Title
Using eye gaze data to examine the flexibility of resource allocation in visual working memory

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 41(0)

Authors
Stewart, Edmond
Donkin, Chris
Le Pelley, Mike

Publication Date
2019

Peer reviewed

Using eye gaze data to examine the flexibility of resource allocation in visual working memory

Edmond Stewart (e.stewart@unsw.edu.au)
Chris Donkin (christopher.donkin@gmail.com)
Mike Le Pelley (m.lepelley@unsw.edu.au)
School of Psychology, University of New South Wales

Abstract

Computational models of visual working memory (VWM) generally fall into two categories: slots-based models and resources-based models. Slots-based models theorise that the capacity of memory is defined by a finite number of items. Each slot can only contain one item and once an item is in memory it is remembered accurately. If an item is not in memory, however, there is no memory of the item at all. By contrast, resources-based models claim that all items, rather than just a few enter memory. However, unlike the slots model they are not necessarily remembered accurately. On the surface, these models appear to make distinct predictions. However, as these models have been developed and expanded to capture empirical data, they have begun to mimic each other. Further complicating matters, Donkin, Kary, Tahir and Taylor (2016) proposed that observers were capable of using either slot- or resource-based encoding strategies. In the current experiment, we aimed to test the claim that observers adapt their encoding strategies depending on the task environment by observing how participants move their eyes in a VWM experiment. We ran participants on a standard colour recall task (Zhang and Luck, 2008) while tracking their eye movements. All participants were asked to remember either 3 or 6 items in a given trial, and we manipulated whether the number of items was held constant for a block of trials, or varied randomly. We expected to see participants use more resource-like encoding when the number of items to remember was predictable. Contrary to these expectations, we observed no difference between blocked and unblocked conditions. Further, the eye gaze data was only very weakly related to behaviour in the task. We conclude that caution should be taken in interpreting eye gaze data in VWM experiments.

Keywords: visual working memory; eye gaze; hierarchical modelling

Introduction

In recent years, there have been a number of attempts to describe visual working memory (VWM) using computational models. These models attempt to address fundamental questions such as whether VWM has a strict capacity limit and how likely a stimulus is to be remembered. Broadly speaking, these models fall into two categories: slots-based and resources-based models.

Slots-based models propose that memory functions like a finite set of slots with each slot able to hold one item. The slots-based model proposed by Luck and Vogel (1997) is the prototypical account of this type. If an item is in a slot then it will be remembered. Critically, this account states that if an item is in a memory slot it will be remembered with a very high precision. If it is not in memory, no information is retained about the item. Therefore, if asked to recall an item that is not in memory, a slots-based account assumes that person – having no information about the item – will be forced to guess. Zhang and Luck (2008) expanded on this basic model to create their slots plus averaging model that makes the additional assumption that when the observer has more slots than items to remember, then items are stored in multiple slots. The information in multiple slots can then be combined to produce a more accurate response, thus leading to better performance when set sizes are small.

The resources-based model, on the other hand, conceptualises memory as being more flexible than does the slots model. Memory is described as a resource that is allocated to different items. This memory resource determines the quality of the memories. The more memory resource an item is allocated, the more precise the memory. According to the standard resources model (Frick, 1988) memory is divided equally between all items in the display. Since the amount of memory resource is constant, the more objects there are in a display, the less memory each item is allocated. Unlike the slots model, all items are remembered however, they are remembered with less accuracy as the number of items increases. Beyond the standard model, a variety of resources-based models exist that allow resources to be distributed more flexibly, or models that favour selecting a few items to focus most resources on (Alvarez & Cavanagh, 2004; Bays & Husain, 2008).

The mimicry problem

Zhang and Luck (2008) compared their slots-plus-averaging model to a slots model and a resources model and found that their model provided a better account of the data in a colour recall task. They concluded that the slots-plus-averaging model provided a favourable account of their VWM data. However, this model was challenged by Van den Berg et al. (2012), who developed the resources-based, variable-precision model. Unlike the standard resources model, the variable-precision model assumed that memory resources could be distributed unequally between items in memory. Van den Berg et al. (2012) compared the resources, slots-plus-averaging and variable-precision models and found that the variable-precision model had a better account of the data. Furthermore, in a large scale study, Van den Berg, Awh, and Ma (2014) tested a host of computational models of VWM against the variable-precision model. Using data from multiple experiments across multiple sites it was found that versions of the variable-precision model tended to provide the best account of the data.

