley, J. M., Poirier, M., & Saint-Aubin, J. (2024). Give me enough time to rehearse: Presentation rate modulates the production effect. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-023-02437-5

Engle, R. W. (2018). Working memory and executive attention: A revisit. Perspectives on *Psychological Science, 13(2)*, 190–193. https://doi.org/10.1177/1745691617720478

Fawcett, J. W., Baldwin, M. M., Whitridge, J. W., Swab, M., Malayang, K., Hiscock, B., Drakes, D. H., & Willoughby, H. V. (2022). Production improves recognition and reduces intrusions in between-subject designs: an updated meta-analysis. *Canadian Journal of Experimental Psychology.*

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3 : A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39(2)*, 175–191. https://doi.org/10.3758/BF03193146

Fernandes, M. A., Wammes, J. D., & Meade, M. E. (2018). The surprisingly powerful influence of drawing on memory. *Current Directions in Psychological Science, 27(5)*, 302–308. https://doi.org/10.1177/0963721418755385

Forrin, N. D., MacLeod, C. M., & Ozubko, J. D. (2012). Widening the boundaries of the production effect. *Memory & Cognition, 40(7)*, 1046–1055. https://doi.org/10.3758/s13421-012-0210-8

Geiselman, R. E., & Bellezza, F. S. (1977). Eye movements and overt rehearsal in word recall. *Journal of Experimental Psychology: Human Learning and Memory, 3(3)*, 305–315. https://doi.org/10.1037/0278-7393.3.3.305

Gionet, S., Guitard, D., & Saint-Aubin, J. (2022). The production effect interacts with serial positions: Further evidence from a between-subjects manipulation. *Experimental Psychology, 69(1)*, 12–22. https://doi.org/10.1027/1618-3169/a000540

Grenfell-Essam, R., Ward, G., & Tan, L. (2013). The role of rehearsal on the output order of immediate free recall of short and long lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39(2)*, 317–347. https://doi.org/10.1037/a0028974

Guérard, K., & Tremblay, S. (2008). Revisiting evidence for modularity and functional equivalence across verbal and spatial domains in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34(3)*, 556–569. https://doi.org/10.1037/0278-7393.34.3.556

Guérard, K., Tremblay, S., & Saint-Aubin, J. (2009a). The processing of spatial information in short-term memory: Insights from eye tracking the path length effect. *Acta Psychologica, 132(2)*, 136–144. https://doi.org/10.1016/j.actpsy.2009.01.003

Guérard, K., Tremblay, S., & Saint-Aubin, J. (2009b). Similarity and binding in memory: Bound to be detrimental. *The Quarterly Journal of Experimental Psychology, 62(1)*, 26–32. https://doi.org/10.1080/17470210802215277

Guitard, D., & Saint-Aubin, J. (2015). A replication of "Functional equivalence of verbal and spatial information in serial short-term memory (1995; Experiments 2 and 3)". *Quantitative Methods for Psychology, 11(2)*, r4-r7. https://doi.org/10.20982/tqmp.11.2.r004

Jamieson, R. K., Mewhort, D. J. K., & Hockley, W. E. (2016). A computational account of the production effect: Still playing twenty questions with nature. *Canadian Journal of Experimental Psychology, 70(2),* 154–164. https://doi.org/10.1037/cep0000081

Jamieson, R. K., & Spear, J. (2014). The offline production effect. *Canadian Journal of Experimental Psychology, 68(1),* 20–28. https://doi.org/10.1037/cep0000009

Jonker, T. R., Levene, M., & Macleod, C. M. (2014). Testing the item-order account of design effects using the production effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40(2),* 441–448. https://doi.org/10.1037/a0034977

Jones, D., Farrand, P., Stuart, G., & Morris, N. (1995). Functional equivalence of verbal and spatial information in serial short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21(4),* 1008–1018. https://doi.org/10.1037/0278-7393.21.4.1008

Kelly, M. O., Ensor, T. M., Lu, X., MacLeod, C. M., & Risko, E. F. (2022). Reducing retrieval time modulates the production effect: Empirical evidence and computational accounts. *Journal of Memory and Language, 123*, 1–14. https://doi.org/10.1016/j.jml.2021.104299

Logie, R. H. (1995). *Visuo-spatial working memory*. Hove, UK: Lawrence Erlbaum Associates Ltd.

MacLeod, C. M., & Bodner, G. E. (2017). The production effect in memory. *Current Directions in Psychological Science*, *26(4)*, 390–395. https://doi.org/10.1177/0963721417691356

MacLeod, C. M., Gopie, N., Hourihan, K. L., Neary, K. R., & Ozubko, J. D. (2010). The production effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36(3)*, 671–685. https://doi.org/10.1037/a0018785

Marin, J.-M., Pudlo, P., Robert, C. P., & Ryder, R. J. (2012). Approximate Bayesian computational methods. *Statistics and Computing, 22(6)*, 1167–1180. https://doi.org/10.1007/s11222-011-9288-2

Morey, C. C. (2018). The case against specialized visual-spatial short-term memory. *Psychological Bulletin, 144(8)*, 849–883. https://doi.org/10.1037/bul0000155

Morey, C. C., Mareva, S., Lelonkiewicz, J. R., & Chevalier, N. (2018). Gaze-based rehearsal in children under 7: A developmental investigation of eye movements during a serial spatial memory task. *Developmental Science, 21(3)*, 1–8.

