Original Article

Visuospatial bootstrapping: spatialized displays enhance digit and nonword sequence learning

Stephen Darling,[1] (ID) Jelena Havelka,[2] (ID) Richard J. Allen,[2] (ID) Elle Bunyan,[1] (ID) and Lise Flornes[1]

[1]Division of Psychology, Sociology and Education, Queen Margaret University, Edinburgh, UK. [2]School of Psychology, University of Leeds, Leeds, UK

Address for correspondence: Stephen Darling, Division of Psychology, Sociology and Education, Queen Margaret University, Queen Margaret University Drive, Musselburgh, Edinburgh EH21 6UU, UK. sdarling@qmu.ac.uk

Visuospatial bootstrapping describes the observation that performance on a verbal memory task is enhanced by presenting the to-be-remembered material in a format with additional embedded spatial information. Thus far, it has only been reported in short-term memory tasks. Here, we report two experiments assessing the impact of spatial information on the learning of sequences in long-term memory. Experiment 1 used digits presented within a familiar numeric keypad as stimuli compared against single digits presented in one location. Experiment 2 used novel nonwords, which were either presented in an unchanging arrangement permitting the building-up of location knowledge or in a constantly changing arrangement. Both experiments demonstrated strong evidence that reliable spatial information facilitated sequence learning, particularly in later sequence positions. It is concluded that the incidental availability of spatialized information during study can facilitate learning of sequences of digits and nonwords. Furthermore, the spatial information can be learned during the task itself and does not need to be preexistent in long-term knowledge.

Keywords: visuospatial bootstrapping; working memory; serial recall; learning; spatialization

Introduction

Serial order recall in memory may be supported by spatialized processing. The "mental whiteboard hypothesis"[1,2] and the "spatial–positional association of response codes (SPoARC)"[3]—both themselves based on observations of a positional effect first reported by Van Dijck and Fias[4]—argue that an internalized spatial representation is used to represent the serial position of items in a sequence.[3–9] There is also evidence of crossmodal integration from the working memory (WM) literature.[10–14]

A series of studies on visuospatial bootstrapping (VSB)[15] have shown evidence that modality-specific subsystems can work together to enhance short-term memory (STM) task performance. VSB involves the comparison of conditions where participants carry out multiple trials of immediate serial recall of digits, responding verbally. The experimental condition is a "keypad" condition, in which the digits 1–9 are all shown on the screen in the form of the stereotypical "T9" mobile phone/PIN entry keypad. The to-be-remembered items are then sequentially highlighted. Typically, performance is improved with keypad displays. Bootstrapping effects have been replicated frequently.[15–23]

VSB has been interpreted in terms of the multicomponent[24–27] model of WM as occurring when a visual presentation contains information that can be used to support—"bootstrap"—the activities of phonological STM. Verbal interference[17] and interference to set-switching processes[19] do not seem to reduce the bootstrapping effect, while spatial interference is associated with its absence. Given that it seems to be relatively independent of higher-level resources, it has been

argued that bootstrapping can be thought of as being one function of the episodic buffer, a component of the WM model that was conceived to account for processes mediating between different long-term memory (LTM) and STM systems as part of the process of encoding episodic information in memory.[25–27] The multicomponent model of WM has some commonality with Paivio's dual coding theory (DCT)[28]; VSB is also consistent with dual coding.

Other influential models of WM do not place such heavy emphasis on modality specificity,[29–32] and further research has argued that the basis for specialized visuospatial memory (VSM) is weak.[33] Although the central observation of bootstrapping—that is, that VSM (of some kind) can support verbal STM when verbal memory is under high load—seems to fit well with modality specificity, it is not incompatible with these other models. The bootstrapping effect represents an example of spatialization—of spatial resources grounding a verbal memory task. Darling and colleagues[15] suggested that such a process of spatialization may be one of the routes whereby STM information can be consolidated into LTMs—one of the purported roles of the episodic buffer. This leads directly to an empirical question that we begin to address in this article: can the bootstrapping effect be observed in studies of learning in LTM? In this paper, we investigate whether bootstrapping can facilitate learning of sequences in LTM by investigating whether bootstrapping would enhance the learning of supraspan sequences.

In two experiments in this paper, the VSB task was adapted to create a supraspan version where the participants aimed to correctly recall sequences over repeated attempts. We predicted that sequence learning would be facilitated in spatialized conditions compared with control conditions. Experiment 1 investigated learning of digit sequences, while Experiment 2 looked at spatialization effects on learning of novel nonword sequences.

Experiment 1

In Experiment 1, participants aimed to correctly learn sequences of 12 digits using the development of the bootstrapping task. Our principal hypothesis was that learning would be facilitated in a condition where digits were presented in a keypad display compared with a control, single-item condition. We

anticipated that we would observe this benefit for keypads in terms of the number of attempts taken to recall the to-be-remembered sequence and that it would also be apparent when we analyzed the number of trials taken to learn each individual item at each position in the sequence.
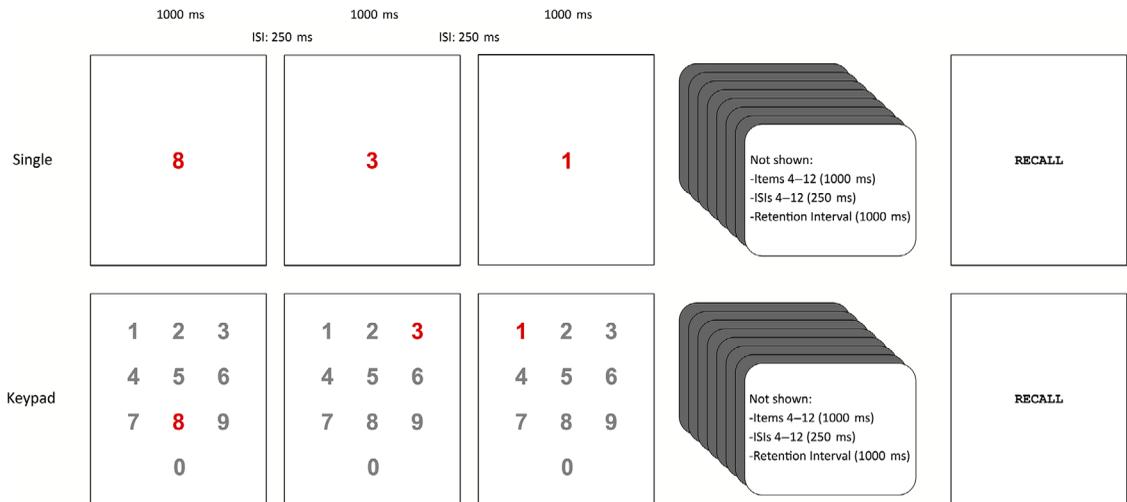
Method

Design. The key variable in this study was the display arrangement used; this was either a single-item display or a keypad display. Participants took part in a set of trials in both display conditions, with the order being counterbalanced across participants. There were five trials in each display condition. In each trial, a sequence of 12 digits was used. This was presented repetitively for immediate recall in a series of subtrials, with the repetition being continued until all 12 items were accurately recalled. Two sets of five pseudorandomly generated sequences were used for this, these sets being counterbalanced across display conditions. Within each set, the sequences were allocated to trials in the same order.

