Navigating the Mind's Eye: Understanding Gaze Shifts in Visuospatial Bootstrapping

Teodor Y. Nikolov, Cardiff University

Richard J. Allen, University of Leeds

Jelena Havelka, University of Leeds

Stephen Darling, Queen Margaret University

Baz van de Vegte, Cardiff University

Candice C. Morey, Cardiff University

Abstract

Visuo-spatial bootstrapping refers to the well-replicated phenomena in which serial recall in a purely verbal task is boosted by presenting digits within the familiar spatial layout of a typical telephone keypad. The visuo-spatial bootstrapping phenomena indicates that additional support comes from long-term knowledge of a fixed spatial pattern, and prior experimentation supports the idea that access to this benefit depends on the availability of the visuo-spatial motor system (e.g., Allen et al., 2015). We investigate this by tracking participants' eye movements during encoding and retention of verbal lists to learn whether gaze patterns support verbal memory differently when verbal information is presented in the familiar visual layout. Participants' gaze was recorded during attempts to recall lists of seven digits in three formats: centre of the screen, typical telephone keypad, or a spatially identical layout with randomized number placement. Performance was better with the typical than with the novel layout. Our data show that eye movements differ when encoding and retaining verbal information that has a familiar layout compared with the same verbal information presented in a novel layout, suggesting recruitment of different spatial rehearsal strategies. However, no clear link between gaze pattern and recall accuracy was observed, which suggests that gazes play a limited role in retention, at best.

Keywords: short-term memory, working memory, eye movements, rehearsal, attention

Working memory (WM) is a limited capacity system which allows the storing and manipulation of information, and which is suggested to require an active rehearsal process to prevent forgetting (Baddeley et al., 2021; Barrouillet & Camos, 2021; Cowan et al., 2021). A common assumption among leading theoretical approaches is for a flexible interplay between visual, verbal, and spatial information, pulling in influences of prior knowledge from long-term memory. The extent to which these are involved in any given situation is likely to depend on task context. Investigating the contributions of these components to task performance and attention is important for better understanding working memory and how it can be enhanced.

Visuospatial bootstrapping (VSB; Darling et al., 2017; Darling & Havelka, 2010) refers to a recall benefit arising from the association of verbal items with meaningful visuospatial information stored in long-term memory. Darling and Havelka (2010) investigated the recall of sequences of digits of three groups of participants. One group was presented with the digits one-by-one in the middle of the screen, another group was presented with a typical numerical keypad (as on an ATM) with each number from the sequence being highlighted one by one, while the third group was presented with a display containing a linear keypad. Verbal recall of the sequence was higher in the group which was presented with a typical keypad compared to the central single digit or linear keypad displays, with no significant difference between the latter display types. The VSB effect refers to this memory benefit observed when long-term representations are utilised to aid recall over the absence of such a pattern.

The VSB benefit has been reliably replicated (Allan et al., 2017; Allen et al., 2015, 2023; Calia et al., 2015; Darling et al., 2012, 2014; Darling & Havelka, 2010; Race et al., 2015, 2023) and investigated further to better identify its source. Darling et al. (2012) presented participants with sequences of digits in four different display types. They included

the central single digit and typical keypad displays mentioned above, and introduced a novel static keypad and a novel changing keypad which included a display similar to the typical keypad but containing rearranged digits that were not in their typical location. In the static condition the location of the digits did not change, whereas they did so with every display in the novel changing condition. Recall of the digit sequences was higher again in the typical keypad condition compared to the other conditions, with no significant difference between the other three conditions. It is noteworthy that recall improved in the later trials of the novel static keypad condition suggesting that new arrays can be learned. The authors suggested that the early disadvantage in the novel static keypad condition could be attributed to overcoming the interference of the long-term knowledge pattern of the typical keypad. Darling et al. (2012) suggested that these findings indicated that the long-term knowledge of the locations associated with specific digits enable the integration of visuospatial and verbal information. Further, developmental research supports this view of long-term knowledge playing an important role in VSB. This effect has been investigated in 6-year-old, 9-year-old children, and young adults (Darling et al., 2014), as well as in older adults (Calia et al., 2015), with the effect observed in every group with the exception of the 6-year-old children. The absence of the benefit in young children could be attributed to the lack of sufficient accumulation of typical keypad pattern knowledge, which further provides evidence for the dependence of VSB on accessing long-term representations. Taken together, these findings have been suggested to support the view that long-term memory for the visuospatial layout is essential for the VSB effect.

Several lines of research have also contributed important findings that illuminate what mechanisms contribute to this phenomenon. Allen et al. (2015) introduced articulatory suppression during the encoding phase of the VSB paradigm and found that it resulted in disruption of recall to a greater extent when digits were presented in a single central location

compared to when presented in a typical keypad display. This finding suggests that performance in the central condition is more heavily dependent on verbal maintenance as it does not contain additional environmental cues, whereas in the keypad condition the familiar spatial pattern could reduce the reliance on verbal working memory. Further, Allen et al. (2015) used concurrent spatial tapping either at encoding or at retrieval during an immediate serial recall task, as well as during retention during a delayed recall task with a 5-second retention period (Allen et al., 2023). The VSB effect was not observed when the spatial tapping occurred during encoding or retention in a delayed recall task but was present if the spatial tapping only occurred at retrieval. This finding suggests that the integration of knowledge occurs at encoding. In another study, Allan et al. (2017) manipulated the path complexity of sequences to examine if the quality of spatial representations also play a role in the VSB, alongside the familiarity of the keypad layout. While they observed effects of both path complexity and VSB, these did not interact. The presence of the path complexity effect has been suggested to show that incidental encoding of spatial path occurs in verbal memory tasks, regardless of the layout familiarity. Allan et al. also analysed recall accuracy by serial position and revealed that the VSB effect specifically boosted memory for items toward the end of the lists. Darling et al. (2020) demonstrated a VSB effect with novel non-words which were presented either in a changing or unchanging layout, with the static one allowing the building up of location knowledge. They found that reliable spatial information facilitated sequence learning, especially later in the sequence. The authors suggested that the incidental availability of spatialized information during encoding can facilitate not only recall of digits, but also nonwords. The findings also indicate that the spatial information can be learned during the task itself and does not depend on already familiar long-term patterns. Alongside long-term acquisition of the boot-strapped knowledge, visuo-spatial resources are specifically implicated in the use of that knowledge when encoding and briefly maintaining verbal lists.

Our present investigation aimed to provide further detail about the characteristics of the visuo-spatial resources that support the VSB effect. An unknown factor is how gaze shifts during the conditions of the VSB task. Research shows that people tend to look towards empty locations which previously contained information that is now retrieved from memory (Altmann, 2004; Ferreira et al., 2008; Hoover & Richardson, 2008; Johansson et al., 2006; Richardson et al., 2009; Richardson & Spivey, 2000; Scholz et al., 2018; Spivey & Geng, 2001; van Ede et al., 2019). This phenomenon has been found to have a functional role by facilitating retrieval of both verbal and visuospatial information (Hollingworth, 2009; Laeng et al., 2014; Laeng & Teodorescu, 2002; Scholz et al., 2016). While covert attentional shifts have been found to interfere with spatial working memory, eye movements have been shown to interfere to a greater extent (Lawrence et al., 2004; Pearson & Sahraie, 2003), which is consistent with the position that movement entails spatial attention but spatial attention does not necessarily require movement (Smyth, 1996). Altogether, this evidence points to oculomotor activity as a potential supporting factor that might underlie the VSB advantage.

Eye movements have been surmised to play a specific role for maintenance in visuospatial working memory (Baddeley, 1986; Morey, 2018; Pearson et al., 2014; Postle et al., 2006; Schut et al., 2017; Theeuwes et al., 2005, 2006; Tremblay et al., 2006), but as yet it remains unclear how and when they may support recall in a serial spatial working memory task. Looking to nonmemorized locations during the retention interval interferes with spatial memory (Hale et al., 1996; Postle et al., 2006), supporting the view that there is a tight link between eye movements and spatial memory maintenance. With serial presentation, maintenance processes begin during the presentation of subsequent list items, and as incoming information is presented for later recall, distinct gaze patterns emerge when spatial position is to-be-remembered, as opposed to identity. Czoschke et al. (2019; see also Lange & Engbert, 2013) observed that during the encoding of verbal information, participants made

precise eye movements to to-be-remembered items, contrasting with spatial information

encoding, where such saccades were scarce. At the end of the encoding period, pronounced

fixations to locations that previously contained to-be-remembered items were linked with

increased spatial memory performance, indicating more looking-at-nothing behaviour for

maintaining spatial as opposed to verbal information. This pattern is supported further by

Staudte and Altmann, (2017), who found fewer eye movements towards locations that

previously contained memory items during the identity (verbal) retrieval condition compared

to the location retrieval condition. These findings highlight a distinct oculomotor pattern

during encoding and maintenance of verbal information versus visuospatial information.