Murray, D. J. (1967). The role of speech responses in short-term memory. *Canadian Journal of Psychology, 21(3)*, 263–276. https://doi.org/10.1037/h0082978

Nairne, J. S. (1988). A framework for interpreting recency effects in immediate serial recall. *Memory & Cognition, 16*, 343–352. doi: https://doi.org/10.3758/BF03197045

Nairne, J. S. (1990). A feature model of immediate memory. *Memory & Cognition, 18*(3), 251–269. https://doi.org/10.3758/BF03213879

Nosofsky, R. (2011). The generalized context model: An exemplar model of classification. In E. Pothos & A. Wills (Eds.), Formal Approaches in Categorization (pp. 18-39). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511921322.002

Oberauer, K. (2009). Design for a working memory. In B. H. Ross (Ed.), *The psychology of learning and motivation*., Vol. 51. (Vol. 51, pp. 45–100). Elsevier Academic Press. https://doi.org/10.1016/S0079-7421(09)51002-X

Page, M. P. A., & Norris, D. (2009). A model linking immediate serial recall, the Hebb repetition effect, and the learning of phonological word forms. *Philosophical Transactions of the Royal Society, 364*, 3737-3753.

Parmentier, F. B. R., Elford, G., & Maybery, M. (2005). Transitional information in spatial serial memory: path characteristics affect recall performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31(3)*, 412–427. https://doi.org/10.1037/0278-7393.31.3.412

Poirier, M., Yearsley, J. M., Saint-Aubin, J., Fortin, C., Gallant, G., & Guitard, D. (2019). Dissociating visuo-spatial and verbal working memory: It's all in the features. *Memory & Cognition, 47(4)*, 603–618. https://doi.org/10.3758/s13421-018-0882-9

Postma, A., & De Haan, E. H. F. (1996). What was where? Memory for object locations. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 49A(1)*, 178–199. https://doi.org/10.1080/027249896392856

Quinlan, C. K., & Taylor, T. L. (2019). Mechanisms underlying the production effect for singing. *Canadian Journal of Experimental Psychology, 73(4)*, 254–264. https://doi.org/10.1037/cep0000179

Quinlan, C. K., & Taylor, T. L. (2013). Enhancing the production effect in memory. *Memory, 21(8)*, 904–915. https://doi.org/10.1080/09658211.2013.766754

Rundus, D. (1971). Analysis of rehearsal processes in free recall. *Journal of Experimental Psychology*, *89*(1), 63–77. https://doi.org/10.1037/h0031185

Saint-Aubin, J., Poirier, M., Yearsley, J. M., Robichaud, J.-M., & Guitard, D. (2023). Modeling verbal short-term memory: A walk around the neighborhood. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 49(2)*, 198–215. https://doi.org/10.1037/xlm0001226

Saint-Aubin, J., Yearsley, J., Poirier, M., Cyr, V., & Guitard, D. (2021). A model of the production effect over the short-term: The cost of relative distinctiveness. *Journal of Memory and Language*, *118*, Article 104219. https://doi.org/10.1016/j.jml.2021.104219

Saint-Aubin, J., Tremblay, S., & Jalbert, A. (2007). Eye movements and serial memory for visual-spatial information: Does time spent fixating contribute to recall? *Experimental Psychology, 54(4)*, 264–272. https://doi.org/10.1027/1618-3169.54.4.264

Simons, D. J. (2014). The value of direct replication. *Perspectives on Psychological Science, 9(1)*, 76–80. https://doi.org/10.1177/1745691613514755

Sisson, S. A., Fan, Y., & Tanaka, M. M. (2007). Sequential Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences of the United States of America, 104(6)*, 1760–1765. https://doi.org/10.1073/pnas.0607208104

Sisson, S. A., Fan, Y., & Tanaka, M. M. (2009). Correction for Sisson et al., sequential Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences of the United States of America, 106(39)*, Article 16889. https://doi.org/10.1073/pnas.0908847106

Souza, A. S., Czoschke, S., & Lange, E. B. (2020). Gaze-based and attention-based rehearsal in spatial working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46(5)*, 980–1003. https://doi.org/10.1037/xlm0000771

Souza, A. S., & Oberauer, K. (2018). Does articulatory rehearsal help immediate serial recall? *Cognitive Psychology*, *107*, 1–21. https://doi.org/10.1016/j.cogpsych.2018.09.002

Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods, 42*, 1096 –1104. http://dx.doi.org/10.3758/BRM.42.4.1096

Stoet, G. (2017). PsyToolKit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, *44*, 24 –31. http://dx.doi.org/10.1177/0098628316677643

Tan, L., & Ward, G. (2000). A recency-based account of the primacy effect in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(6), 1589–1625. https://doi.org/10.1037/0278-7393.26.6.1589