The main measure of learning was a measure of the number of attempts needed to achieve a fully correct recall of the entire 12-item sequence. Second, we measured the number of attempts taken to register the first correct recall of a digit in a given sequence position. Finally, in order to make an assessment of participants' immediate serial recall performance, we measured the total number of correct responses made in each sequence position on the first run of each of the five sequences used. Response times were not measured in this study due to the challenge of measuring and interpreting them from a sequential verbal response.

Participants. Forty-eight participants took part in the experiment.[a] Participants were recruited

Figure 1. Display conditions illustrating the first three items in a sequence beginning with the digits 8,3,1… Each digit was visible for 1000 ms followed by an interstimulus interval (ISI) of 250 milliseconds. There was a retention interval of 1000 ms after the final interstimulus interval. Sequence items 4–12 are not shown in this figure to save space, but were presented in an identical way.

from the Queen Margaret University community, and were offered a small remuneration of expenses in return for the time taken to participate. Mean age of participants was 32.31(SD = 15.94, range = 18–80). Thirty-two participants answered a binary choice to identify their sex as female and 16 responded as male. When asked to select whether they were left- or right-handed, eight identified as left-handed and 40 as right-handed. All participants provided written informed consent to participate, and the research was approved by the Research Ethics Panel at Queen Margaret University.

Materials and procedures. A PC computer with a 19-in (48-cm) 5:4 ratio display set to a resolution of 1024 × 768 pixels was used to present the stimuli, which were compiled using E-PRIME® 2.0.[37] In the first trial, sequence presentation began with the message "Ready!" being presented in the center of the screen for 1500 ms, followed by a fixation cross presented in the center of the screen for 500 milliseconds. This was followed by a presentation of the to-be-remembered sequence. Twelve digits were presented visually, one after the other, each digit being visible for 1000 ms with a blank interdigit interval of 250 ms after all digits. After the final interdigit interval, there was a further interval of 1000 ms before the presentation of

the message "Recall" in the middle of the screen. At this point, participants attempted to recall the entire digit sequence verbally in order. The experimenter scored whether the participant had correctly identified each item in the sequence.

If the participant had failed to recall the sequence correctly, then the next presentation would repeat the sequence to give the participant another recall attempt in the following subtrial. Alternatively, if the participant recalled all items correctly, the next presentation would advance to the next trial, with the next sequence being presented. Once a participant had completed five trials, they completed the five trials in the other display condition.

Figure 1 illustrates the two different display conditions. In the keypad condition, every presentation of an individual digit was achieved by presenting the T9 keypad on the screen, with the digits represented in bold Arial font at point size 48. The keypad included all digits, including zero at the bottom. Digit centers were separated by 130 pixels in both horizontal and vertical directions. The center point of the keypad was centered in the screen. The one to-be-remembered digit was represented entirely in red (RGB 255,0,0), while the other digits were shown in gray (RGB 128,128,128).

In the single-item display condition, every presentation of the to-be-remembered digit was carried

out by showing that specific digit in red (Arial font, point size 48).

Results

Morey and Rouder's[38] *BayesFactor* package was used to calculate Bayes factors in this report.

Mean attempts to criterion for the keypad condition was 4.10 (SE = 0.21), while in the single-item condition, it was 4.86 (SE = 0.17).

A Bayes factor analysis compared the null hypothesis (keypad = single) with the directional experimental hypothesis (keypad > single). The Bayes factor of 161 indicated extremely strong evidence in favor of the experimental hypothesis. Cohen's d^b effect size was 0.79 (95% CI: 0.38–1.21).

Potential effects related to the order in which participants took part in display conditions were assessed by using a Bayesian repeated-measures analysis based on a 2 × 2 ANOVA design on the number of attempts to criterion. The best model included both main effects and interaction. This was preferred to the next best model, which included only the display (BF = 18.84), hence contributing strong evidence for an interaction between the order and display. When the keypad display condition was presented first, there was a substantial difference in attempts to criterion between the keypad display (3.82, SE = 0.30) and the single-item display conditions (5.22, SE = 0.24), but when the single-item displays were presented first, the difference between the keypad (4.39, SE = 0.31) and single (4.50, SE = 0.23) display conditions was comparatively small. Effect size (η_p^2) for the main effect of display was 0.3 (90% CI: 0.12–0.48); for the main effect of order, it was 0.00 (90% CI: 0–0.05); and for the interaction, it was 0.24 (90% CI: 0.08–0.42).

Figure 2 summarizes the number of attempts to first correct recall in each sequence position. A 2 × 12 Bayesian ANOVA showed that the best model included both the main effects of display and sequence position, as well as the interaction between them (BF$_{10}$ = Inf^c). This model was preferred over the model including both main effects
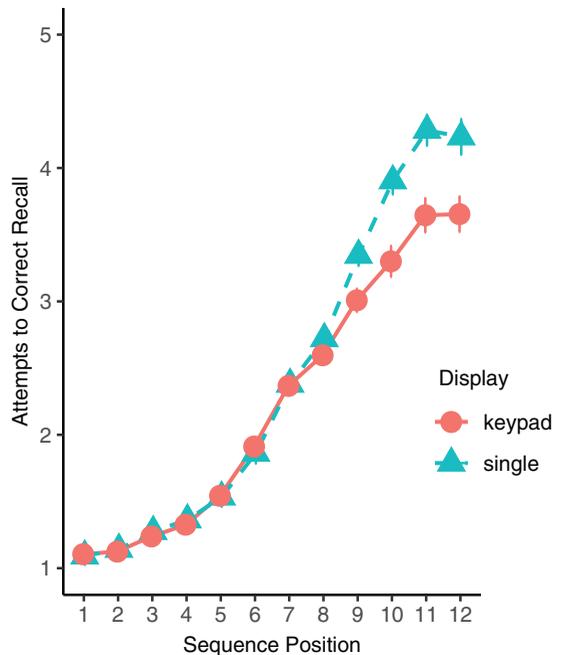
Figure 2. The plot of the number of attempts taken to first correct recall in each sequence position, for keypad and single displays. Error bars show within-subjects standard error (SE).[55]

but without the interaction (BF$_{10}$ = 81,227). This model itself was preferred over both models containing only a single main effect (versus sequence position only, BF$_{10}$ = 131,217; versus display only, BF$_{10} \approx$ Inf). The model including sequence position was preferred to the null model (BF$_{10} \approx$ Inf), but the model containing only the condition was not (BF$_{10}$ = 2.15). Effect size (η_p^2) for the main effect of display was 0.16 (90% CI: 0.03–0.34); for the main effect of sequence position, it was 0.85 (90% CI: 0.83–0.87); and for the interaction, it was 0.17 (90% CI: 0.10–0.20).