However, while these patterns may occur consistently, whether they support recall

accuracy in a meaningful way is less clear. Lower oculomotor activity is often observed

during the retention period of a spatial memory task (e.g., Pearson and Sahraie, 2003,

Experiment 5; Morey et al., 2018), which is unexpected if eye movements or oculomotor

planning make a robust contribution to maintaining locations. Placeholders for remembered

locations increase spontaneous eye movements during retention (Loaiza & Souza, 2022),

which does not seem wholly consistent with the idea that eye movements are actively used

for maintenance. Even when spontaneous eye movements are observed during retention, their

utility is not clear. Loaiza and Souza found no correspondence between looking at the

location of a to-be-remembered item and recall precision. In contrast to these findings,

Tremblay et al. (2006) and Morey et al. (2018) found that spatial serial reconstruction

accuracy increased when participants fixated elements of the list in order during retention.

However, in both investigations, participants engaged in quite limited amounts of ordered

looking despite having a free interval of several seconds. Altogether, this body of evidence

suggests that looking back toward previously presented locations might modestly strengthen

memory for a small subset of them, which makes support from gaze a plausible contributor to the VSB effect.

In the present study, we aimed to investigate gaze during a VSB task. We eye-tracked participants during presentation of seven digits shown in three configurations (familiar telephone keypad, dynamic randomised novel keypad, and centrally presented single digit) and during a 5-second retention interval. We imposed the 5-second retention period between the end of list presentation and the test prompt to learn whether participants use free time to fixate the encoded positions of to-be-remembered digits. During this retention period we presented an empty keypad grid in the typical and novel-changing keypad conditions, and a blank display in the central condition. Differences in performance would reflect contributions from the presence of a spatial layout. If the visuo-spatial bootstrapping advantage occurs because of domain-specific spatial encoding, then we may see deviations in the patterns of eye movements made to familiar keypads where the spatial layout confers information about the to-be-remembered digits compared to random arrangements where the spatial locations provide no relevant clues.

Method

Participants

Forty-one participants took part in the experiment. Three participants were excluded due to poor eye tracking calibration and validation values (mean spatial accuracy was worse than 0.5° or a maximum spatial accuracy worse than 1° for each performed validation procedure). The 38 remaining participants (14 male) ranged from 19 to 40 ($M= 24.53$; $SD= 5.31$). This sample size is comparable to that of similar published studies in which clear advantages were observed for typical keypad displays (Allan et al., 2017; Allen et al., 2015, 2023; Darling et al., 2012, 2020; Darling & Havelka, 2010). All participants were students at Cardiff University and were recruited via the School of Psychology's participant panel, social media

and word of mouth. All had normal or corrected-to-normal vision. Participants received a £10

honorarium.

Design, materials and procedure

   The task was administered using OpenSesame (Mathôt et al., 2012; paradigm can be

found at https://osf.io/hu9a8/) on a display monitor with a resolution of 1024x768 pixels

(width 53.2 cm; height: 30 cm; refresh rate 60Hz), viewed at a distance of 60 cm. Eye

movements were recorded using a desk-mounted SR Research EyeLink 1000 Plus eye tracker

which recorded monocular eye movements at 1000Hz using pupil and corneal reflection

tracking. Fixations were detected using the standard SR Research algorithm. Participants

were tested individually in a dimly illuminated room. Participants completed three practice

trials, one per display condition, before beginning the experiment. They were given the

opportunity to ask questions before beginning the experimental procedure. A 5-point

calibration grid, followed by a 5-point validation grid was used to fit and test the spatial

accuracy of the eye-tracker at the start of experiment. If validation showed a mean spatial

accuracy worse than 1 visual degree or a maximum spatial accuracy worse than 2 visual

degrees, calibration and setup were repeated. Each trial began with a drift check (a central

point which the participant had to fixate) that was manually accepted by the researcher, and if

both eyes had an error of more than 0.50 degrees, then the calibration and validation

procedure was repeated. Participants completed 3 randomly ordered blocks of 8 trials each, 2

blocks per display condition (48 trials in total).

   To-be-remembered 7-digit lists were randomly determined at run time, selected from

the set {0, 1, 2, 3, 4, 5, 6, 7, 8, 9} without replacement. In all conditions, a white 3 x 3 x 3 +

1-cell grid resembling the layout of a telephone number pad was presented onscreen, against

a dark grey background (#2E3436). Each cell in the grid measured 60x60 pixels, separated by

130 pixels from centre to centre. In the typical keypad display, digits were arranged as they

would be found on a telephone keypad. In the novel-changing display, digits were allotted to grid cells randomly on each trial. In the central display, each digit of the 7-item list appeared successively in the middle cell of the grid (i.e., where 5 would be in the typical keypad), with the remaining cells blank. The digits were presented in black font Droid Sans Mono with a size 18. Because the cells in the central display condition do not have contextual meaning, our gaze analyses that aim to determine whether looking benefited accuracy contrast the typical keypad and random displays only. We include responses from the central display condition in descriptive accuracy analyses to illustrate the expected boosts to performance with typical keypad displays and the expected detriments with the random displays in comparison with the central display. We also include an initial analysis of saccade amplitude across all three layout conditions to confirm that participants look around much less in the central display condition.

Each trial began by presenting 7 digits visually, one by one, each for 800ms with an interitem interval of 400ms after each digit, which contained an empty keypad grid without any cells highlighted for all conditions. To-be-remembered digits were highlighted successively by changing the colour of the background of their cells to blue (#0000FF) for the 800-ms presentation period. In the central condition, this meant that the middle cell was highlighted as each new digits appeared for its 800-ms presented period. After the final digit in the list was presented (including the subsequent 400-ms unhighlighted display), all numbers disappeared from the grid and the blank grid remained onscreen for 5000ms. Finally, participants saw a prompt to recall the digits orally in order and the researcher typed in each response spoken by the participant. No changes to the response were allowed, but participants were allowed to indicate if they did not know a position by saying "blank" instead of a digit. After completing each 8-trial block, the participant was offered a break.

Results

Data Analysis

We conducted Bayesian Analysis of Variance (Rouder et al., 2009, 2012) with the default settings of the BayesFactor package (Morey et al., 2022). This was implemented in R studio (R Core Team, 2013). The Bayes Factor (*BF*) is the relative likelihood of two models given the data, with models with highest *BF* value being preferred by the data and referred to as the best models. Null models contain only between-subject variance. By taking the ratio of the best model and other models that omit or include a given effect, we can compute evidence for or against including a predictor. For interpretation, we use the guidelines established by Lee & Wagenmakers (2014) where a *BF* of 3-10 represents moderate evidence in favour of the alternative hypothesis, and a *BF* of over 10 represents strong evidence. In addition to performance accuracy, we considered several eye movement measures. Saccade amplitude was taken as a broad indication of the amount of oculomotor activity, with higher amplitude indicating a greater distance from the start to the end of a saccade. The probability of fixating the interest areas of presented items and the probability of revisiting memorized interest areas were taken as indicators of where oculomotor activity was directed.

To investigate how looking behaviour and recall accuracy relate to each other, we employed mixed-effects regression models, which allowed us to account for the nested structure of our data, accommodate individual differences, and provide robust estimates of the relationship between these variables. These were conducted using the lmer function from the lme4 package (Bates et al., 2015). The coefficients (*β* values), standard error (*SE*) values, *t*-values, and *p*-values for each metric were reported, with the *p*-values being calculated using the lmerTest library for LMMs (Kuznetsova et al., 2017). For effects that did not reach significance ($p < .05$), only *t* and *p* values were reported without *β* or *SE* values, and these

effects were not emphasized in interpretations. Anonymized data and analysis scripts can be found at https://osf.io/3yath/[1]

Recall Accuracy

Figure 1 shows the recall accuracy as a function of layout (central, novel-changing keypad, and typical keypad) and serial position (1-7). For consistency with the gaze analyses reported below, which sometimes also include accuracy of recall as a factor, this inferential analysis contrasts the typical and novel-changing keypad conditions.[2] The best model of this data included the main effects of serial position and layout $BF = 1.59 \times 10^{48}$. The best model was preferred over a model including only the main effect of serial position by BF of 3.64, and over the full model including the main effects and their interaction by BF of 85.69. The best model was also favoured over a model including only the main effect of layout by a BF of 1.82×10^{48}. Although this boot-strapping advantage appears rather small, recall was improved in the typical compared to the novel-changing condition. Although the interaction between layout and serial position was not supported, the descriptive pattern appeared comparable to serial position functions previously shown (Allan, et al., 2017).