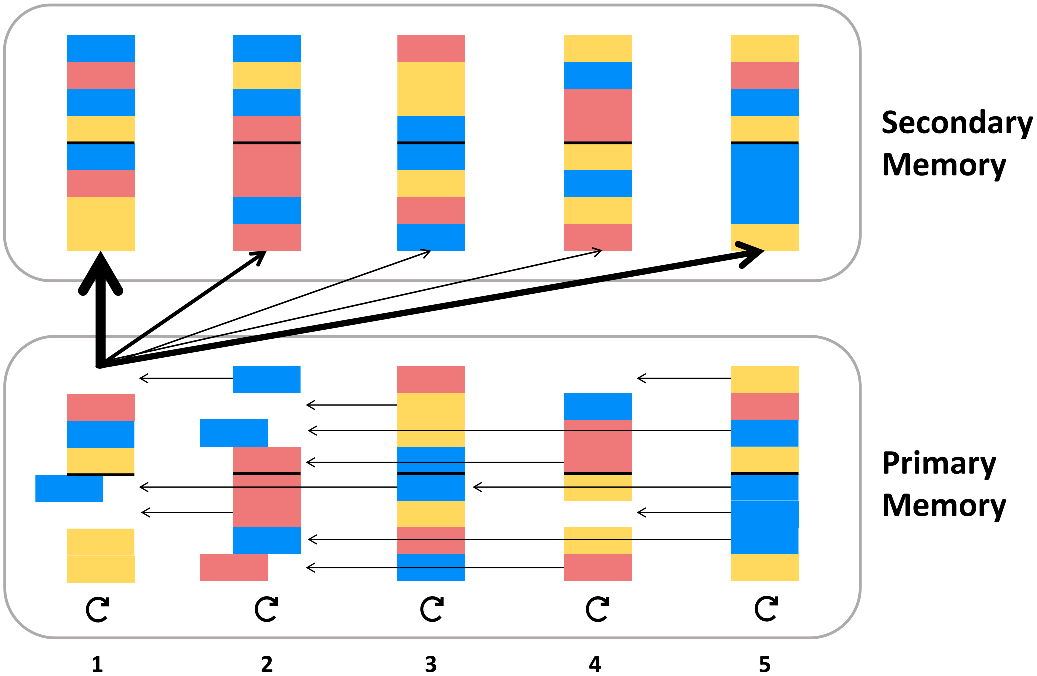
Tremblay, S., Saint-Aubin, J., & Jalbert, A. (2006). Rehearsal in serial memory for visual-spatial information: Evidence from eye movements. *Psychonomic Bulletin & Review, 13(3),* 452–457. https://doi.org/10.3758/BF03193869

Turner, B. M., & Van Zandt, T. (2012). A tutorial on approximate Bayesian computation. *Journal of Mathematical Psychology, 56(2),* 69–85. https://doi.org/10.1016/j.jmp.2012.02.005

Ward, G. (2002). A recency-based account of the list length effect in free recall. *Memory & Cognition, 30(6)*, 885–892. https://doi.org/10.3758/BF03195774

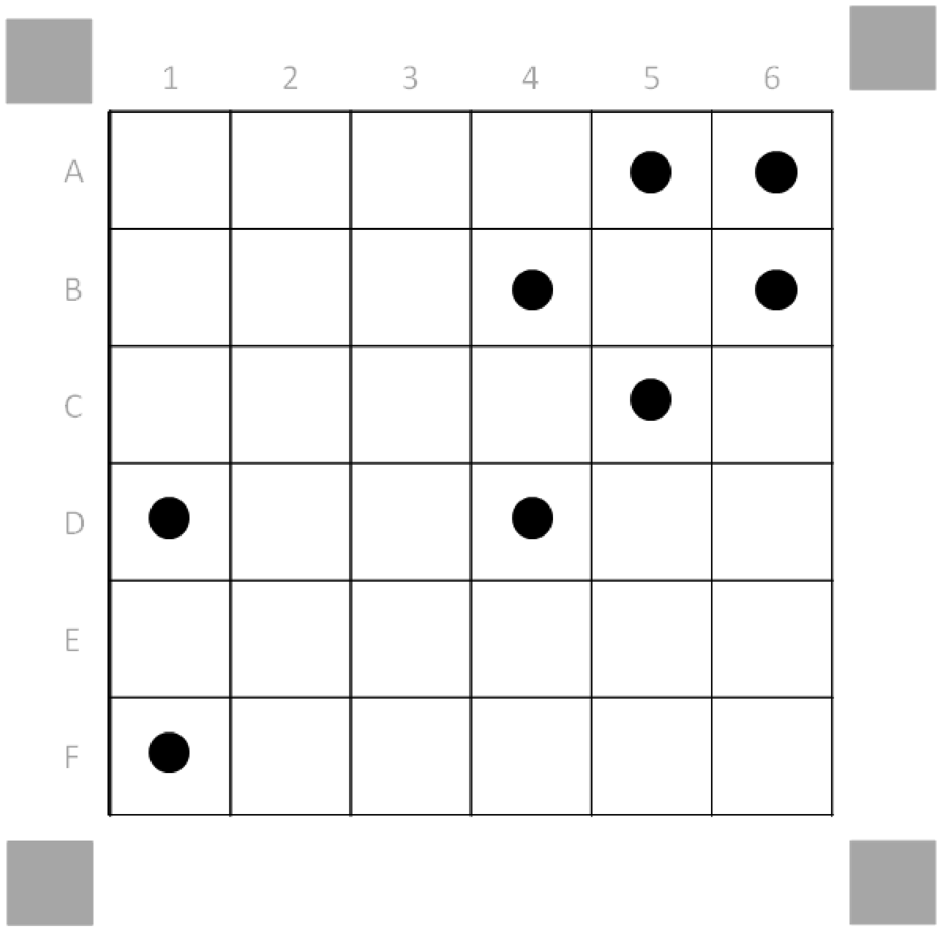
Figure 1.