These Bayes factors demonstrate extremely strong evidence in favor of a model in which both display type and sequence position, and the interaction between them influence the data. To understand the interaction, we conducted a series of within-subjects Bayesian *t*-tests[38] between single and keypad displays at all 12 sequence positions. The results are summarized in Table 1. There was moderate evidence that there was no difference

---

^{b}Mean difference divided by pooled variance with no adjustment for the within-subjects design is used throughout this paper.

^{c}The term "≈Inf" is used to indicate a value exceeding 1.80 x10ˆ308∼, which is the maximum possible value of a floating-point double number and is used for clarity in the

---

manuscript, we note that the Bayes factors are not infinite but are so large as to approach infinity.

Table 1. Table showing the Bayes factor in favor of the hypothesis that there is a difference (in either direction) between single and keypad displays at each sequence position in the learned sequence in Experiment 1

Sequence position	BF_{10}	d	95% CI for d
1	0.17	−0.09	−0.49 to 0.31
2	0.18	0.10	−0.30 to 0.50
3	0.21	0.16	−0.24 to 0.56
4	0.20	0.14	−0.26 to 0.54
5	0.16	−0.03	−0.43 to 0.37
6	0.18	−0.10	−0.50 to 0.30
7	0.16	0.02	−0.38 to 0.42
8	0.27	0.22	−0.18 to 0.62
9	3.59	0.54	0.13 to 0.95
10	137.82	0.83	0.41 to 1.25
11	136.88	0.83	0.41 to 1.25
12	45.08	0.75	0.33 to 1.16

between display conditions in sequence positions up to position 8, after which there was moderate evidence of a difference at position 9, and then very strong or extremely strong evidence over positions 10–12. In all cases where there was a difference, it was in favor of a benefit to learning (i.e., fewer attempts were needed) in the keypad condition.

Figure 3 summarizes performance on the initial presentation of each sequence.

Analysis of these data using a 2 × 12 Bayesian ANOVA showed that the best model included only the sequence position ($BF_{10} \approx$ Inf). This model was preferred to one including the main effects of both sequence position and display, but this evidence was only of anecdotal strength ($BF_{10} = 2.01$). There was evidence against the interaction when comparing the model with both main effects and interaction to the model including the main effects only ($BF_{10} = 0.0004$), and there was evidence against the model containing only the main effect of display ($BF_{10} = 0.10$). Essentially, these observations confirm the obvious pattern from the graph that there was strong evidence for an effect of position—with earlier positions being associated with better recall—but there was no evidence for an effect of display. Effect size (η_p^2) for the main effect of display was 0.02 (90% CI: 0.00–0.13); for the main effect of sequence position, it was 0.88 (90% CI: 0.85–0.89); and for the interaction, it was 0.01 (90% CI: 0.00–1.00).

The sample on Experiment 1 was unintentionally somewhat skewed by a small number of older participants. In order to ensure that this did not influence results, we reanalyzed our data for this study with participants older than 50 years excluded. Eight participants were excluded, bringing the sample mean age to 25.83 years old (SD = 5.65). The pattern of results observed was the same as when the whole sample was analyzed, so we report data here only from the whole sample.

Discussion

Participants learned a supraspan sequence with fewer attempts when the display contained the extra spatial information from the T9 keypad. This represents a clear extension of VSB from STM to learning. We also observed that the benefit of the spatial displays was greater in later sequence positions.

Evidence for or against a bootstrapping effect in immediate serial recall on the very first presentation was far from compelling. Numerous previous studies[16–23] have replicated bootstrapping effects in STM tasks, so it is worth considering what may have been the reason that the effect was not readily apparent in the current study. One plausible explanation is simply that the task of recalling a 12-item sequence in immediate recall with no prior knowledge of the sequence is too difficult, an explanation that is entirely consistent with the very low recall performance in later sequence positions. It is also possible that the fact that task demands emphasized learning may have influenced performance. Nonetheless, the fact that this could potentially be interpreted as a situation where the bootstrapping effect has not been effectively replicated in STM should be noted, though it should also be held in mind that this experiment is substantively different from a direct replication of the STM bootstrapping effect.

This conclusion has some limits that we sought to expand in Experiment 2. The first is that Experiment 1 only applies to the learning of sequences of digits. Digit sequence memory is important as it allows us to understand learning effects within a highly familiar set, which itself has a highly familiar stereotypical layout (the T9 keypad). However, it is, therefore, difficult to know whether the observed effect specifically requires this very familiar culturally defined and likely overlearned stimulus, or
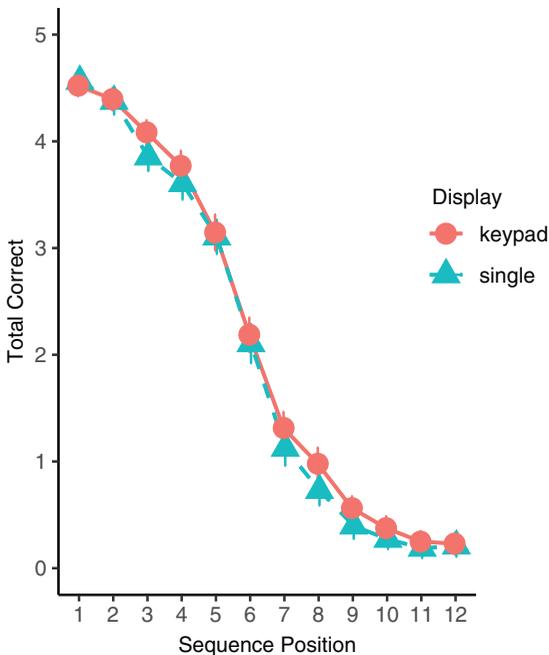
Figure 3. The plot of total correct responses on initial sequence exposure in each sequence position, for keypad and single displays. Error bars show within-subjects SE.[55]

whether it could be observed in other arrays that exhibit spatial consistency over multiple repetitions.