---

[1] Note that there is a basic preregistration document in our OSF repository but because our analysis plan was not specified in detail and we modified it after rounds of piloting and design adjustment, we do not consider this analysis pre-registered.

[2] We ran the same model including the central condition and the best model for the data included only a main effect of serial position.
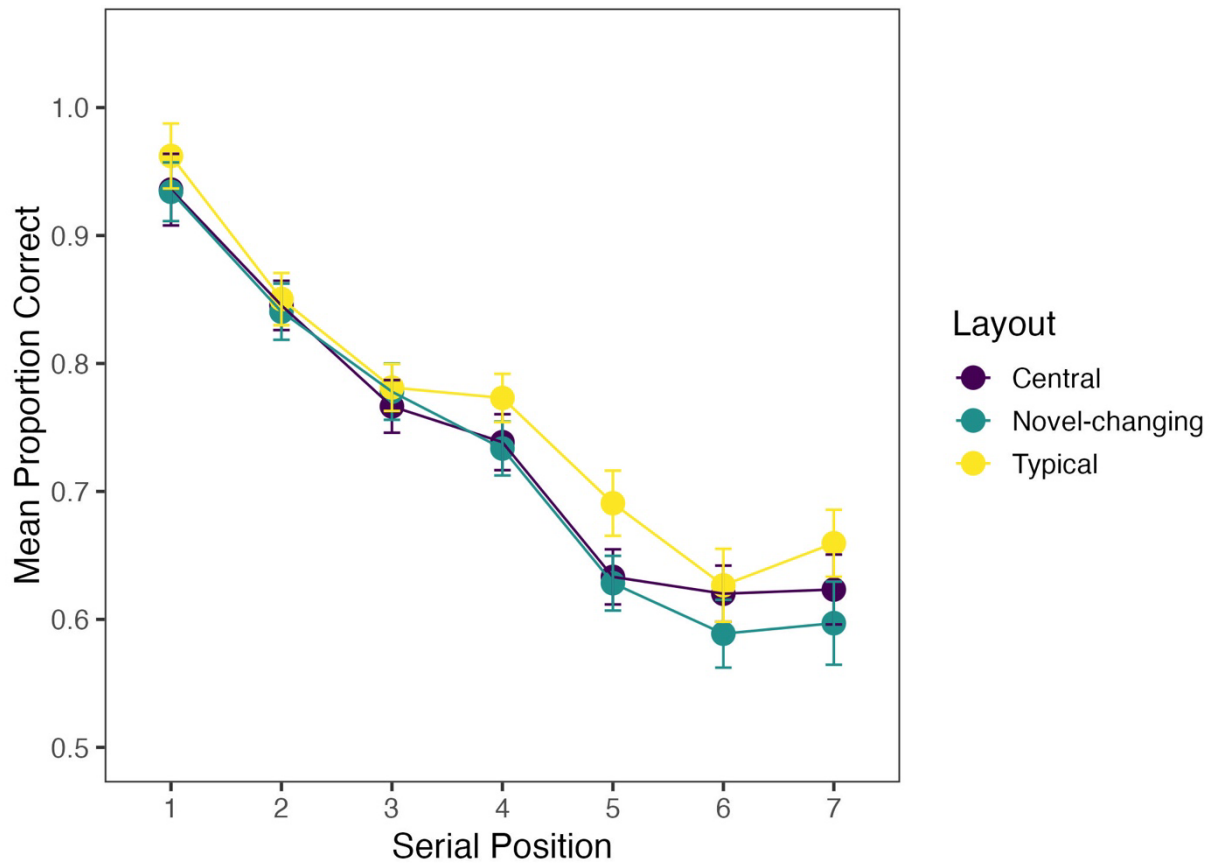
Figure 1. Mean proportion correct as a function of condition (central, novel-changing, typical) and serial position (1-7). Error bars mark within-participant standard errors around the mean calculated with the Cousineau-Morey method.

How much oculomotor activity occurred during presentation and retention?

Figure 2 shows the mean saccade amplitude as a function of layout (central, novel-changing, and typical) for the presentation and retention periods. We performed a factorial Bayesian ANOVA with layout and interest period and found that the best model included the main effects of layout and interest period, and their interaction ($BF = 3.19 \times 10^{416}$). This model was preferred over the next best model including only the main effects of layout and interest period by a BF of 3.22×10^{34}. As there is a clear pattern of higher saccade amplitude in the novel-changing and typical conditions, we further investigated the looking behaviour by including only the novel-changing and typical conditions. The best model included a main effect of layout ($BF = 45014.66$), providing evidence for higher saccade amplitude in the novel-changing condition compared to the typical condition. This model was only marginally

preferred over including a main effect of interest period by a *BF* of 1.34. The best model was

also favoured over the full model including both main effects of layout and interest period,
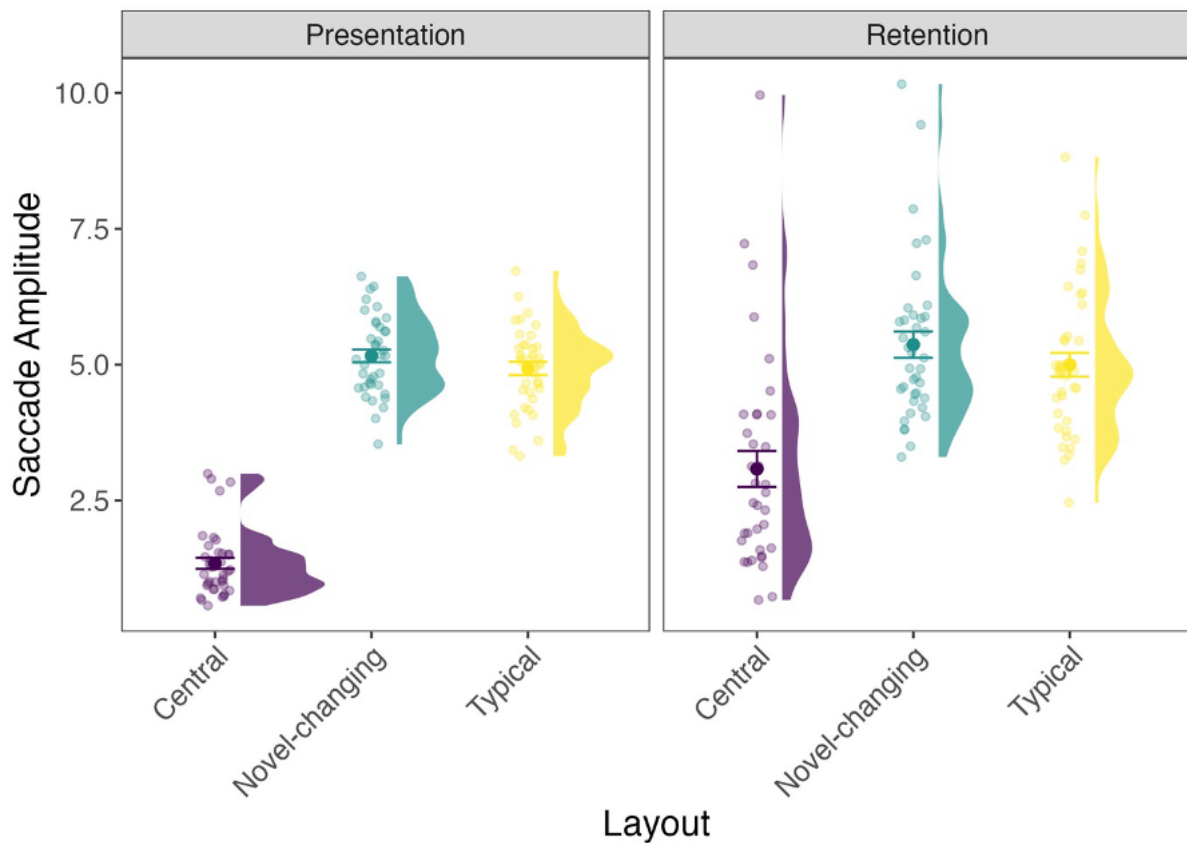
and their interaction (*BF* = 9.89).

Figure 2. Mean Saccade amplitude in visual degrees for each layout condition (central, novel-changing, typical) during the presentation and retention interest periods. Error bars show one standard error around the overall mean per condition. Points represent participant-specific means. Density curves along the continuous axis reveal data distribution and density across conditions.