Due to its flexible allocation of memory, the variable-precision model can produce many memory states that closely resemble what is predicted by a slot based model. For example, it is possible that in a 6 items display the mnemonic resource could be allocated equally between four items with no memory resource allocated to the other two. The result of such a memory state would be that four items are remembered with high precision and two are remembered with very low precision. If these very low precision memories are probed, the predictions of the model are indistinguishable from guessing. Such a memory state appears very much like a slots model. On the one hand, such overlap between models is problematic, due to model mimicry. Despite making the same predictions, the interpretations from the slots and resource models are very different. The variable-precision model states that "random" responses are caused by extremely low precision memories, while a slots model says that such responses are not based on memory. At our current level of understanding it is not possible to distinguish between these models. On the other hand, it could be that the mimicry between models represents what is actually shown in individuals. That is, perhaps observers do alternate between slot- and resource-like encoding of VWM displays.

A slot and resource model of encoding in VWM

A recent finding from Donkin, Kary, Tahir, and Taylor (2016) suggested that participants may be able to change their memory "strategy" in VWM tasks. Specifically, they argued that if people know how many items they will be required to remember, they are more likely to use a resource-like encoding, attempting to remember information about all items in the display, compared to if they don't know the set size of the next trial.

In their study, Donkin et al. (2016) analysed data from old and two new experiments with a model that used a mixture of slots-based and resources-based memory processes. The experiments were change detection experiments in which participants were tasked with recalling 2, 4, 6 or 8 items. In one experiment, set size varied from trial to trial, with an equal number of each size in each block of trials (the 'unblocked' condition). In the other experiment, set size was constant within each of four one-hour sessions (the 'blocked' condition). Compared to the experiments with unblocked set size, participants in the blocked experiment appeared more likely to use resource-like encoding (Figure 1). By contrast, participants in the unblocked condition were better accounted for by a slot-like encoding.

The authors suggest that VWM may be more flexibly applied than previously thought. It is possible that the task environment affects how people apply their memory. Perhaps if people know the number of items presented in a trial they will attempt to remember all items instead of focusing on a few. This would increase the chance of an item being in memory, but lower precision on blocked trials relative to unblocked trials and thus following a more resources-like pattern. While the behavioural data in a change detection task appear consistent with this suggestion (when analysed with these particular models), such a claim warrants more evidence to its support.

Here, we present data from a continuous production task in which participants were presented with items in either a blocked or unblocked conditions. To replicate the general results in Donkin et al. (2016), we expect that the number of items remembered should increase in the blocked compared to the unblocked condition (with a corresponding decrease in the precision of memory). As a further test of this prediction, we also use eye tracking to see whether eye movements differ between blocked and unblocked trials, thus suggesting endogenous attention is able to change the strategy used in VWM. We expect that participants in the blocked condition would move their eyes to more items in the display, presumably spending less time fixating on any given item.

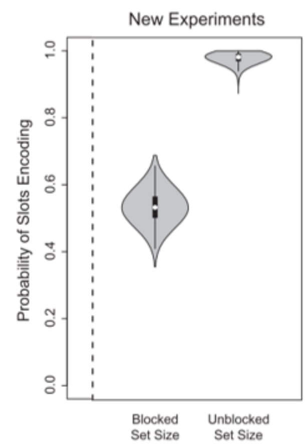

Figure 1: Results from Donkin et al. (2016) depicting the likelihood that participants used slots-like compared to resources-like encoding in the blocked and unblocked (new) experiments.

The current experiment

This experiment aimed to examine Donkin et al.'s (2016) claims that participants may be able to change their memory strategy depending on task environment. We wanted to determine whether 1) flexible memory allocation could be seen in a continuous report task, and 2) eye gaze could provide evidence of a change in memory strategy.

The task used was an adaptation of the colour recall task used by Zhang and Luck (2008), with the addition of eye tracking as well as a between-subjects condition of blocked or unblocked trials. The colour stimuli used in the standard production experiments are very simple to encode (Eng, Chen, & Jiang, 2005). As a result, a participant may be able to encode items quickly. We thought that more complex stimuli would encourage longer fixations and thus provide more data to assist our analysis. While more complex stimuli were desirable, it was also necessary to have stimuli that could be reproduced from a continuous range (the key benefit of colour stimuli). To this end, we used a "ring" set of stimuli. Shown in Figure 2a, these stimuli consisted of a coloured ring with a "bead" placed randomly on the ring's circumference. Participants were asked to place the bead on the ring as it appeared during study (Figure 2b). The stimuli were presented in set sizes of N = 3 or 6 for 1000ms.