Schematic illustration of the Revised Feature Model (Saint-Aubin et al., 2021).



Note: Each multi-coloured column represents an item vector in which coloured rectangles stand for distinct features. The yellow, red, and pink rectangles represent the values of 1, 2, and 3. The top four colours represent modality-dependent features, and the bottom four colours represent the modality-independent features. For illustrative purpose, four features are presented in each category whereas, in reality, the numbers are not necessarily the same in each category and there could be many more than four. The arrows to the left illustrate the retroactive interference process. When the same feature, shown here by a coloured rectangle, occupies the same position in two items, the feature of the previous item can be overwritten by the corresponding feature of the subsequent item. Overwriting is illustrated by the white rectangles. As shown by the smaller number of white rectangles with long arrows than with short arrows, the overwriting probability is inversely proportional to the distance between the items. After each item presentation, there is a rehearsal attempt of all items presented so far. The rehearsal attempt is represented by the clockwise open circle arrows. Rehearsal can restore some of the overwritten features shown by the half-retracted blocks. At recall, each degraded vector in Primary Memory is compared with all intact vectors in Secondary Memory and the vector with the highest relative similarity is selected. The recall process of the first item is shown above. The thickness of the lines is proportional to the similarity between the degraded first item and the intact traces in Secondary Memory.

Figure 2

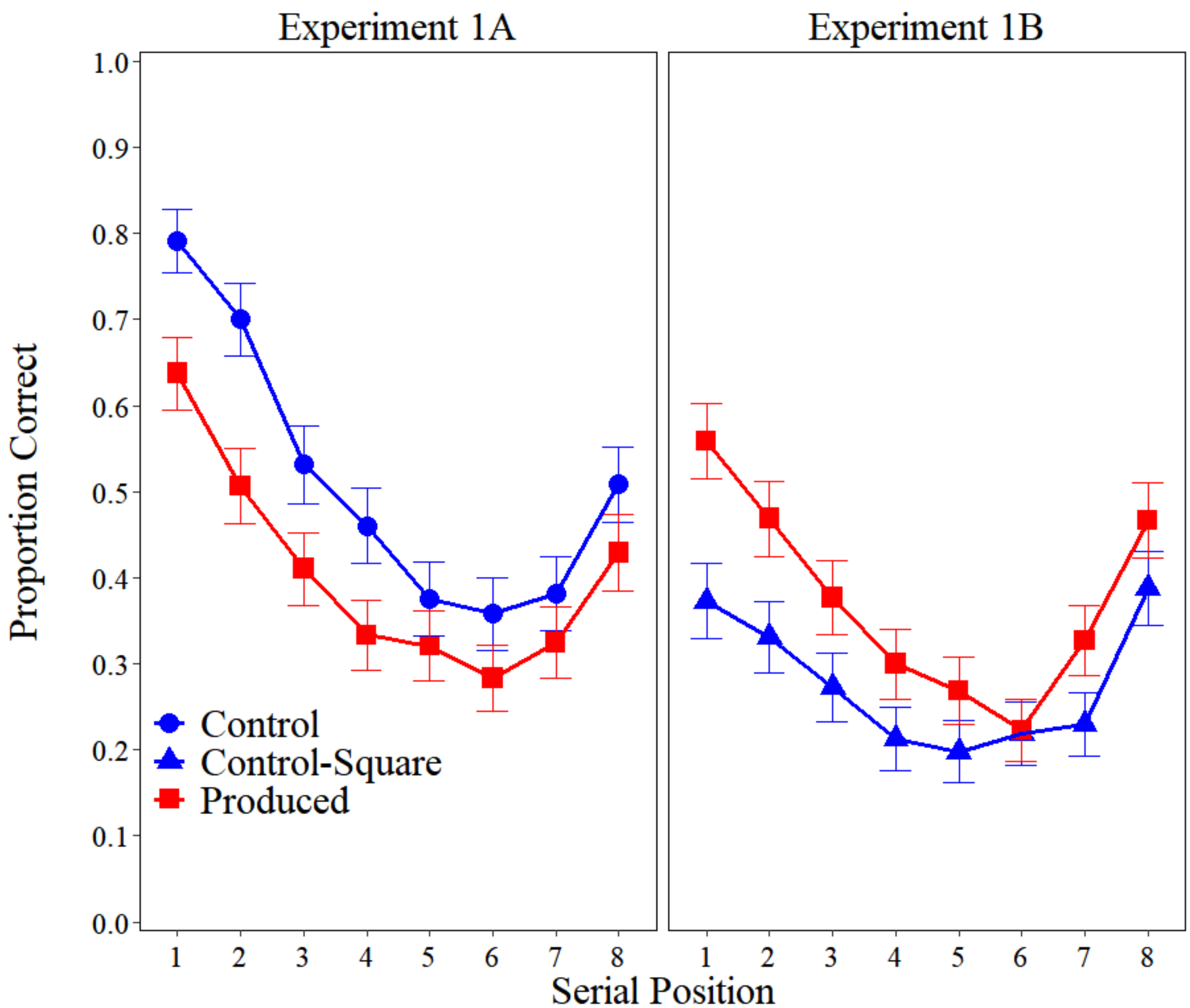


Experiment 1b (with grid): Illustration of the display at the beginning of the recall phase.