What was fixated during encoding?

Figure 3 shows the average proportion of time participants fixated the incoming item's

interest area during its 1200-ms presentation period. Here the denominator was the sum of

fixation durations, which exclude time in blink or looking offscreen. It is also important to

note that the only relevant interest area for the central condition was the central position. In

any further inferential analysis we focus only on the novel-changing and typical layouts as

Figure 2 shows little looking around in the central condition and the meaning of looking

toward one of the blank cells in the central display condition would be unclear. Furthermore, Figure 3 shows that the fixation probability pattern is vastly different in the central condition compared to the conditions with a spatial array.

To explore the probability of looking at each currently presented item (interest areas were defined to match the size of each cell of the grid) we ran a two-way Bayesian ANOVA including layout (novel-changing and typical) and serial position (1-7). When considering proportions of current fixations during presentation to each item's location, the best model was the full model including the main effects of layout and serial position, as well as their interaction ($BF = 1.77$ x 10^{47}), which was decisively favoured over the next best model including only a main effect of serial position ($BF = 6.25$ x 10^{14}). The pattern shown in the left panel of Figure 3 suggests that the interaction must reflect a stronger impact of serial position in the typical layout than in the novel-changing layout. This decrease in looking directly at the currently presented item in the typical keypad condition is consistent with patterns previously observed when participant focus on encoding a series of spatial locations (Lange & Engbert, 2013).

To investigate if during the presentation of items participants revisit the locations of previously presented items, we calculated a cumulative measure which considered the sum of proportion of fixations to each of the previously presented items. There was strong evidence that the best model was the full model including the main effect of layout (novel-changing and typical) and serial position (1-7) and their interaction ($BF = 5.54$ x 10^{214}). The inclusion of the interaction in the model was strongly preferred ($BF = 9.66$ x 10^{16}). These results suggest that participants look back more frequently to previously-presented items in the typical keypad layout than with the novel-changing layout (Figure 3, right panel). A linear mixed-effects model was fitted to the data to examine the effects of the number of correct responses (0-7) and layout (typical keypad or novel dynamic) on the cumulative oculomotor

pattern. Random intercepts were specified for participants and trials to account for the nested structure of the data. Results showed a significant effect of layout ($b = -.121$, $SE = .008$, $t = -15.99$, $p < .001$), with a higher cumulative fixation probability in the typical keypad ($M = .66$, $SD = 0.17$) compared to the novel keypad condition ($M = .43$, $SD = 0.17$). However, there was no significant main effect of number of correct responses ($t = - .30$, $p = 0.76$), nor an interaction between the two main effects ($t = 1.16$, $p = 0.25$). This model was chosen on both theoretical and empirical considerations, with removing the fixed number of correct responses not leading to a significant change in model fit ($p = 0.498$). Although participants appear to look back toward the positions of previously presented items in both conditions, the extent of this behaviour does not vary with the extent of recall accuracy.
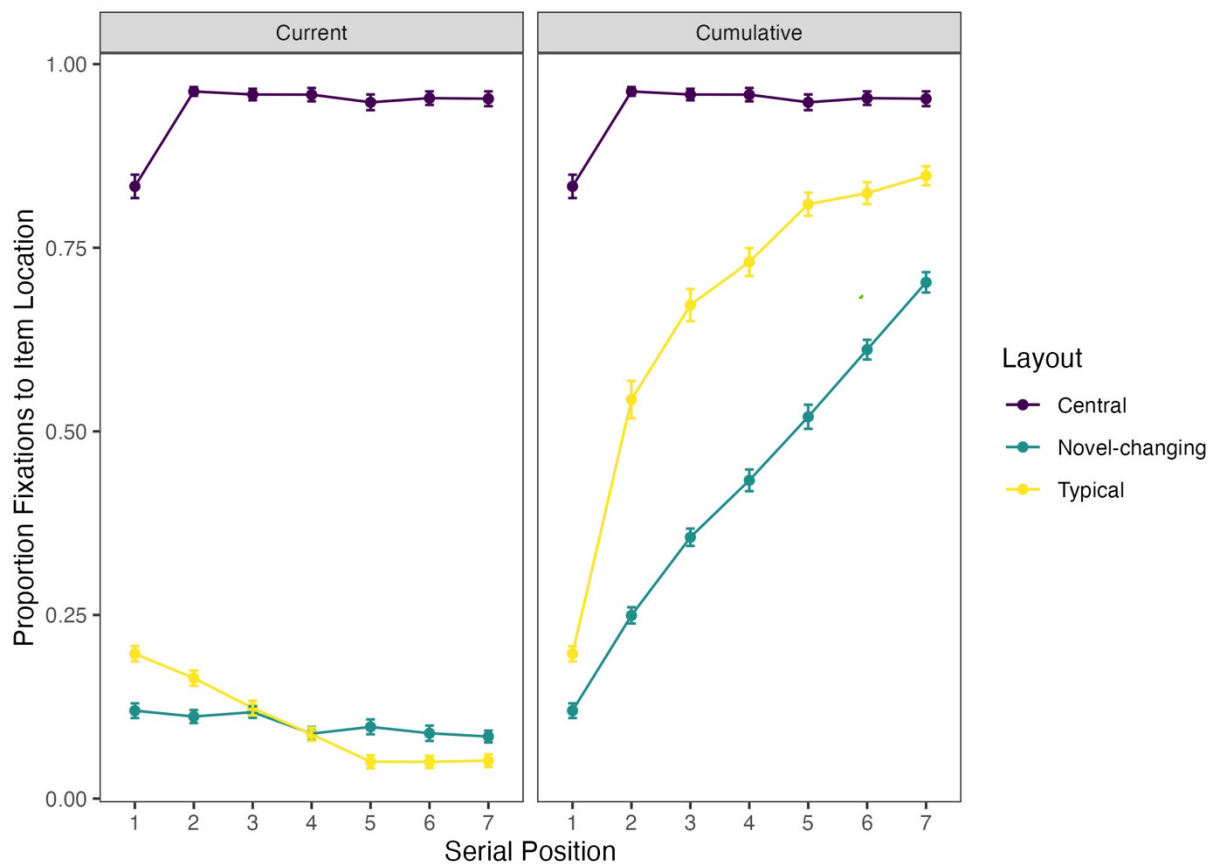

Figure 3. Mean proportion of fixations to current (left panel) and cumulative (right panel) item locations as a function of keypad condition (central, novel-changing, typical) and serial position (1-7). Error bars mark within-participant standard errors around the mean calculated

with the Cousineau-Morey method. Values for the Central condition are included for comparison only.

Did the number of fixated ordered pairs during retention influence accuracy?

The overall saccade amplitude data showed that participants tended to shift their gazes during the 5-second retention period as well as during encoding. We investigated gazes during retention in further detail to understand whether these gazes supported recall accuracy. Inspired by the analyses of Tremblay et al. (2006), we identified trials in which adjacent pairs within the 7-item sequence were fixated in the same order as originally presented, and compared accuracy of trials based on whether 0, 1, or more such ordered fixations were observed during the retention period (see Figure 4). If ordered fixations support memory, possibly because they reflect an oculomotor-based spatial rehearsal process, then retention periods that include more instances of ordered fixations should be recalled more accurately.