By introducing eye gaze to this task, we hope to observe some differences in attention between our blocked and unblocked conditions. In tasks in which participants can move their eyes freely – such as a visual search task – there is little evidence that participants utilise peripheral attention (Findlay & Gilchrist, 2001; Rayner, 2009). As such, where participants fixate their gaze provides a proxy for what they are attending. If participants are able to change their memory strategy it seems likely that there would be differences in attention allocation as well. In order to see every item on a trial, a participant must move their eyes faster for a 6-item trial than for a 3-item trial. In the blocked condition the participant knows the set size of the next item. With this knowledge, it is possible that they prepare to move their eyes more quickly in the 6-item blocks. In the unblocked condition, participants are unsure of the set size on the next trial. While they might encode set sizes of 3 fairly easily, without additional preparation they may not able to see every item when the set size is 6.

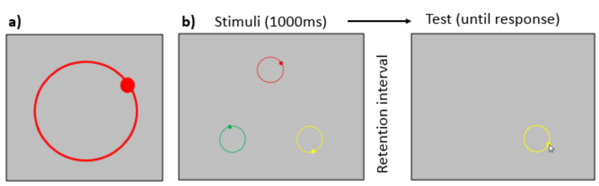

Figure 2: a) An example of the stimuli used. Stimuli varied in colour and location of the bead. b) The trial sequence. 3 or 6 stimuli were presented on a grey background for 1000ms followed by a retention interval (mask then blank screen) of 700ms. Participants were presented with a ring at the study location and were asked to place the bead on the ring as it appeared during study.

It was predicted that 1) similar to Donkin et al.'s (2016) results, we would find an increased probability resources-like encoding in the blocked condition of this experiment. This would be measured by higher chances on an item being in memory and lower precisions when compared to an unblocked condition. 2) We predicted we would see eye gaze data that supported more resources like encoding in the blocked condition. Specifically, more fixations but lower fixation durations compared to the unblocked condition.

Method

Participants 40 participants were recruited from the UNSW sign-up system SONA to complete a single one-hour session. Participants received $15 in exchange for participating.

Apparatus A Tobii TX300 eye-tracker, with 300 Hz temporal and 0.15° spatial resolution, mounted on a 23-inch widescreen monitor (1920 x 1080 resolution, refresh rate 60 Hz) was used. Participants' heads were positioned in a chinrest 60 cm from the screen.

Stimuli The stimuli (Figure 2a) were coloured rings with a filled circle placed on the ring's circumference. They could be one of eight distinct colours (red, yellow, green, cyan, blue, magenta, brown or salmon pink) and were presented on a grey background. All rings had a fixed diameter of 120 pixels (visual angle = 3.03°) and a thickness of 2 pixels (visual angle = 0.05°). Beads had a fixed diameter of 20 pixels (visual angle = 0.51°). Stimuli were presented randomly around the circumference of an invisible circle at the centre of the screen (diameter 600 pixels, visual angle = 15.1°). The angle between items was equal. Beads were randomly placed on the ring for each item, on each trial. The target was indicated by presenting just the ring of the stimulus in the location it had appeared in during study.

Design This recall task followed Zhang and Luck's (2008) design. 420 trials were divided into 14 blocks of 30 trials. Either N = 3 or 6 items were presented on each trial. In the unblocked condition, presentation was randomised per block with an equal number of 3 or 6 item displays per block. In the blocked condition, the first half of the experiment consisted of trials of all one set size and the second half consisted of only the other set size (counter balanced between participants).

Procedure A fixation cross was presented for 500ms at the start of each trial, followed by a blank screen for 400ms. The study array of N rings were presented for 1000ms. This was followed by a mask for 200ms then a blank screen for 500ms. The participant was then presented with a ring they saw at study (same location and colour) and was asked to place a bead on the ring where it appeared during the trial. The participant indicated where they believed the bead was with the mouse and confirmed their selection with the spacebar. Participants received feedback on their selection for 1000ms. Their deviation from the correct bead location was given in degrees alongside verbal feedback ("OUTSTANDING!" for deviations less than 10°, "Very good!" between 10° and 20°, "Good" between 20° and 35°, "OK" between 35° and 45° and deviations greater than 45° were labelled "Poor"). Figure 2b depicts the trial sequence. After each block, participants were given a break for a minimum of 20s before continuing.