Note: The letters and digits providing the coordinates of the dots were not presented. Here, they are displayed to illustrate how participants could have used a verbal recoding strategy in Experiment 1. In Experiments 2 and 3 the grid was not presented and in Experiments 1A, 2A, and 3A, the grey squares were not presented.

Figure 3

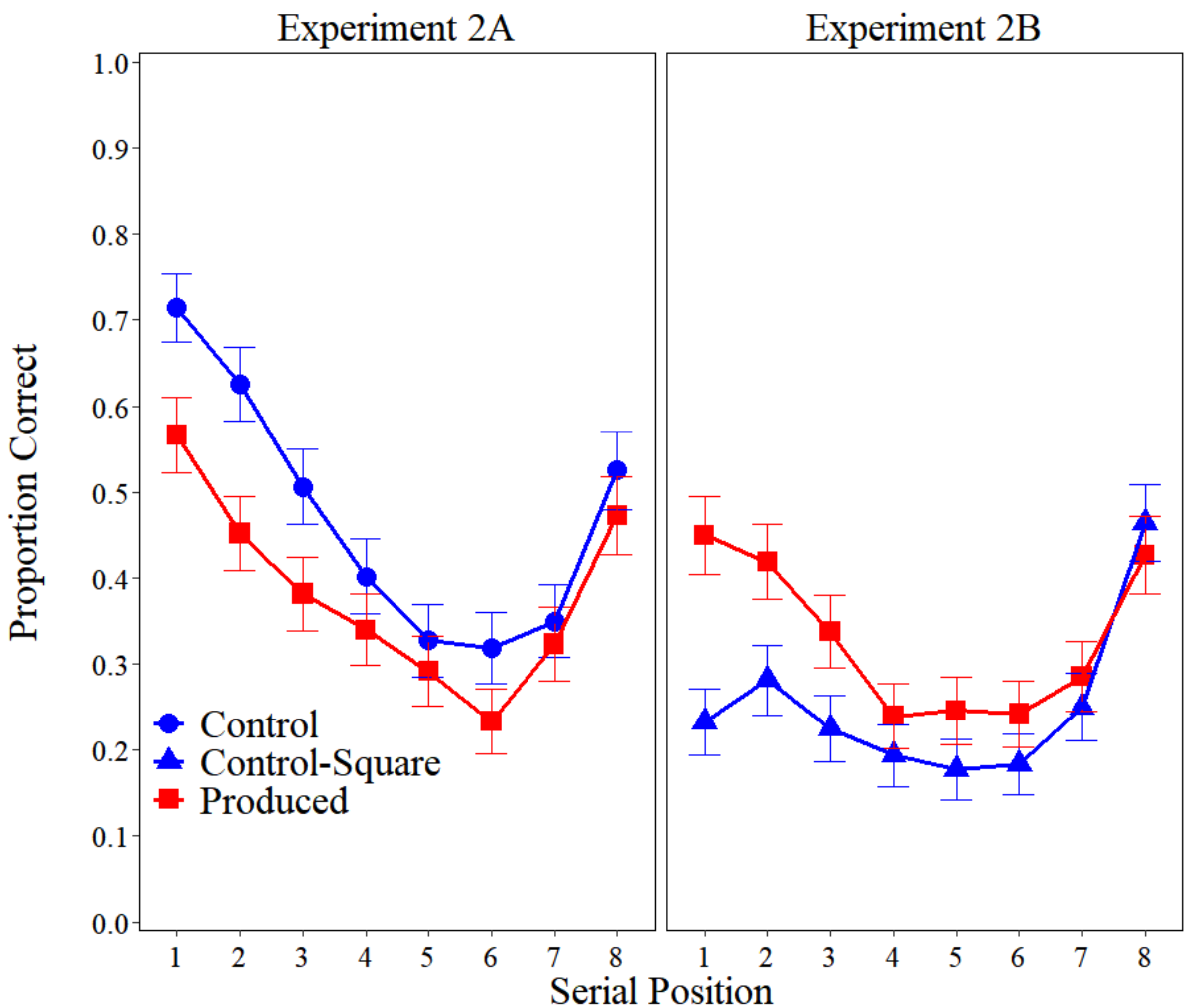
The proportion of correct responses as a function of production condition (control, control-square, produced) and serial position (1 to 8) in Experiment 1A (left panel) and Experiment 1B (right panel) with grid.