We ran an ANOVA including layout (novel-changing or typical), serial position (1-7), and number of fixated adjacent pairs (0, 1, 2+; more extensive ordered fixations were not so frequently observed to merit additional categories) to investigate if looking toward ordered pairs during retention provided any boost in accuracy. The best model included main effects of layout and serial position ($BF = 3.68 \times 10^{154}$), favoured over a model including the main effects of layout, serial position, and adjacent pairs fixated with a BF of 116.93. Including the main effect of serial position was strongly favoured ($BF = 4.75 \times 10^{152}$), and including the layout was preferred with a BF of 183.09. While this shows no evidence that paired looking improved performance, these findings provide further support for an accuracy benefit in the typical compared to novel-changing condition, even without evidence that this benefit was due to paired looking.
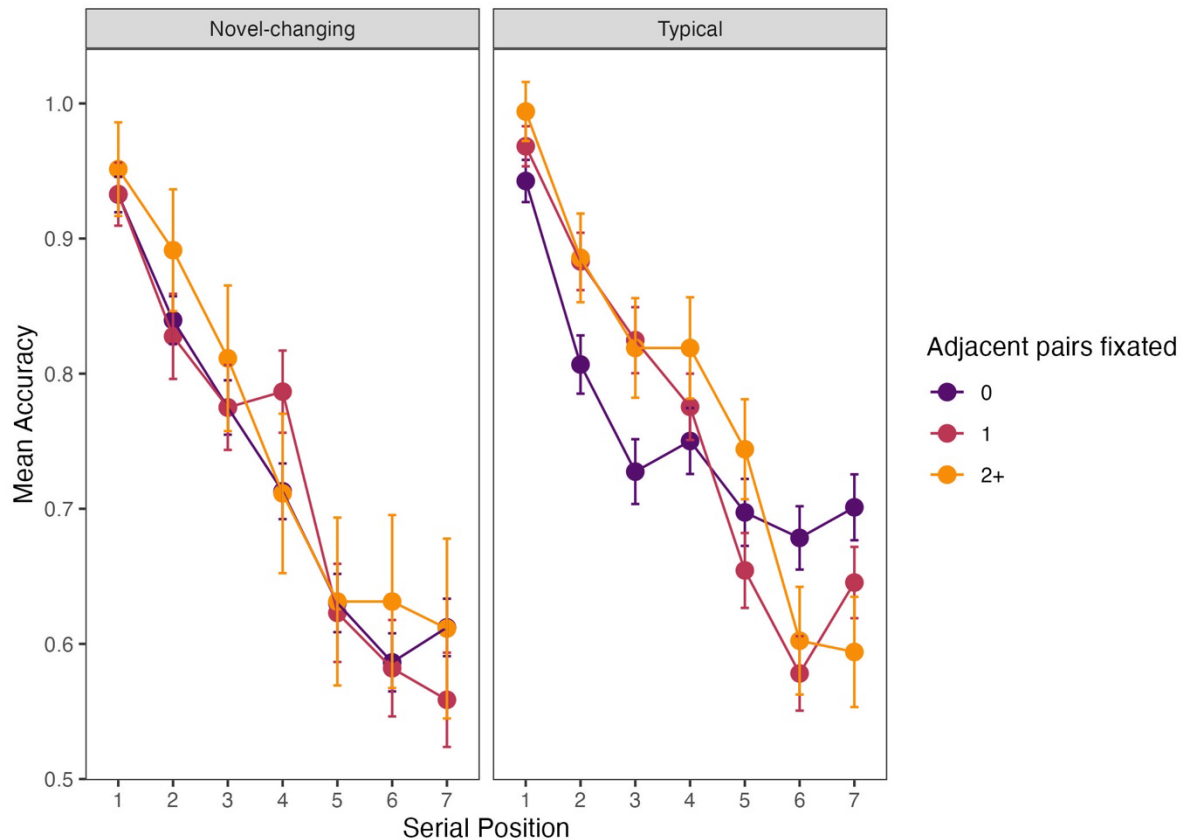
Figure 4. Mean accuracy as a function of serial position (1-7) and number of adjacent pairs fixated (0, 1, 2+). Error bars mark within-participant standard errors around the mean calculated with the Cousineau-Morey method.

Which pairs were fixated during retention, and did fixating particular pairs influence recall accuracy?

Although looking at more pairs did not reliably predict better overall recall accuracy, perhaps focusing on particular parts of the list benefited recall selectively. We compared the proportion of trials on which each possible ordered pair within the 7-item list was fixated (e.g., proportion on which participants fixated the location of the first-presented item and then the second-presented item, proportion on which they fixated the second and then the third, etc.), broken down further by display format and whether list recall was correct or not (Figure 5). The best model included the main effects of layout and serial position ($BF = 3212.15$), which was preferred over the next best model including only the main effect of layout with a BF of 2.83. Excluding the interaction between layout and serial position was favoured ($BF = 13.81$). Including the effect of layout was strongly favoured, suggesting that

more paired looking occurred in the typical keypad condition ($BF = 1415.13$). In the novel-changing layout, participants were less likely to fixate ordered pairs overall, with no obvious preference for fixating pairs from particular positions in the list. In the typical keypad condition, fixating pairs was more likely, particularly for early- and late-list pairs. More ordered looking did not lead to better overall recall accuracy ($BF = 9.02$).
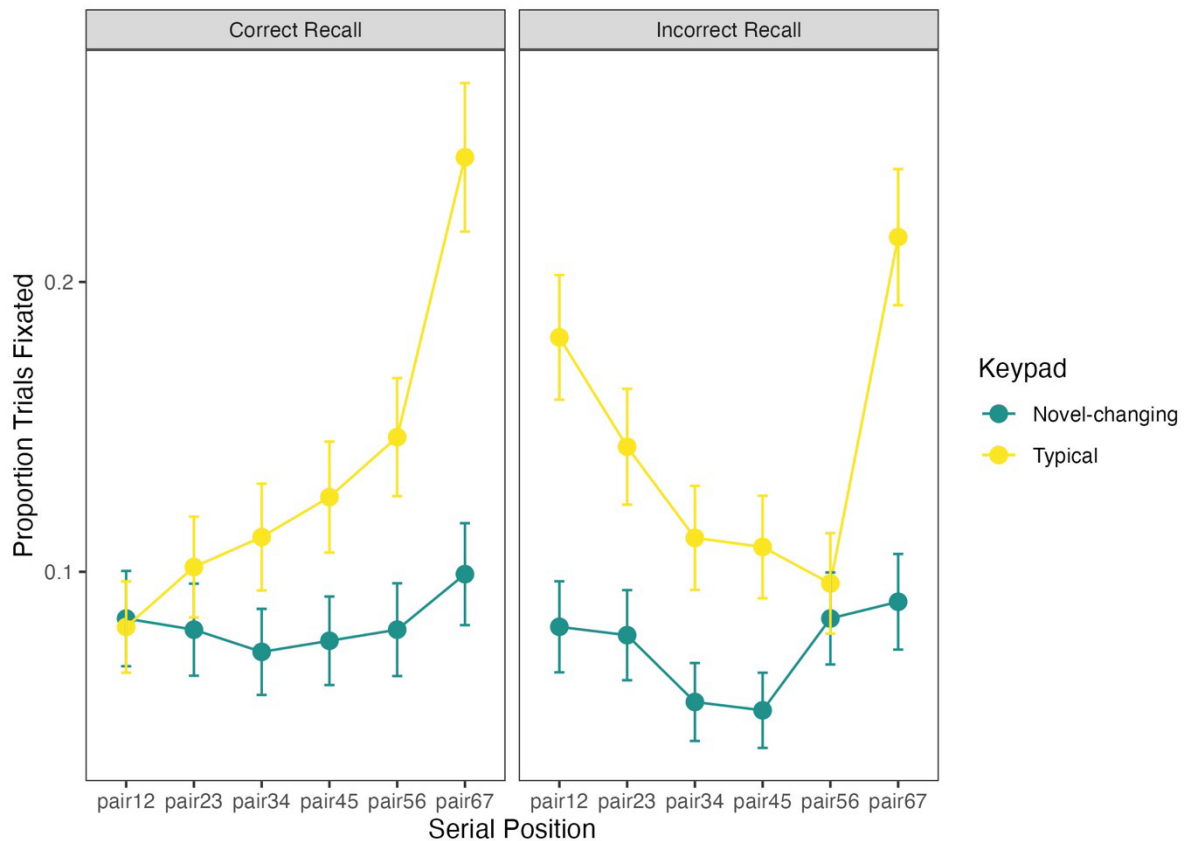
Figure 5. Mean proportion of trials as a function of fixated pairs and keypad condition during the retention interval. Error bars mark within-participant standard errors around the mean calculated with the Cousineau-Morey method.

Which serial positions were revisited during the retention interval?

Given previous findings that show stronger visuospatial boot-strapping benefits with immediate recall, it could be the case that gaze-based support for maintenance is rather short-lived, and therefore may be concentrated early during our retention interval. To investigate the time course of where oculomotor activity is directed throughout the retention interval, we considered the probability of fixating each of the relevant digits continuously across the 5-

second retention period. The fixation probability measure was based on all samples in

fixations and saccades and was calculated against all samples in each interest area in 20ms-

bins.