Model procedure We used a model to allow us to compare the probability of an item being in memory (P_m) and the precision of memories (Prec) between the blocked and unblocked conditions. The model was a Bayesian hierarchical version of the Zhang and Luck (2008) mixture model (Oberauer, Stoneking, Wabersich, & Lin, 2017). The model assumed that the deviation between given response and the correct response either came from memory or from a separate guessing process. Responses based on memory were associated with Von Mises distributions with a mean that was centred on the correct response and a precision that varied depending on condition (blocked and unblocked), set size (3 and 6) and individual participant. Responses based on guessing were uniformly distributed around the circle for all conditions and all participants. The model allocated responses to either memory or guessing process by taking a value from a Bernoulli distribution with a probability of using memory equal to P_m. The parameters P_m, like Prec, also varied with condition, set size and individual participant. Thus, four values were estimated for each participant, P_m and Prec for set sizes 3 and 6 (remembering that blocked and unblocked conditions are between subjects). Rather than estimating parameters separately for each participant, we instead constrained individual-participant level parameters such that they came from their own population-level Normal distributions (i.e., one for each parameter in each set size and blocked/unblocked condition). We focus our analysis on the population-level posterior distributions of P_m and Prec across the four conditions of our experiment.

Results

Prior to analysis, trials with no eye gaze data collected were removed (544 trials or 3.24% of trial data). Trials with more than 10 fixations during the presentation window were also removed (964 trials, 5.74% of the data) as were trials where the average fixation duration was less than 125ms (2703 trials, 16.09% of the data).

Behavioural results For each trial, the deviation between the participant's answer and the true bead location was recorded. Since the range of answers varied around the circumference of the circle, the deviation was expressed in radians (π radians = 180 degrees). Figure 3 shows the frequency distribution of deviations for set sizes 3 and 6 (green and red lines respectively) for unblocked and blocked set sizes. Both conditions displayed the typical response pattern for this task (e.g. Zhang and Luck, 2008) with most responses clustered around the correct response for both set sizes but with more accurate responses for set size 3.

Modelling results Figure 4 shows plots of the population-level posterior distribution for P_m and Prec across condition and set size. There was no visible difference in P_m values for set size 3 between the blocked and unblocked conditions. There was a slight indication of a difference between the unblocked and blocked conditions for set size 6, with smaller P_m values in the blocked compared to the unblocked condition. Note that this pattern is the opposite of what we expected. Prec values appear to differ across set size, with higher precision in set size 3 compared to 6. However, there was no observable difference between the blocked and unblocked conditions.

The differences in Pm and Prec values between conditions for set size 6 only are presented in Figure 5. The difference between the posterior distributions for Prec centres on zero, suggesting no difference between conditions in precision for set size 6. The plot of P_m difference shows higher values for P_m in the unblocked condition compared to the blocked condition. However, this difference is small. Since an appreciable mass of the posterior distribution surrounds zero, there is little evidence of a difference between the conditions.

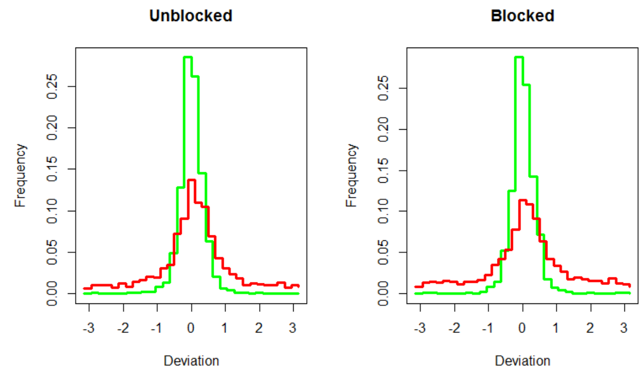

Figure 3: The frequency of responses by deviation from actual bead location for the unblocked and blocked conditions. The green line represents set size 3 the red line represents set size 6.