Second, the comparison of a single-item presentation with a familiar keypad incorporates some potential confounds. One is that the keypad display is quite different from the single-item display: it contains more elements and thus may require more processing. This may lead to different depth of processing effects, or greater engagement of attention, or other factors that make it a problematic comparison with a two-dimensional array. Finally, in a single-item display, the to-be-remembered digit is constantly overwritten in the same location. One solution to both problems is to use a different type of the control display in which overwriting does not occur.

Experiment 2

Experiment 2 attempted to address these issues by requiring participants to learn sequences of nonwords using a design modeled on Experiment 1. These were presented either in a "static" 3 × 3 keypad array or in a "random" 3 × 3 array. The static array included the nonwords in a pseudo-random arrangement that remained consistent in

item–location mapping throughout all learning attempts in the block. In the random array, the location of each nonword was randomized within the grid every time a new sequence item was presented. In this latter condition, the location would be useless as a predictor of identity.

Nonwords are considerably more difficult to remember sequentially than individual digits. Hence, the supraspan sequence length was set at nine items, rather than the 12 used in Experiment 1. Nonwords also lack a cardinal arrangement, so any benefits to learning from spatial information in an array must derive from participants' extraction of regularity available in the array—any emergence of a bootstrapping pattern would suggest that participants could extract useful novel spatial information relatively quickly and that the bootstrapping effect on sequence learning was not limited to overlearned sequences. In Experiment 2, we predicted that this regular spatial information, present in the static but not the random condition, would be associated with enhanced memory performance.

Method

Design. In Experiment 2, the design was very similar to Experiment 1 except that the displays used were either a random keypad or a static keypad display, and that the sequences comprised nine nonwords instead of 12 digits.[d]

The same dependent measures were adopted to those used in Experiment 1.

Participants. Twenty-eight participants took part in the experiment.[e] Participants were recruited from the Queen Margaret University community as part of a student research project; some of the participants received course credit for participating. The mean age of participants was 21.71 years old (SD = 3.35, range = 18–34). Twenty-three participants responded to a forced-choice question

---

[d]Note that the first two participants completed eight trials per condition as had been initially intended but the study took too long for participants to tolerate. Therefore, all subsequent participants viewed only five trials per condition—these five trials per condition were the same sequences shown in the same order as the first two participants had been presented with in their first five trials in each condition.

[e]power = 0.80, alpha = 0.05, tails = 1.[34]
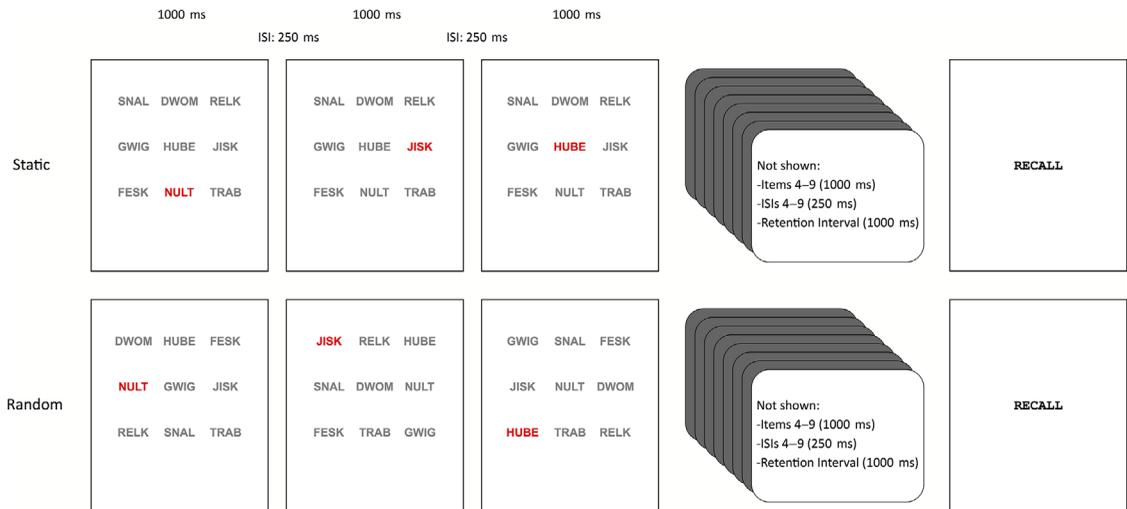
Figure 4. Display conditions illustrating the first three items in a sequence beginning with the nonwords "NULT," "JISK," "HUBE"… Each item was visible for 1000 ms followed by an ISI of 250 milliseconds. There was a retention interval of 1000 ms after the final ISI. Sequence items 4–9 are not shown in this figure to save space but were presented in an identical way.

as being female and five responded as male. When asked to select whether they were left- or right-handed, five identified as left-handed and twenty-three as right-handed. All participants provided written informed consent to participate, and the research was approved by the Research Ethics Panel at Queen Margaret University.

Materials and procedures. The following nine nonwords were used as stimuli in this study: "GWIG," "RELK," "DWOM," "TRAB," "HUBE," "JISK," "SNAL," "FESK," and "NULT." These words were selected from the ARC Nonword Database.[39] They were chosen as monosyllabic four-word nonwords with monomorphemic structure and without illegal bigrams. None of the nonwords shared a starting letter. All the words consist of three consonant phonemes with a vowel between two of them.

An identical computer setup was used to Experiment 1. A very similar procedure was also adopted, with the same timing parameters, with sequence length set at nine nonwords. The experimenter scored a correct recall response for each nonword if a recognizable pronunciation of all phonemes of the nonwords was achieved. The structure of the study was also the same as in Experiment 1—participants repeated each sequence until it was fully recalled, at which point, they moved to the next trial in the set

of five. Once they completed one display condition, they then completed the other.

Both display conditions adopted similar 3×3 keypad-like arrangements where nonwords were presented in bold Arial font at point size 24. Centers were separated by 130 pixels in both horizontal and vertical directions. The central point of the keypad was centered on the screen horizontally and at slightly above the midpoint vertically (the array was located in the exact coordinates used in Experiment 1, but the lack of an item in the "0" position on the keypad meant that the center of the display was no longer in the vertical center of the screen). The one to-be-remembered nonword was represented entirely in red (RGB 255,0,0), while the other nonwords were shown in gray (RGB 128,128,128).

The difference between the two conditions reflected the consistency of the nonword-location mappings. In the static condition, this was randomized at the start of the set of five trials but then remained static across all trials in that condition. In the random keypad condition, the location of each nonword was shuffled each time a new nonword was presented in every sequence. Hence, in the static condition, each participant was exposed to consistent location—identity mapping, whereas in the random condition, the location—identity mapping was uninformative. Figure 4 illustrates the two different display conditions.

Results

Mean attempts to criterion for the static condition was 5.21 (SE = 0.34), while in the random item condition, it was 7.90 (SE = 0.51).