After a visual inspection of Figure 6 we focused our analysis on the first 1000ms of

the retention interval. We performed an ANOVA with layout and interest areas

corresponding to the seven serial positions during encoding as factors. The best model was

the full model including the main effects of layout, interest area, and their interaction ($BF =$

2.98×10^{35}). The best model was strongly favoured over the next best model which included

the main effect of interest area by a BF of 6.23×10^{12}. Including the interaction was favoured

by a BF of 7.35×10^{13}. The best model was also preferred over a model including only the

main effect of layout by 3.64×10^{36}. To further investigate if looks to the last presented digit

are driving the VSB effect, we excluded the fixation probability of the last digit from the

analysis and ran a 2 (novel-changing and typical) by 6 (interest areas of serial positions 1-6)

Bayesian ANOVA across the first 1000ms of their retention period. The best model for this

data included only the main effect of layout, but there was weak evidence for favouring it

over the null (which included only between-subject variance) with a BF = 0.40. Taken

together these findings suggest that any difference between the novel-changing and typical

keypad layouts is coming from an increase in the looks to the last presented digit in the

typical keypad condition. A linear mixed-effects model was fitted to the data to examine the

effects of the number of correct responses (0-7) and layout (typical keypad or novel dynamic)

on the probability of looks at the last digit during the first 1000ms of the retention period.

Random intercepts were specified for participants and trials to account for the nested

structure of the data. Results did not reveal a significant effect of layout ($t = -1.79$, $p = .07$),

or a significant effect of number of correct responses ($t = -.63$, $p = .53$). There was also no

interaction between these two main effects ($t = -.58$, $p = .56$). This model was chosen on both

theoretical and empirical considerations, with removing the fixed number of correct responses not leading to a significant change in model fit ($p = .68$).

We also investigated the link between accuracy for the last item and the probability of fixating the same item during the first 1000ms of the retention period. A linear mixed-effects model with the effects of accuracy for the last presented item (0 or 1) and layout (typical keypad or novel dynamic) on the probability of fixating the same item during the first 1000ms of the retention period. We also specified participants and trials as random intercepts to account for the nested structure of the data. Results did not reveal a significant effect of layout ($t = -1.32$, $p = .19$), or a significant effect of accuracy ($t = .75$, $p = .45$). There was also no interaction between the two main effect ($t = -.52$, $p = .61$).
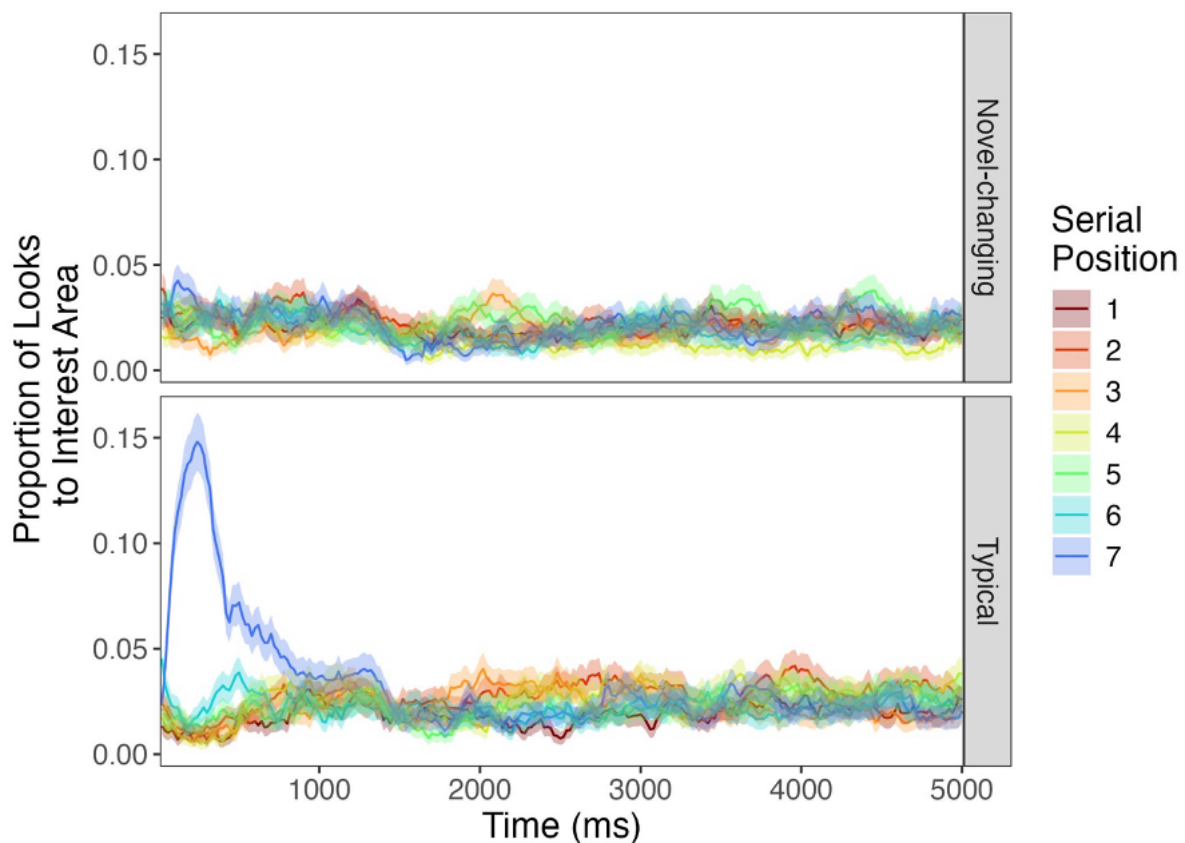

Figure 6. Time course of the proportion of looks to each digit interest area as a function of layout (novel-changing and typical). Shaded areas represent within-participant standard errors around the mean calculated with the Cousineau-Morey method.

Discussion

The present study aimed to explore gaze when encoding and maintaining a list of seven digits in a visuospatial bootstrapping task and found that the familiar keypad layout was associated with a unique gaze pattern. When encoding digits in the typical keypad layout, participants were more likely to look back toward previously presented locations than when digits were arranged in a novel layout. They were also more likely to fixate ordered pairs of digit locations during retention when digits had been presented in the typical layout, particularly towards the end of the list. However, neither of these tendencies directly predicted list recall accuracy, nor recall of the end-of-list items. These findings confirm that participants approach the typical layout lists differently than the novel or centrally presented lists and suggests that oculomotor activity may subtly distinguish these conditions. However, we did not observe direct evidence that the gazes themselves reinforce accurate memory.

The finding that in both the novel-changing and typical condition fixation probability of incoming items is low compared to the central condition is consistent with previous research which found patterns consistent with saccadic suppression during spatial compared to verbal memory encoding (Czoschke et al., 2019; Lange & Engbert, 2013; Patt et al., 2014). Taken together with our findings that lower saccade amplitude was found in the central condition compared to the conditions with a spatial array, this suggests that participants looked at the incoming items precisely in the central condition without much looking around. In the novel-changing and typical keypad conditions there was more oculomotor activity, and while the probability of fixating each incoming item was low, part of the reason for that were targeted looks back to previously presented items, shown by the cumulative increase in the proportion of fixations landing on relevant positions as the list progressed, which was especially pronounced in the typical keypad condition. This suggests that there is a unique

oculomotor behaviour when a spatial array is present, allowing the revisiting of locations associated with prior list items. In a task asking participants to either recall five bigrams, their unique locations, or the combination of both the bigrams and their locations, Czoschke et al. (2019) observed a pattern of lower fixation probability in the location recall condition compared to the verbal and combined conditions. However, they observed the opposite pattern when they considered looking back to previous items during the encoding phase, with lower regression probability in the verbal recall compared to the spatial and combined recall conditions. Even though our task was primarily verbal, the observed gaze patterns in the grid presentation conditions, especially where the mapping between digit and location was familiar, were consistent with those found by Czoschke et al. (2019) for spatial memory tasks. The correspondence between our gaze patterns in displays where visuospatial bootstrapping was possible and Czoschke et al.'s spatial recall conditions suggests that presenting verbal information in unique locations of a spatial array afforded participants opportunities to look back to previous items, which was most prevalent when the layout was already known. Therefore the gaze patterns we observed during visuospatial bootstrapping were consistent with those observed during spatial encoding, which is in keeping with the hypothesis that a partial spatial trace is "bootstrapped" onto to an otherwise verbal task, which may then boost recall. However, we found no direct evidence that this pattern was associated with a memory cost or benefit.