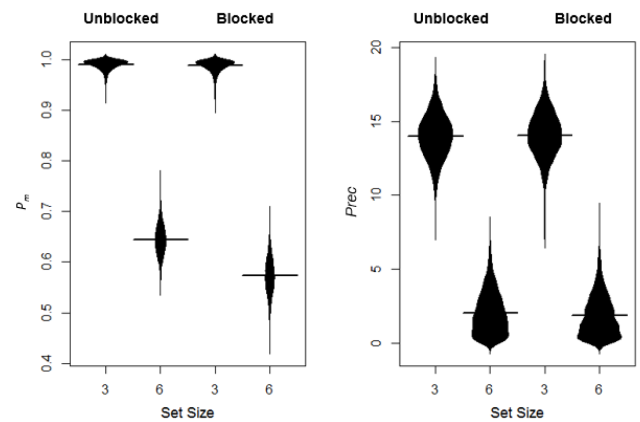

Figure 4: Posterior distribution for Pm and Prec parameters in blocked and unblocked conditions for both set size 3 and 6. Horizontal lines show the mean of the distributions.

Eye gaze results We now compare unblocked and blocked conditions using the average fixation duration per trial and the average number of fixations for each set size. The mean values of each measure in each condition are plotted in Figure 6. On average, the unblocked condition had more fixations and less fixation duration compared to the blocked condition. Again, the qualitative pattern, if present, is in the opposite direction of what was expected.

An analysis of variance (ANOVA) on average number of fixations and average fixation duration yielded no significant effect of condition ($F(1,37) = 1.441$, $p = 0.238$; $F(1,37) = 1.443$, $p = 0.237$ respectively) or set size ($F(1,37) = 0.337$, $p = 0.543$; $F(1,37) = 0.170$, $p = 0.682$ respectively) on either measure. We find no strong evidence for a difference between the average number of fixations or average fixation duration between trials of different set size or condition.

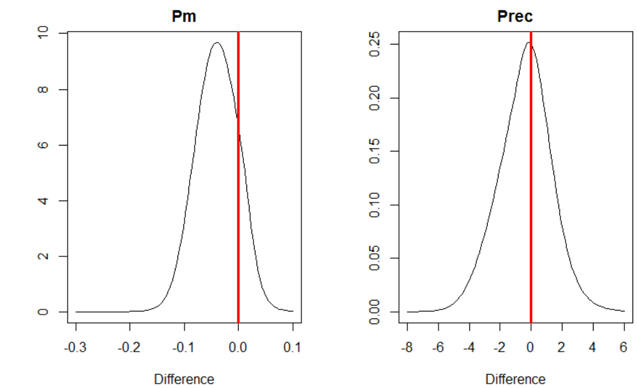

Figure 5: Difference in the posterior distributions of Pm and Prec between the unblocked and blocked conditions (set size 6 only).

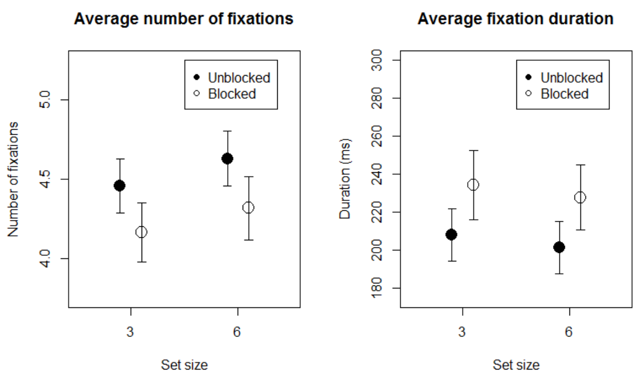

Figure 6: Average number of fixations and average fixation duration per trial for blocked and unblocked conditions for set sizes 3 and 6. Error bars indicate the standard deviations.

Discussion

The behavioural results for both the blocked and unblocked conditions were similar. The modelling results indicated very little difference in the memory strategies between conditions. For each condition, similar values for the probability of an item being in memory and for precision were found. Similarly, there was little difference in eye gaze patterns. There was a suggestion that there was a higher probability of an item being in memory in the unblocked condition. However, there were more fixations with lower durations in this condition as well. This trend is counter to our prediction, that the blocked condition would have a higher probability of items in memory, more fixations and lower average fixation duration.