Note. Control-square: Participants had to click on the irrelevant red squares presented in one of the four external corners outside the grid while the dots were presented. Produced: Participants had to ignore the irrelevant red squares and click on the dots while they were presented. Error bars represent 95% within-participant confidence intervals computed according to Morey's (2008) procedure.

Figure 4

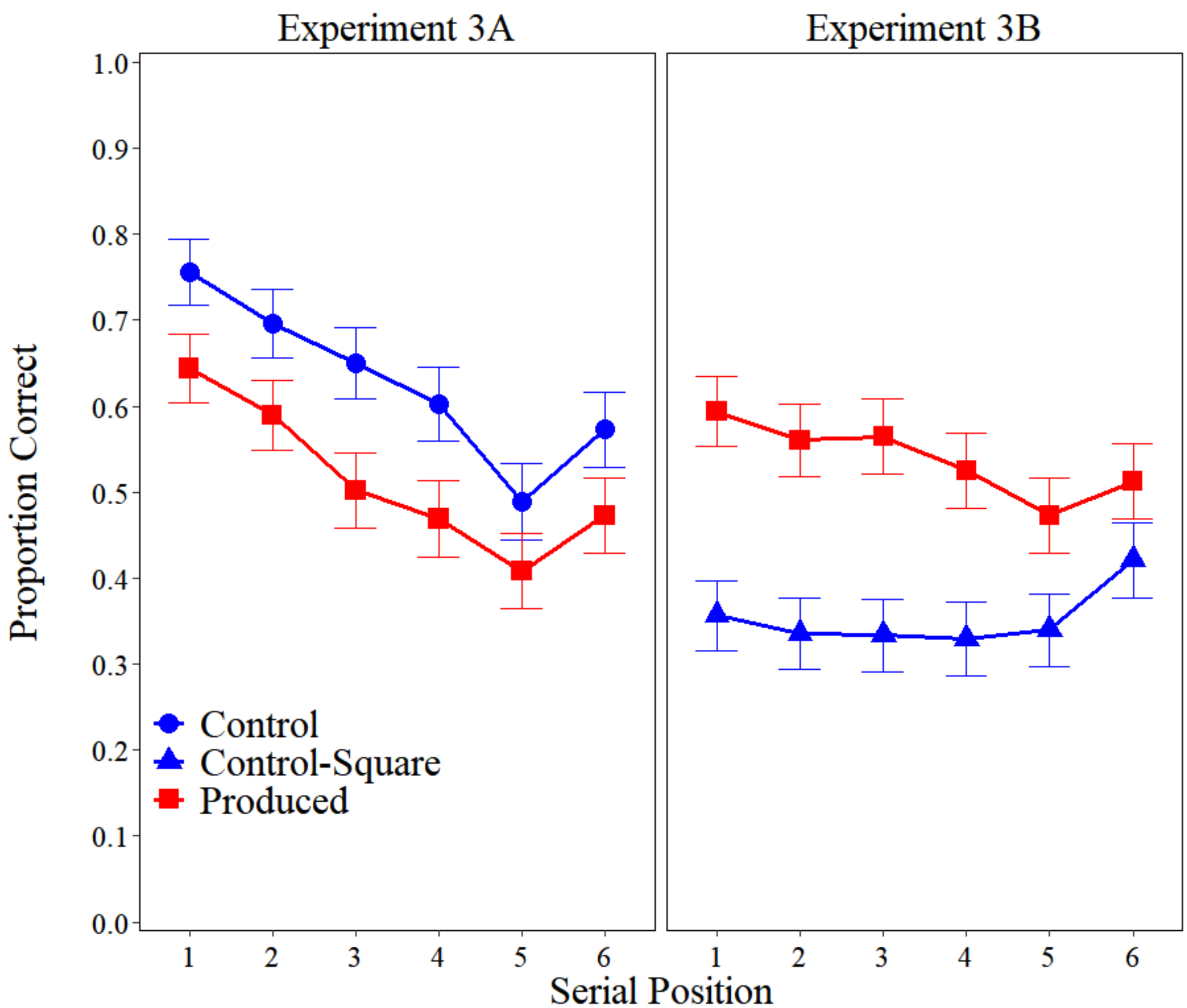
The proportion of correct responses as a function of production condition (control, control-square, produced) and serial position (1 to 8) in Experiment 2A (left panel) and Experiment 2B (right panel) without the grid.



Note. Control-square: Participants had to click on the irrelevant red squares presented in one of the four external corners outside the grid while the dots were presented. Produced: Participants had to ignore the irrelevant red squares and click on the dots while they were presented. Error bars represent 95% within-participant confidence intervals computed according to Morey's (2008) procedure.