Investigating the retention period, we observed a higher proportion of trials where pairs of digits were fixated in the typical compared to novel-changing condition, with the biggest difference observed for the last presented pair. This could be consistent with use of gaze to rehearse a subset of the 7-item lists, which might boost individuals' span slightly in the typical keypad condition. But again, we did not observe a direct memory benefit associated with trials in which this pattern was observed. Further, we demonstrated that at the

beginning of the retention interval gaze is biased towards the last-presented item in the typical compared to the novel-changing condition. While it is possible that this only reflects lingering at the location of the last presented digit, the absence of the same bias in the novel-changing layout suggests that there is more to it than that: possibly, these gazes could reflect attempts to reinforce memory for end of list items, which might sometimes result in better performance. Though we did not observe direct evidence for that claim, we note that the visuospatial bootstrapping effect we observed in this study was rather small ($d = 0.14$) compared to other investigations (e.g. Allan et al., 2023: $d = 0.30$). Though we observed a statistically significant difference between the familiar and novel layouts consistent with many previous studies, this difference did not emerge in analyses including the intermediate centrally presented condition. This could reflect that the boost we observed in the typical keypad condition compared to central and likewise the cost from central to the novel-changing condition were rather modest and focused toward the end of list (though this is also consistent with previous findings; Allan et al., 2017). The overall low probability of revisiting previously presented item locations during the retention period is consistent with previous findings of low oculomotor activity during the retention period (e.g., Pearson and Sahraie, 2003, Experiment 5; Morey et al., 2018). However, the bias towards the last-presented item in the familiar layout condition lends support to Tremblay et al.'s (2006) argument that eye movements may play a role in rehearsing visuo-spatial information and to previous research suggesting participants may intentionally revisit locations associated with memorized items (e.g., Altmann, 2004; Richardson & Spivey, 2000; Ferreira et al., 2008; van Ede et al., 2019). Although previous research has associated this bias with verbal and visuospatial performance benefits (Hollingworth, 2009; Laeng et al., 2014; Laeng & Teodorescu, 2002; Scholz et al., 2016), our study did not provide clear evidence supporting such advantages.

Our data clearly show that eye movements differ when encoding verbal information that has a familiar layout compared with the same verbal information presented in a novel layout. These differences in gaze pattern could be taken as the recruitment of oculomotor infrastructure to apply a spatial rehearsal strategy, in which eye movements are used to re-activate the positions of previously viewed items, and this positional information can presumably only possibly benefit recall in the typical keypad condition where the positions map to learned digits. However, if we consider these gaze patterns evidence of visuo-spatial rehearsal, then we must conclude that gaze patterns are not impacting performance strongly because more looking was not correlated with better recall in any of our analyses. Possibly, looking back toward the position of a to-be-remembered item serves to attempt to trigger retrieval of unknown items and the stochasticity of this process hinders detecting a benefit. Nonetheless, observing different gaze patterns with the typical compared to novel keypad layout implies that the familiar mappings, which are the only thing differing between these scenarios, indeed change participants' approach to the task. Baddeley et al. (2021) suggest that a visuospatial component of the multiple-component working memory system ought to be considered more dependent on executive resources than the verbal component is believed to be, based on consistent findings that visuospatial memory is impacted more by dual-task interference than verbal memory (e.g., Morey, 2018; Morey et al., 2013; Morey & Mall, 2012). Our analysis of effects of gaze on recall in a task where spatial support should be beneficial is consistent with this idea. If domain-specific support from the oculomotor system is affecting recall, the impact is smaller than we can reliably detect and possibly quite short-lived. These severe limits on the extent to which eye movements might support spatial memory suggest that additional support would be required for successful maintenance in most cases. Regardless, the observation of different patterns of gaze for the typical layout are consistent with the view that this known spatial mapping is used to augment verbal recall,

even if the eye movements themselves do not directly support memory. Overall, this result seems most consistent with the assumption that domain-specific resources can be brought together to influence and augment immediate memory but offers little support for the idea that the oculomotor system supports robust and sustained serial spatial rehearsal that has measureable effects on recall accuracy.

Our investigation into visuospatial bootstrapping revealed unique patterns of eye movements during the encoding and retention phases providing further insight into the cognitive processes involved in processing spatially arrayed verbal items. Our findings demonstrated the occurrence of regressions to previously presented items during the encoding stage, suggesting a refinement of mental representations based on their spatial arrangement. During the retention interval, we observed a distinct increase in gaze towards the last presented item. This temporal bias in attention during the maintenance phase implies a dynamic engagement with the memorized spatial array, possibly indicative of privileged treatment of the later-presented items. Our findings of different eye movement patterns during verbal encoding, although with no direct links between those eye movements and recall, prompts a more nuanced consideration of the functional role of eye movements in the visuospatial bootstrapping process. Future research may benefit from exploring potential moderating factors that could elucidate the limitations and boundaries of the observed eye movement patterns. Additionally, an examination of individual differences in the susceptibility to the influence of eye movements on encoding and recall may contribute valuable insights, further refining our understanding of the variability in cognitive strategies that may augment recall in verbal tasks. The current study adds to the growing body of literature on visuospatial bootstrapping and sets the stage for more nuanced inquiries into the specific roles played by eye movements in the encoding and retrieval of spatially organized verbal information.

Declaration of conflicting interests
The Authors declare that there is no conflict of interest.

References

Allan, A., Morey, C. C., Darling, S., Allen, R. J., & Havelka, J. (2017). On the Right Track? Investigating the Effect of Path Characteristics on Visuospatial Bootstrapping in Verbal Serial Recall. *Journal of Cognition*, *1*(1), Article 1. https://doi.org/10.5334/joc.2

Allen, R. J., Havelka, J., Falcon, T., Evans, S., & Darling, S. (2015). Modality specificity and integration in working memory: Insights from visuospatial bootstrapping. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(3), 820–830. https://doi.org/10.1037/xlm0000058

Allen, R. J., Havelka, J., Morey, C. C., & Darling, S. (2023). Hanging on the telephone: Maintaining visuospatial bootstrapping over time in working memory. *Memory & Cognition*. https://doi.org/10.3758/s13421-023-01431-5

Altmann, G. T. M. (2004). Language-mediated eye movements in the absence of a visual world: The 'blank screen paradigm'. *Cognition*, *93*(2), B79–B87. https://doi.org/10.1016/j.cognition.2004.02.005

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

Baddeley, A., Hitch, G., & Allen, R. (2021). A Multicomponent Model of Working Memory. In A. Baddeley, G. Hitch, & R. Allen, *Working Memory* (pp. 10–43). Oxford University Press. https://doi.org/10.1093/oso/9780198842286.003.0002

Barrouillet, P., & Camos, V. (2021). The Time-Based Resource-Sharing Model of Working Memory. In P. Barrouillet & V. Camos, *Working Memory* (pp. 85–115). Oxford University Press. https://doi.org/10.1093/oso/9780198842286.003.0004

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Calia, C., Darling, S., Allen, R. J., & Havelka, J. (2015). Visuospatial bootstrapping: Aging
and the facilitation of verbal memory by spatial displays. *Archives of Scientific
Psychology*, *3*(1), 74. https://doi.org/10.1037/arc0000019

Cowan, N., Morey, C. C., & Naveh-Benjamin, M. (2021). An Embedded-Processes
Approach to Working Memory: How Is It Distinct From Other Approaches, and to
What Ends? In N. Cowan, C. C. Morey, & M. Naveh-Benjamin, *Working Memory*
(pp. 44–84). Oxford University Press.
https://doi.org/10.1093/oso/9780198842286.003.0003

Czoschke, S., Henschke, S., & Lange, E. B. (2019). On-item fixations during serial encoding
do not affect spatial working memory. *Attention, Perception, & Psychophysics*, *81*(8),
2766–2787. https://doi.org/10.3758/s13414-019-01786-5

Darling, S., Allen, R. J., & Havelka, J. (2017). Visuospatial Bootstrapping: When
Visuospatial and Verbal Memory Work Together. *Current Directions in
Psychological Science*, *26*(1), 3–9. https://doi.org/10.1177/0963721416665342

Darling, S., Allen, R. J., Havelka, J., Campbell, A., & Rattray, E. (2012). Visuospatial
bootstrapping: Long-term memory representations are necessary for implicit binding
of verbal and visuospatial working memory. *Psychonomic Bulletin & Review*, *19*(2),
258–263. https://doi.org/10.3758/s13423-011-0197-3

Darling, S., & Havelka, J. (2010). Visuospatial bootstrapping: Evidence for binding of verbal
and spatial information in working memory. *Quarterly Journal of Experimental
Psychology*, *63*(2), 239–245. https://doi.org/10.1080/17470210903348605

Darling, S., Havelka, J., Allen, R. J., Bunyan, E., & Flornes, L. (2020). Visuospatial
bootstrapping: Spatialized displays enhance digit and nonword sequence learning.
Annals of the New York Academy of Sciences, *1477*(1), 100–112.
https://doi.org/10.1111/nyas.14429

Darling, S., Parker, M.-J., Goodall, K. E., Havelka, J., & Allen, R. J. (2014). Visuospatial

   bootstrapping: Implicit binding of verbal working memory to visuospatial

   representations in children and adults. *Journal of Experimental Child Psychology*,

   119, 112–119. https://doi.org/10.1016/j.jecp.2013.10.004

Ferreira, F., Apel, J., & Henderson, J. M. (2008). Taking a new look at looking at nothing.