A Bayes factor was computed comparing the null hypothesis (static = random) with the directional experimental hypothesis (static > random). The Bayes factor of 105,467 indicated extremely strong evidence in favor of the experimental hypothesis. Cohen's *d* effect size was 3.19 (95% CI: 2.58–3.79).

Order effects were investigated by using a mixed-design 2 × 2 ANOVA on the number of attempts to criterion. The best model included both just the main effect of display. This was preferred anecdotally to the next best model, which included the main effects of display and order (BF = 2.09), moderately to the model including both effects and the interaction (BF = 3.55), and extremely to the model including only order (BF = 177,608.29). There was no evidence of the presence of an effect or interaction involving order—and there was moderate evidence against an interactive effect. There was little difference between attempts to criterion depending on order for either the static displays (static first: 5.63, SE = 0.41, random first: 4.80, SE = 0.53) or for the random displays (static first: 7.86, SE = 0.88, random first: 7.94, SE = 0.54). Effect size (η_p^2) for the main effect of display was 0.64 (90% CI: 0.41–0.79); for the main effect of order, it was 0.01 (90% CI: 0–0.15); and for the interaction, it was 0.05 (90% CI: 0.00–0.24).

A 2 × 9 Bayesian ANOVA on attempts to correct recall in position (see Fig. 5) showed that the best model included both the main effects of display and sequence position, as well as the interaction between them (BF$_{10}$ = 7.22e+154). This model was preferred over the model including both main effects but without the interaction (BF$_{10}$ = 4.29e + 15). This model itself was preferred over both models containing only a single main effect (versus sequence position only, BF$_{10}$ = 4.85e + 26; versus display only, BF$_{10}$ = 2.34e + 132). The model including sequence position was preferred to the null model (BF$_{10}$ = 3.47e + 112, and the model including only display was preferred to the null model 7.19e + 06). Effect size (η_p^2) for the main effect of display was 0.7 (90% CI: 0.49–0.82); for the main effect of sequence position, it was 0.86 (90% CI: 0.82–0.88); and for the interaction, it was 0.56 (90% CI: 0.46–0.61).
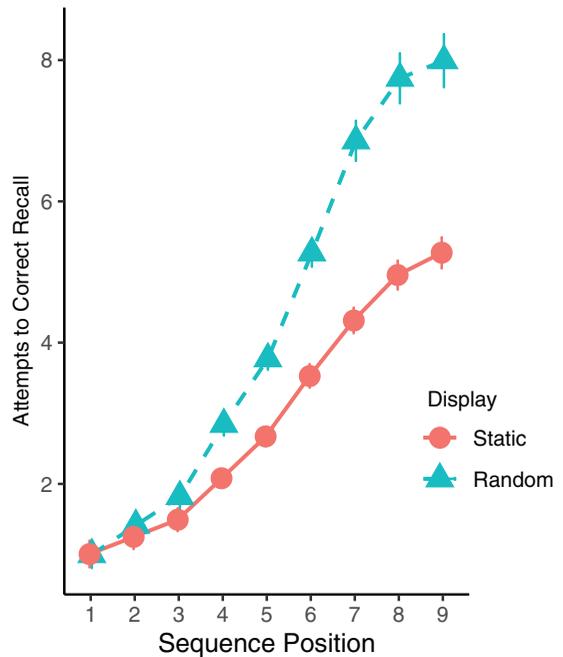
Figure 5. The plot of the number of attempts taken to first correct recall in each sequence position, for static and random displays. Error bars show within-subjects SE.[55]

These Bayes factors demonstrate extremely strong evidence in favor of a model in which both display type, sequence position, and the interaction between them influence the data. A series of within-subjects Bayesian *t*-tests were conducted between static and random displays at all nine sequence positions. The results are summarized in Table 2. There was limited evidence either toward or away from the null hypothesis at positions 1 and 2, but for the remaining sequence positions (3–9), there was very strong evidence that fewer attempts were required for correct recall in the static display condition. The mean difference between conditions tended to increase in later positions. The overall benefit to learning related to the use of a static array driven by a performance benefit to keypads that occurred to a greater degree in later sequence positions.

A 2 × 9 Bayesian ANOVA on correct responses to initial presentations in each position (see Fig. 6) showed that the best model included the two main effects of sequence position and display (BF$_{10}$ = 2.25e + 176). This model was only marginally better than the model including the interaction though, so there was no evidence for the

Table 2. Table showing the Bayes factor in favor of the hypothesis that there is a difference (in either direction) between static and random displays at each sequence position in the learned sequence in Experiment 2

Sequence position	BF_{10}	d	95% CI for d
1	0.32	−0.20	−0.60 to 0.20
2	1.25	0.42	0.02 to 0.82
3	19.17	0.70	0.29 to 1.11
4	425.75	0.97	0.55 to 1.39
5	1141.01	1.05	0.62 to 1.48
6	16,442.72	1.28	0.83 to 1.71
7	408,099.95	1.55	1.09 to 2.01
8	304,801.50	1.53	1.07 to 1.98
9	94,818.49	1.43	0.97 to 1.87

presence or absence of interaction ($BF_{10} = 0.99$). This model was also preferred by several orders of magnitude to the one including only sequence position ($BF_{10} = 8.85e + 05$), which itself was favored over the null hypothesis ($BF_{10} = 2.54e + 170$). There was no evidence in favor of a model containing only display ($BF_{10} = 1.80$). Effect size (η_p^2) for the main effect of display was 0.27 (90% CI: 0.06–0.5); for the main effect of sequence position, it was 0.92 (90% CI: 0.89–0.93); and for the interaction, it was 0.11 (90% CI: 0.03–0.15).

These observations confirm the pattern from the graph that there was extremely strong evidence for the effects of position display. However, there was no evidence of any great strength for the interaction between these two factors.

Discussion

Experiment 2 replicated the effect of display on the number of attempts needed to learn a sequence of information. Learning in the static condition took on average 0.66 fewer attempts than in the random condition in terms of attempts to criterion. Again replicating what was seen in Experiment 1, this effect interacted with sequence position, with spatial information reducing the number of attempts until correct recall more in later positions. Evidence also indicated the presence of a bootstrapping effect in immediate serial recall where participants in the static display condition outperformed those in the random display condition, though there was no evidence of an interaction with sequence position.

The evidence from Experiment 2 indicates that the observation of the benefits of spatialized pre-

sentation to sequence learning seen in Experiment 1 was robust. Experiment 2 additionally demonstrated that spatialized facilitation in VSB is neither restricted to digits nor to the presence of the culturally archetypical spatial arrangement of digits. Instead, they affirm the case that learning sequences of nonwords is contributed to by the presence of consistent spatialized information in the to-be-remembered information.