Figure 5

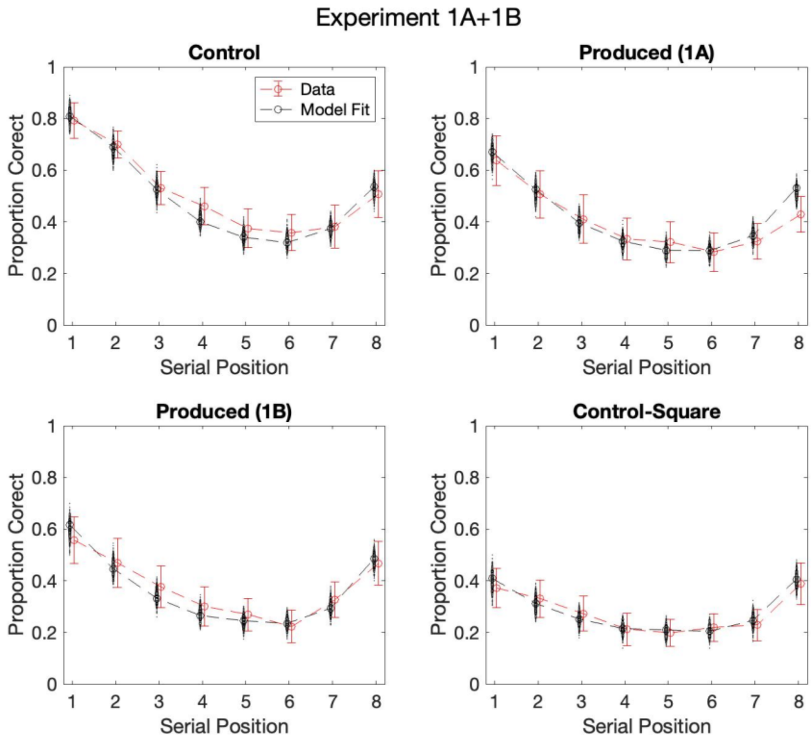
The proportion of correct responses as a function of production condition (control, control-square, produced) and serial position (1 to 8) in Experiment 3A (left panel) and Experiment 3B (right panel) without the grid and recall.



Note. Control-square: Participants had to click on the irrelevant red squares presented in one of the four external corners outside the grid while the dots were presented. Produced: Participants had to ignore the irrelevant red squares and click on the dots while they were presented. Error bars represent 95% within-participant confidence intervals computed according to Morey's (2008) procedure.

Figure 6

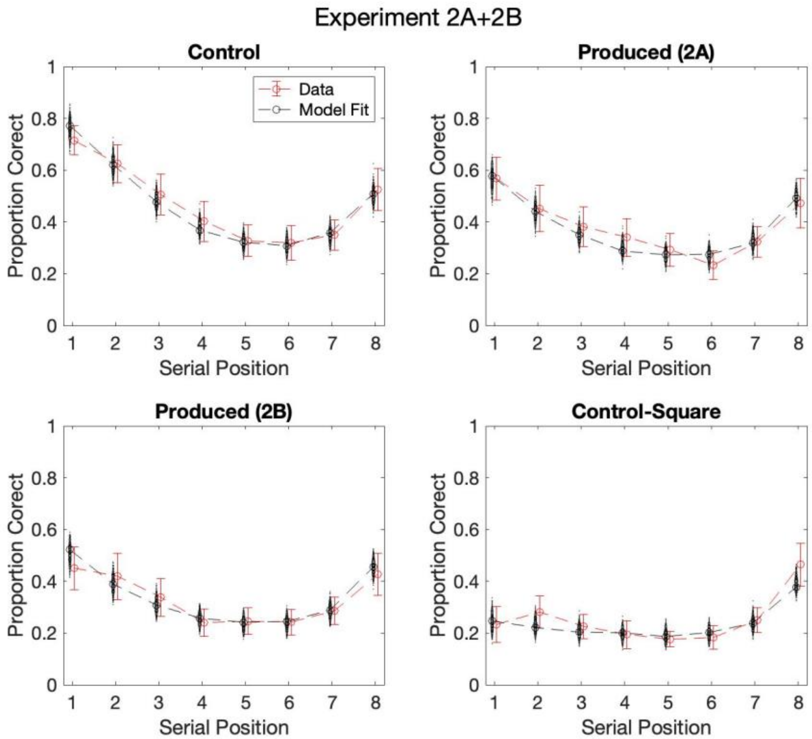
Model fits for Experiments 1A and 1B.



Note: Model fits for each of the four conditions that make up Experiments 1A and 1B. Red lines with error bars are the data and the dashed black line is the mean result of simulating the model 10,000 times using the medians of the posterior distributions for the model parameters. Overall, the model fits match the data well, with little systematic misfitting visible.