Overall, these results are not consistent with what we expected based on Donkin et al.'s (2016) finding that memory can be flexibly allocated based on task environment. There were a number of differences in the experimental design between the Donkin et al. experiments and those reported here. The largest difference seems to be that here we used a continuous production task. It may be that participants are less able or willing to adapt their mnemonic resources in production tasks. On the face of it, production tasks require a more precise response than in a recognition/change detection task (in which there are only two responses). In the change detection experiments reported in Donkin et al., it was the blocked condition that was unlike previous experiments. It may have been that participants in our blocked condition did not spread their resources more diffusely in an attempt to remember more items because of the resultant cost to the precision of their memories. Future experiments could encourage participants to accept more error in their response, giving positive feedback whenever a response falls within a particular region around the correct response. Perhaps participants would adjust their mnemonic allocation in blocked conditions (where the number of items to remember is predictable) in such lenient environments. That said, such an explanation is obviously post-hoc, and so we do interpret this data as problematic for a model of VWM that proposes that mnemonic allocation is flexible and under strategic control.

In future work, we aim to connect the eye gaze data and the behaviour of individuals on individual trials. We have conducted preliminary analyses in which we see a weak correlation between fixation duration and the deviation between the correct response and the response given by the participant. We also see that whether an item was fixated during study is a weak predictor of deviation accuracy. These results were much weaker than we had anticipated, and so we will follow up these analyses with more refined methods. In particular, we will use summary statistics from eye gaze data as predictors for the parameters of the Zhang and Luck (2008) mixture model. For example, we might expect that an item not fixated during study would be more likely to come from a guessing process in the mixture model. We would also expect the fixation duration to affect the precision parameter of the memory process in the model. We have carried out versions of these analyses that we are not yet confident enough to report here, but were again very surprised by the lack of relationship between the eye gaze data and the behaviour of participants in the task.

Some of the flaws in the current design need to be addressed to convincingly link eye gaze and memory in this task. For example, one of the problems with the eye gaze data is perhaps that there is not enough distinction in where people are looking (their fixation locations) and their fixation durations. In this task, we suspect it is possible for participants to encode more than one stimulus in a single fixation as these relatively simple stimuli can be encoded quickly. We anticipate that either spatially separating items or more complex stimuli would therefore help distinguish which items a participant has looked at and thus attended and encoded.

Conclusions

Given participants did not move their eyes as much as anticipated, this seems to have impacted the collection of eye gaze information. In turn, the value of using eye gaze as our proxy for attention was thus diminished. As a result, we did not observe the difference in memory strategy between unblocked and blocked conditions as seen by Donkin and colleagues (2016).

Logically, vision must be helpful in encoding visual items into memory. The lack of a connection between memory and eye gaze in this study is likely due to methodological reasons. As mentioned, it might be necessary to make items more complex or make the display array more separated. However, to what extent alterations need to be made in order to observe an effect of eye gaze on memory remains to be seen. Future experiments could include gaze contingent presentations. Such a paradigm could require participants to fixate on a stimuli for a set period of time within a study array. As a result, there would be more certainty in what participants have looked at and perhaps encoded.

Presently, the current experiment serves as a caution to those interested in investigating VWM tasks using eye gaze.

Acknowledgments

This research was funded by the Australian Research Council (project number DP DP170101684)

References

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological science, 15*(2), 106-111.

Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science, 321*(5890), 851-854.

Donkin, C., Kary, A., Tahir, F., & Taylor, R. (2016). Resources masquerading as slots: Flexible allocation of visual working memory. *Cognitive psychology, 85*, 30-42.

Eng, H. Y., Chen, D., & Jiang, Y. (2005). Visual working memory for simple and complex visual stimuli. *Psychonomic Bulletin & Review, 12*(6), 1127-1133.

Findlay, J. M., & Gilchrist, I. D. (2001). Visual attention: The active vision perspective. In *Vision and attention* (pp. 83-103): Springer.

Frick, R. W. (1988). Issues of representation and limited capacity in the visuospatial sketchpad. *British Journal of Psychology, 79*(3), 289-308.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*(6657), 279.

Oberauer, K., Stoneking, C., Wabersich, D., & Lin, H.-Y. (2017). Hierarchical Bayesian measurement models for continuous reproduction of visual features from working memory. *Journal of Vision, 17,* 1-27.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The quarterly journal of experimental psychology, 62*(8), 1457-1506.

Van den Berg, R., Awh, E., & Ma, W. J. (2014). Factorial comparison of working memory models. *Psychological Review, 121*(1), 124.

Van den Berg, R., Shin, H., Chou, W.-C., George, R., & Ma, W. J. (2012). Variability in encoding precision accounts for visual short-term memory limitations. *Proceedings of the National Academy of Sciences, 109*(22), 8780-8785.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature, 453*(7192), 233-235. doi:10.1038/nature06860