General discussion

These experiments establish that the provision of ostensibly task-irrelevant spatial information during presentation of sequences facilitates verbal learning for those sequences. VSB facilitated long-term sequence learning to a considerable degree. The replication of this effect in Experiment 2 using nonwords demonstrated that the bootstrapping effect on learning was not limited to numeric or single-character stimuli. Previous research[15–23] has indicated that spatialized information at presentation can facilitate immediate serial recall; these experiments extend that finding to sequence learning.

In Experiment 2, we have observed a bootstrapping effect (in long-term learning) by using regular presentations of a layout that was previously unfamiliar to participants. This suggests that bootstrapping is not limited to familiar displays and that spatial representations capable of supporting it can be learned over relatively short timescales. We note that it is necessary to test this hypothesis fully in the subspan context to be sure it will generalize to the STM bootstrapping case—but the analysis of first attempt recall in Experiment 2 suggests that it should do.

The present experiments focus on the learning of sequences within closed sets of items (either the digits 1–9, or nine nonwords). All previous bootstrapping studies had used closed sets of stimuli, and in Experiment 2 we maintained this approach in order to reduce the changes made between the two experiments. It has recently been noted that SPoARC WM effects can be observed with open sets, but that similar effects are not seen in episodic memory.[8] Whether bootstrapping effects occur in nonclosed sets remains an important question for future research.

These results show that sequence learning benefits from spatialized presentation. What is more, the
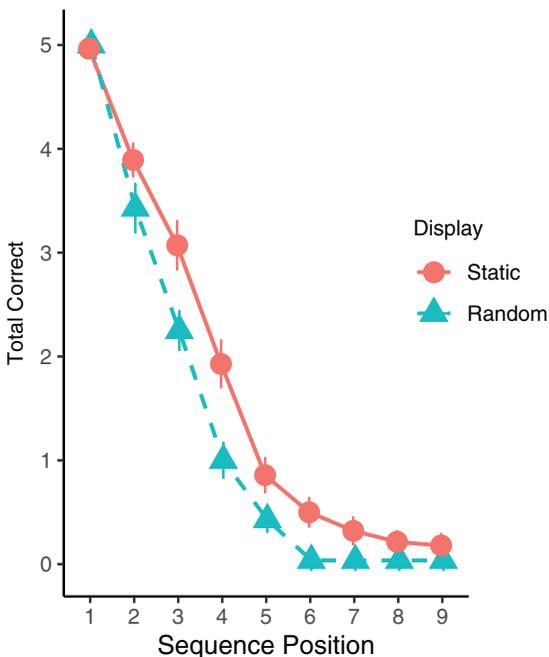
Figure 6. The plot of total correct responses on initial sequence exposure in each sequence position, for keypad and single displays. Error bars show within-subjects SE.[55]

effect size is larger in these two experiments than has typically been reported in the subspan STM versions. The locus of this pattern is not clear. One possibility is that this is caused by the extensive repetition of the same processes that underlie bootstrapping in immediate serial recall. Another is that the effect occurs to some extent within the associative processes underlying LTM—a view consistent with the idea that WM closely interacts with LTM,[24–27] or represents activation within LTM.[31,40] Presently, there are no data to weigh conclusively on this important question. Evidence around VSB-based learning is also consistent with DCT,[28] which is evidenced by phenomena such as the picture superiority effect. Picture superiority seems more linked to slower recollective processes than faster familiarity-oriented processes.[41] Although this is compatible with the results presented here, it does perhaps sit a little less comfortably with emerging evidence that boostrapping may require limited higher-level processing.[19]

The memory benefit from spatialized displays occurred more for items that were later in the presented sequences. This is likely to be largely a con-

sequence of recency effects. Early sequence positions were recalled well on initial exposures. As participants were repeatedly exposed to the to-be-remembered sequences, they became better at recalling later sequence items, and this enhancement was greater in the spatialized conditions. It is possible that sequence positions that were beyond the capacity of WM may have benefitted particularly—but it is also possible that the benefit is the result of accumulated gains of enhanced WM performance on every presentation. A related issue is that although we think it is likely that mechanisms related to bootstrapping in STM (such as, potentially, the episodic buffer) are at the root of the improved learning, this need not be the case: the facilitation may instead rely on LTM phenomena, such as dual coding[28] or on-cue search in LTM.[42] At present, it is not possible to settle this issue conclusively, and it should be a target of future research in this area. Additionally, the link between an individual's immediate memory span and spatialized learning benefits like VSB is unknown—and would also be a useful target for future work.

An *a posteriori* analysis in Experiment 1 suggested that order effects interacted with display condition. A similar analysis in Experiment 2 showed no evidence of the presence of order effects that interacted with display. Tentatively, we suggest that this indicates that the use of single-location presentations (where spatial overwriting may occur) as a comparison with two-dimensional keypads may have underlaid order effects in Experiment 1, and that Experiment 2 demonstrated that the spatialized benefits in these studies can be separated from order effects when such overwriting is eliminated. Nonetheless, this issue requires further investigation within the context of a study where hypotheses around order effects form part of the *a priori* predictions.

The nature of spatialized representations that sustain aspects of sequential encoding has been debated since the emergence of the spatial–numerical association of response codes effect,[43] and more actively since the emergence of ideas around the importance of spatial coding to sequential memory (SM). Although initial descriptions of the mental whiteboard/SPoARC effects[1,3,4] identified a left–right dimension, this is not necessarily thought to be universal. Guida and colleagues[9] have identified that the direction of the

SPoARC effect is related to reading direction. The present research adds two key points to this conversation. First, the spatialized framework that supports sequence learning can be two-dimensional, in that the spatialized information supporting enhanced sequence learning is supported by a two-dimensional array. Second, culturally acquired representations can support sequence learning. The identification of the influence of learned aspects of cognition like reading direction on the SPoARC effect has already demonstrated the cultural nature of spatialization. However, the observation in Experiment 2 that participants can quickly learn to use spatial regularities in a display to facilitate performance on a verbal learning task demonstrates that the cultural grounding needed to sustain these kinds of effects could potentially be learned quickly, rather than being a consequence of lifelong immersion in culturally determined habits. Individuals can learn useful spatial representations for supporting verbal sequence learning over the course of a half-hour experimental run.