Figure 7

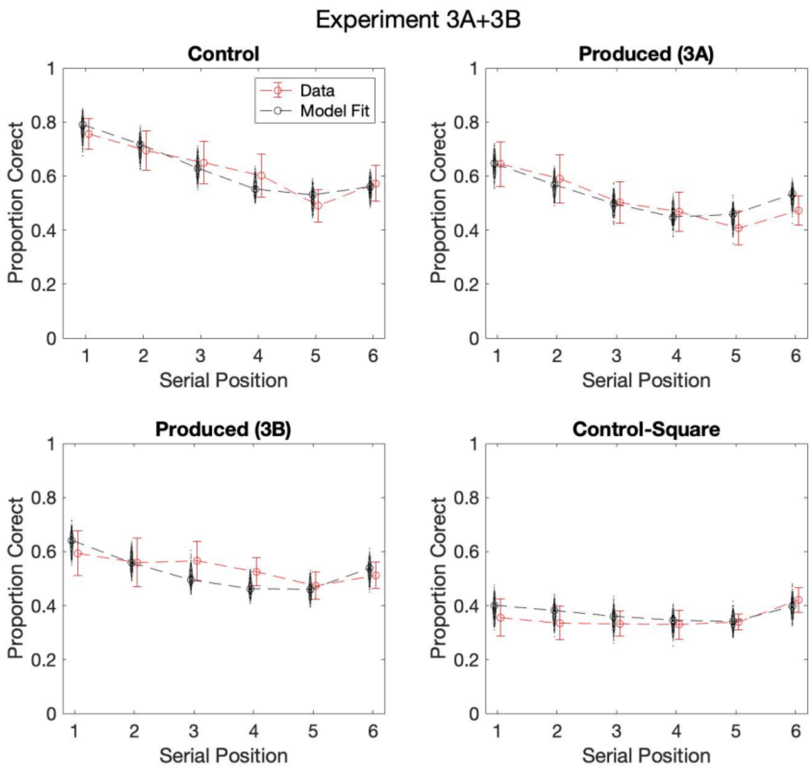
Model fits for Experiments 2A and 2B.



Note: Model fits for each of the four conditions that make up Experiments 2A and 2B. Red lines with error bars are the data and the dashed black line is the mean result of simulating the model 10,000 times using the medians of the posterior distributions for the model parameters. Overall, the model fits match the data well, with little systematic misfitting visible.

Figure 8

Model fits for Experiments 3A and 3B.



Note: Model fits for each of the four conditions that make up Experiments 3A and 4B. Red lines with error bars are the data and the dashed black line is the mean result of simulating the model 10,000 times using the medians of the posterior distributions for the model parameters. Overall, the model fits match the data well, with little systematic misfitting visible.

Appendix: Model Fitting Details.

The RFM is too complex for an analytic expression for the likelihood to be derived, so as with all previous attempts to fit the model to data we called on the methods of Approximate Bayesian Computation (ABC) (see Turner & Van Zandt, 2012, or Marin et al., 2012, for a review).

Following Poirier et al. (2019), Saint-Aubin et al. (2021, 2023), and Cyr et al (2021) we used ABC Partial Rejection Control (ABC-PRC) (Sisson et al., 2007, 2009). ABC-PRC works by repeatedly sampling from a prior over the parameter space until it finds a set of parameters which generate a set of summary statistics (in our case serial position curves) sufficiently close to the data, as determined by the discrepancy function. When this happens, the algorithm stores these parameter values, and moves on to the next particle in the generation. Once all particles in a generation have been associated with parameter sets, the algorithm gives each particle a weight depending on the prior, and then begins a new generation, sampling from the previous generation with probabilities given by the weights, and repeatedly perturbing around the previous parameter values until a set is found producing summary statistics even closer to the data. Once the required number of generations have elapsed posterior estimates for the parameters can be obtained as the fraction of particles in the final generation with that parameter value. Posterior predicted distributions of the summary statistics are also easily obtained. For full details see Sisson et al. (2007) (also note the errata, Sisson et al., 2009).

The important parameters for ABC-PRC are the number of particles (set to 1000 for all fits reported here), the details of the prior, the proposal distributions, and the minimum tolerances for each fit. The proposal distribution and tolerances can be found in the code in the OSF project. Priors, and resulting posterior distributions are summarized in Table A1. In addition the model includes a number of parameters that are not varied and set to standard values across different studies. For example, the theta (θ) parameter is set to 5%, and the number of modality independent features for an item is set to 20.

Table A1

Parameter	Prior	Experiment 1 Median (95% HDI)	Experiment 2 Median (95% HDI)	Experiment 3 Median (95% HDI)
Distance Scaling Parameter a_a	$Normal(4,1)$	4.06 (3.54, 4.53)	3.21 (2.74, 3.70)	2.09 (1.44, 2.76)
Distance Scaling Parameter a_b	$Normal(4,1)$	4.64 (4.15, 5.10)	3.79 (3.35, 4.29)	2.08 (1.41, 2.69)
Overwriting Parameter λ	$Normal(1,0.3)$	0.329 (0.287, 0.379)	0.376 (0.318, 0.437)	0.468 (0.376, 0.574)
Rehearsal Parameter, Control $r_{Control}$	$Beta(1.5,1.5)$	0.982 (0.938, .998)	0.731 (0.630, 0.816)	0.881 (0.797, 0.946)

Rehearsal Parameter, Produced $r_{Produced}$	$Beta(1.5,1.5)$	0.651 (0.571, 0.737)	0.414 (0.343, 0.498)	0.474 (0.387, 0.566)
Rehearsal Parameter, Control-Square $r_{Control-Square}$	$Beta(1.5,1.5)$	0.399 (0.312, 0.483)	0.119 (0.059, 0.197)	0.127 (0.075, 0.192)
Temperature Parameter τ	$HalfNormal(0,0.3$	0.089 (0.075, 0.103)	0.109 (0.094, 0.122)	0.177 (0.163, 0.191)

Note. Table of parameters in the RFM which were estimated in the model fitting, together with the prior distributions and Medians and 95% HDIs of the posterior distributions.