   Trends in Cognitive Sciences, *12*(11), 405–410.

   https://doi.org/10.1016/j.tics.2008.07.007

Hale, S., Myerson, J., Rhee, S. H., Weiss, C. S., & Abrams, R. A. (1996). Selective

   interference with the maintenance of location information in working memory.

   Neuropsychology, *10*(2), 228–240. https://doi.org/10.1037/0894-4105.10.2.228

Hollingworth, A. (2009). Two forms of scene memory guide visual search: Memory for

   scene context and memory for the binding of target object to scene location. *Visual

   Cognition*, *17*(1–2), 273–291. https://doi.org/10.1080/13506280802193367

Hoover, M. A., & Richardson, D. C. (2008). When facts go down the rabbit hole: Contrasting

   features and objecthood as indexes to memory. *Cognition*, *108*(2), 533–542.

   https://doi.org/10.1016/j.cognition.2008.02.011

Johansson, R., Holsanova, J., & Holmqvist, K. (2006). Pictures and Spoken Descriptions

   Elicit Similar Eye Movements During Mental Imagery, Both in Light and in

   Complete Darkness. *Cognitive Science*, *30*(6), 1053–1079.

   https://doi.org/10.1207/s15516709cog0000_86

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). **lmerTest** Package: Tests

   in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13).

   https://doi.org/10.18637/jss.v082.i13

Laeng, B., Bloem, I. M., D'Ascenzo, S., & Tommasi, L. (2014). Scrutinizing visual images:

The role of gaze in mental imagery and memory. *Cognition*, *131*(2), 263–283.

https://doi.org/10.1016/j.cognition.2014.01.003

Laeng, B., & Teodorescu, D. S. (2002). Eye scanpaths during visual imagery reenact those of

perception of the same visual scene. *Cognitive Science*, *26*(2), 207–231.

https://doi.org/10.1207/s15516709cog2602_3

Lange, E. B., & Engbert, R. (2013). Differentiating between Verbal and Spatial Encoding

using Eye-Movement Recordings. *Quarterly Journal of Experimental Psychology*,

66(9), 1840–1857. https://doi.org/10.1080/17470218.2013.772214

Lawrence, B. M., Myerson, J., & Abrams, R. A. (2004). Interference with spatial working

memory: An eye movement is more than a shift of attention. *Psychonomic Bulletin &*

Review, *11*(3), 488–494. https://doi.org/10.3758/BF03196600

Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian Cognitive Modeling: A Practical Course*

(1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139087759

Loaiza, V. M., & Souza, A. S. (2022). The eyes don't have it: Eye movements are unlikely to

reflect refreshing in working memory. *PLOS ONE*, *17*(7), e0271116.

https://doi.org/10.1371/journal.pone.0271116

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical

experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–

324. https://doi.org/10.3758/s13428-011-0168-7

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., & Mall, J. T. (2012). Cross-Domain Interference Costs during Concurrent

Verbal and Spatial Serial Memory Tasks are Asymmetric. *Quarterly Journal of*

Experimental Psychology, *65*(9), 1777–1797.

https://doi.org/10.1080/17470218.2012.668555

Morey, C. C., Morey, R. D., Van Der Reijden, M., & Holweg, M. (2013). Asymmetric cross-domain interference between two working memory tasks: Implications for models of working memory. *Journal of Memory and Language*, *69*(3), 324–348. https://doi.org/10.1016/j.jml.2013.04.004

Morey, R. D., Rouder, J. N., Jamil, T., Urbanek, S., Forner, K., & Ly, A. (2022). *BayesFactor: Computation of Bayes Factors for Common Designs* (0.9.12-4.4) [Computer software]. https://CRAN.R-project.org/package=BayesFactor

Patt, V. M., Thomas, M. L., Minassian, A., Geyer, M. A., Brown, G. G., & Perry, W. (2014). Disentangling working memory processes during spatial span assessment: A modeling analysis of preferred eye movement strategies. *Journal of Clinical and Experimental Neuropsychology*, *36*(2), 186–204. https://doi.org/10.1080/13803395.2013.877123

Pearson, D. G., Ball, K., & Smith, D. T. (2014). Oculomotor preparation as a rehearsal mechanism in spatial working memory. *Cognition*, *132*(3), 416–428. https://doi.org/10.1016/j.cognition.2014.05.006

Pearson, D., & Sahraie, A. (2003). Oculomotor Control and the Maintenance of Spatially and Temporally Distributed Events in Visuo-Spatial Working Memory. *The Quarterly Journal of Experimental Psychology Section A*, *56*(7), 1089–1111. https://doi.org/10.1080/02724980343000044

Postle, B. R., Idzikowski, C., Sala, S. D., Logie, R. H., & Baddeley, A. D. (2006). The selective disruption of spatial working memory by eye movements. *Quarterly Journal of Experimental Psychology*, *59*(1), 100–120. https://doi.org/10.1080/17470210500151410

Race, E., Palombo, D. J., Cadden, M., Burke, K., & Verfaellie, M. (2015). Memory
integration in amnesia: Prior knowledge supports verbal short-term memory.
Neuropsychologia, *70*, 272–280.
https://doi.org/10.1016/j.neuropsychologia.2015.02.004

Race, E., Tobin, H., & Verfaellie, M. (2023). Leveraging Prior Knowledge to Support Short-
term Memory: Exploring the Role of the Ventromedial Prefrontal Cortex. *Journal of
Cognitive Neuroscience*, *35*(4), 681–691. https://doi.org/10.1162/jocn_a_01965

Richardson, D. C., Altmann, G. T. M., Spivey, M. J., & Hoover, M. A. (2009). Much ado
about eye movements to nothing: A response to Ferreira et al.: Taking a new look at
looking at nothing. *Trends in Cognitive Sciences*, *13*(6), 235–236.
https://doi.org/10.1016/j.tics.2009.02.006

Richardson, D. C., & Spivey, M. J. (2000). Representation, space and Hollywood Squares:
Looking at things that aren't there anymore. *Cognition*, *76*(3), 269–295.
https://doi.org/10.1016/S0010-0277(00)00084-6

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes
factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374.
https://doi.org/10.1016/j.jmp.2012.08.001

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests
for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*,
16(2), 225–237. https://doi.org/10.3758/PBR.16.2.225

Scholz, A., Klichowicz, A., & Krems, J. F. (2018). *Covert shifts of attention can account for
the functional role of " eye movements to nothing "*. 230–243.
https://doi.org/10.3758/s13421-017-0760-x

Scholz, A., Mehlhorn, K., & Krems, J. F. (2016). Listen up, eye movements play a role in

verbal memory retrieval. *Psychological Research*, *80*(1), 149–158.

https://doi.org/10.1007/s00426-014-0639-4

Schut, M. J., Van der Stoep, N., Postma, A., & Van der Stigchel, S. (2017). The cost of

making an eye movement: A direct link between visual working memory and saccade

execution. *Journal of Vision*, *17*(6), 15. https://doi.org/10.1167/17.6.15

Spivey, M. J., & Geng, J. J. (2001). Oculomotor mechanisms activated by imagery and

memory: Eye movements to absent objects. *Psychological Research*, *65*(4), 235–241.

https://doi.org/10.1007/s004260100059

Staudte, M., & Altmann, G. T. M. (2017). Recalling what was where when seeing nothing

there. *Psychonomic Bulletin & Review*, *24*(2), 400–407.

https://doi.org/10.3758/s13423-016-1104-8

Team, R. C. (2013). R: A language and environment for statistical computing. R Foundation

for Statistical Computing, Vienna, Austria. *Http://Www. R-Project. Org/*.

Theeuwes, J., Olivers, C. N. L., & Chizk, C. L. (2005). Remembering a Location Makes the

Eyes Curve Away. *Psychological Science*, *16*(3), 196–199.

https://doi.org/10.1111/j.0956-7976.2005.00803.x

Theeuwes, J., Van der Stigchel, S., & Olivers, C. N. L. (2006). Spatial working memory and

inhibition of return. *Psychonomic Bulletin & Review*, *13*(4), 608–613.

https://doi.org/10.3758/BF03193970

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

van Ede, F., Chekroud, S. R., & Nobre, A. C. (2019). Human gaze tracks attentional focusing

in memorized visual space. *Nature Human Behaviour*, *3*(5), 462–470.

https://doi.org/10.1038/s41562-019-0549-y