The origin of the bootstrapping paradigm is in models of WM; indeed, the original motivation was to test hypotheses around modality specificity in the multicomponent model.[20] However, so far, VSB research has not provided a lever by which to separate the main theoretical approaches. Although the observation of LTM learning effects in a bootstrapping study does tend to argue for models that eschew highly rigid modality specificity, applying this to current theories is something of a strawman argument because it is hard to find a recent model of WM that requires inflexible modality specificity. Even the so-called "multicomponent" model includes multimodal components: the central executive and the episodic buffer.[25,26] Nonetheless, the bootstrapping phenomenon is replicable and extends now to long-term sequence learning, and future iterations of models of WM need to accommodate it. One aspect of the bootstrapping phenomenon that does need to be addressed in theoretical descriptions of WM is the fact that over multiple different occasions, and now in both learning and immediate serial recall versions of the task, it is clearly the case that when verbal memory capacity is exhausted, VSM systems can bootstrap performance, an observation that is consistent with recent ideas from other fields of WM research.[44,45] At some level, this observation must argue for some separa-

tion between visuospatial and verbal memory processing in tasks that involve WM, and, therefore, would tend to lend itself to the case for specialized visual–spatial STM and to argue contrary to some recent claims.[33]

The present study was focused on asking questions about the relation of spatialization and long-term learning. However, there are a couple of broader implications that should motivate future research toward potentially useful outcomes. First, VSB is a method that can enhance the speed of sequential learning (SL) and SM. This is a potentially exciting finding because implicit SL is important in and of itself to reading.[46,47] Bootstrapping also enhances the SM of nonwords. Nonword learning has been linked[48–54] to the development of effective reading and vocabulary skills. Going forward, we await with interest the possible application of this approach to other groups of participants where support for memorization may be useful, such as perhaps children or older adults. We note in this context that data demonstrating that bootstrapping is relatively resilient to aging and hippocampal damage[18,21,23] suggest that it may form a useful tool, but note that, of course, such an extrapolation requires further work to substantiate it.

While the generalization of the present results toward benefits in educational and reading contexts needs to await future research—not least into the longevity of the learning effects—the present experiments stand by themselves in clearly evidencing a benefit of spatially distributed displays to SL, and in demonstrating that the spatialized benefit to verbal learning can come from spatial arrangements that are themselves easily acquired. People learn sequential information more readily when the material is presented in such a way as to allow visual and spatial memory to work together.

Data and preregistered materials

A copy of a predata collection registration of Experiment 1 of this research (including experimental materials) can be downloaded from here: https://osf.io/em9kt/?view_only=798b700a6c4f470 a854d989613a29f9c

A copy of a predata collection registration of Experiment 2 of this research (including experimental materials) can be downloaded from here: https://osf.io/6xd5e/?view_only=be0481cf920f423 68017a7470f536684

An R Markdown version of this manuscript, including the raw data and reproducible analyses, can be downloaded from: https://osf.io/7pwfe/?view_only=371eea496b234363b3df3946bcb67680.

Author contributions

The contribution of the authors to this project is as follows: S.D. wrote the paper and conducted the reported analyses. S.D., R.J.A., and J.H. devised the plan of research, designed and interpreted the studies. E.B. contributed to the design and implementation of Experiment 1. L.F. contributed to the design and implementation of Experiment 2. All authors revised and agreed on the manuscript.

Acknowledgments

Competing interests

The authors declare no competing interests.

References

1. Abrahamse, E., J.-P. van Dijck, S. Majerus, *et al.* 2014. Finding the answer in space: the mental whiteboard hypothesis on serial order in working memory. *Front. Hum. Neurosci.* **8:** 932.
2. Abrahamse, E.L., J.-P. van Dijck & W. Fias. 2017. Grounding verbal working memory: the case of serial order. *Curr. Dir. Psychol. Sci.* **26:** 429–433.
3. Guida, A., A. Leroux, M. Lavielle-Guida, *et al.* 2016. A SPoARC in the dark: spatialization in verbal immediate memory. *Cogn. Sci.* **40:** 2108–2121.
4. van Dijck, J.-P. & W. Fias. 2011. A working memory account for spatial–numerical associations. *Cognition* **119:** 114–119.
5. Antoine, S., M. Ranzini, T. Gebuis, *et al.* 2017. Order information in verbal working memory shifts the subjective midpoint in both the line bisection and the landmark tasks. *Q. J. Exp. Psychol.* **70:** 1973–1983.
6. De Belder, M., E. Abrahamse, E. Kerckhof, *et al.* 2015. Serial position markers in space: visuospatial priming of serial order working memory retrieval. *PLoS One* **10:** e0116469.
7. Ginsburg, V., K. Archambeau, J.-P. van Dijck, *et al.* 2017. Coding of serial order in verbal, visual and spatial working memory. *J. Exp. Psychol.* **146:** 632–650.
8. Guida, A., S. Carnet, M. Normandon, *et al.* 2018. Can spatialisation be extended to episodic memory and open sets? *Memory* **26:** 922–935.
9. Guida, A., A.M. Megreya, M. Lavielle-Guida, *et al.* 2018. Spatialization in working memory is related to literacy and reading direction: culture "literarily" directs our thoughts. *Cognition* **175:** 96–100.
10. Brown, L.A., D. Forbes & J. McConnell. 2006. Short article: limiting the use of verbal coding in the visual patterns test. *Q. J. Exp. Psychol.* **59:** 1169–1176.
11. Brown, L.A. & R.W. Wesley. 2013. Visual working memory is enhanced by mixed strategy use and semantic coding. *J. Cogn. Psychol.* **25:** 328–338.
12. Langerock, N., E. Vergauwe & P. Barrouillet. 2014. The maintenance of cross-domain associations in the episodic buffer. *J. Exp. Psychol. Learn. Mem. Cogn.* **40:** 1096–1109.
13. Morey, C.C. 2009. Integrated cross-domain object storage in working memory: evidence from a verbal-spatial memory task. *Q. J. Exp. Psychol.* **62:** 2235–2251.
14. Pearson, D.G., R.H. Logie & K.J. Gilhooly. 1999. Verbal representations and spatial manipulation during mental synthesis. *Eur. J. Cogn. Psychol.* **11:** 295–314.
15. Darling, S., R.J. Allen & J. Havelka. 2017. Visuospatial bootstrapping: when visuospatial and verbal memory work together. *Curr. Dir. Psychol. Sci.* **26:** 3–9.
16. Allan, A., C.C. Morey, S. Darling, *et al.* 2017. On the right track? Investigating the effect of path characteristics on visuospatial bootstrapping in verbal serial recall. *J. Cogn.* **1:** 3.
17. Allen, R.J., J. Havelka, T. Falcon, *et al.* 2015. Modality specificity and integration in working memory: insights from visuospatial bootstrapping. *J. Exp. Psychol. Lear. Mem. Cogn.* **41:** 820–830.
18. Calia, C., S. Darling, R.J. Allen, *et al.* 2015. Visuospatial bootstrapping: aging and the facilitation of verbal memory by spatial displays. *Arc. Sci. Psychol.* **3:** 74–81.
19. Calia, C., S. Darling, J. Havelka, *et al.* 2019. Visuospatial bootstrapping: binding useful visuospatial information during verbal working memory encoding does not require set-shifting executive resources. *Q. J. Exp. Psychol.* **72:** 913–921.
20. Darling, S. & J. Havelka. 2010. Visuospatial bootstrapping: evidence for binding of verbal and spatial information in working memory. *Q. J. Exp. Psychol.* **63:** 239–245.
21. Darling, S., M.-J. Parker, K.E. Goodall, *et al.* 2014. Visuospatial bootstrapping: implicit binding of verbal working memory to visuospatial representations in children and adults. *J. Exp. Child. Psychol.* **119:** 112–119.
22. Darling, S., R.J. Allen, J. Havelka, *et al.* 2012. Visuospatial bootstrapping: LTM representations are necessary for implicit binding of verbal and visuospatial working memory. *Psychon. Bull. Rev.* **19:** 258–263.
23. Race, E., D.J. Palombo, M. Cadden, *et al.* 2015. Memory integration in amnesia: prior knowledge supports verbal short-term memory. *Neuropsychologia* **70:** 272–280.
24. Baddeley, A.D., G.J. Hitch & R.J. Allen. A multicomponent model of working memory. In *Working Memory: State of the Science*. R.H. Logie, V. Camos & N. Cowan, Eds. Oxford: Oxford University Press. In press.
25. Baddeley, A.D., R.J. Allen & G.J. Hitch. 2011. Binding in visual working memory: the role of the episodic buffer. *Neuropsychologia* **49:** 1393–1400.
26. Baddeley, A. 2000. The episodic buffer: a new component of working memory? *Trends Cogn. Sci.* **4:** 417–423.
27. Hitch, G.J., R.J. Allen & A.D. Baddeley. 2020. Attention and binding in visual working memory: two forms of attention and two kinds of buffer storage. *Atten. Percept. Psychophys.* **82:** 280–293.

28. Paivio, A. 2007. *Mind and its Evolution: A Dual Coding Theoretical Approach*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

29. Barrouillet, P. & V. Camos. 2012. As time goes by: temporal constraints in working memory. *Curr. Dir. Psychol. Sci.* **21:** 413–419.

30. Engle, R.W. & M.J. Kane. 2004. Executive attention, working memory capacity, and a two-factor theory of cognitive control. In *The Psychology of Learning and Motivation: Advances in Research and Theory*. B.H. Ross, Ed.: Vol. **44**. 145–199. New York: Elsevier Science.

31. Oberauer, K. 2009. Chapter 2 Design for a working memory. In *Psychology of Learning and Motivation*. B.H. Ross, Ed.: 45–100. San Diego, CA: Academic Press.

32. Cowan, N. 2010. The magical mystery four: how is working memory capacity limited, and why? *Curr. Dir. Psychol. Sci.* **19:** 51–57.

33. Morey, C.C. 2018. The case against specialized visual-spatial short-term memory. *Psychol. Bull.* **144:** 849–883.

34. Faul, F., E. Erdfelder, A.-G. Lang, *et al.* 2007. G∗power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Meth.* **39:** 175–191.

35. Rouder, J.N. 2014. Optional stopping: no problem for Bayesians. *Psychon. Bull. Rev.* **21:** 301–308.

36. Stefan, A.M., Q.F. Gronau, F.D. Schönbrodt, *et al.* 2019. A tutorial on Bayes Factor Design Analysis using an informed prior. *Behav. Res. Meth.* **51:** 1042–1058.

37. 2015. E-Prime 2.0 (SP2).

38. Morey, R.D. & J.N. Rouder. 2018. BayesFactor: computation of Bayes factors for common designs.

39. Rastle, K., J. Harrington & M. Coltheart. 2002. 358,534 nonwords: the ARC Nonword Database. *Q. J. Exp. Psychol. A* **55:** 1339–1362.

40. Cowan, N. 2016. *Working Memory Capacity: Classic Edition*. New York: Routledge.

41. Boldini, A., R. Russo, S. Punia, *et al.* 2007. Reversing the picture superiority effect: a speedAccuracy trade-off study of recognition memory. *Mem. Cognit.* **35:** 113–123.

42. Unsworth, N. & R.W. Engle. 2007. The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychol. Rev.* **114:** 104–132.

43. Dehaene, S., S. Bossini & P. Giraux. 1993. The mental representation of parity and number magnitude. *J. Exp. Psychol.* **122:** 371–396.

44. Logie, R.H. 2011. The functional organization and capacity limits of working memory. *Curr. Dir. Psychol. Sci.* **20:** 240–245.

45. Logie, R.H., C. Belletier & J.M. Doherty. Integrating theories of working memory. In *Working Memory: State of the Science*. R.H. Logie, V. Camos & N. Cowan, Eds. Oxford: Oxford University Press. In press.

46. Bogaerts, L., A. Szmalec, M. De Maeyer, *et al.* 2016. The involvement of long-term serial-order memory in reading development: a longitudinal study. *J. Exp. Child. Psychol.* **145:** 139–156.

47. van der Kleij, S.W., M.A. Groen, E. Segers, *et al.* 2019. Sequential implicit learning ability predicts growth in reading skills in typical readers and children with dyslexia. *Sci. Stud. Read.* **23:** 77–88.

48. Baddeley, A., S. Gathercole & C. Papagno. 1998. The phonological loop as a language learning device. *Psychol. Rev.* **105:** 158–173.

49. Baddeley, A., C. Papagno & G. Vallar. 1988. When long-term learning depends on short-term storage. *J. Mem. Lang.* **27:** 586–595.

50. Gathercole, S.E., E. Lamont & T.P. Alloway. 2006. Chapter 8 — Working memory in the classroom. In *Working Memory and Education*. S.J. Pickering, Ed.: 219–240. Burlington: Academic Press.

51. Gathercole, S.E. & A.D. Baddeley. 1989. Evaluation of the role of phonological STM in the development of vocabulary in children: a longitudinal study. *J. Mem. Lang.* **28:** 200–213.

52. Gathercole, S.E. & A.D. Baddeley. 1990. Phonological memory deficits in language disordered children: is there a causal connection? *J. Mem. Lang.* **29:** 336–360.

53. Masoura, E.V. & S.E. Gathercole. 1999. Phonological short-term memory and foreign language learning. *Int. J. Psychol.* **34:** 383–388.

54. Papagno, C., T. Valentine & A. Baddeley. 1991. Phonological short-term memory and foreign-language vocabulary learning. *J. Mem. Lang.* **30:** 331–347.

55. Chang, W. 2020. Plotting means and error bars (ggplot2). *Cookbook for R - Graphs - Plotting means and error bars (ggplot2).*