cortex 122 (2020) 108–114

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Special issue: Research report

Evidence for the world as an external memory: A trade-off between internal and external visual memory storage

Rosyl S. Somai[*], *Martijn J. Schut* and *Stefan Van der Stigchel*

Experimental Psychology, Helmholtz Institute, Utrecht University, Utrecht, the Netherlands

ARTICLE INFO

Article history:
Received 31 May 2018
Reviewed 25 September 2018
Revised 9 November 2018
Accepted 11 December 2018
Published online 3 January 2019

Keywords:
Visual working memory
Trade-off
Eye movements
Cost-efficiency
External memory

ABSTRACT

We use visual working memory (VWM) to maintain the visual features of objects in our world. Although the capacity of VWM is limited, it is unlikely that this limit will pose a problem in daily life, as visual information can be supplemented with input from our external visual world by using eye movements. In the current study, we influenced the trade-off between eye movements and VWM utilization by introducing a cost to a saccade. Higher costs were created by adding a delay in stimulus availability to a copying task. We show that increased saccade cost results in less saccades towards the model and an increased dwell time on the model. These results suggest a shift from making eye movements towards taxing internal VWM. Our findings reveal that the trade-off between executing eye-movements and building an internal representation of our world is based on an adaptive mechanism, governed by cost-efficiency.

1. Introduction

Introspectively, we experience a complete representation of our surroundings. Contrary to this introspection, however, representations of our visual surroundings are imprecise and severely limited, as internal representations are constrained by capacity limits (Gameiro, Kaspar, König, Nordholt, & König, 2017; Mack & Rock, 1998; Simons & Ambinder, 2005; Simons & Levin, 1997). The internal representation of our visual world is maintained in visuospatial working memory, which is divided into a spatial and a visual component. Whereas *spatial working memory* maintains relevant locations in the visual world, *visual working memory* (VWM) is responsible for maintaining the visual features of objects (Baddeley, 2000; Baddeley & Hitch, 1974). Although VWM, and consequently our internal representation, is limited (e.g., Luck & Vogel, 1997; Ma, Husain, & Bays, 2014), it is likely that this maximum capacity will generally not pose a problem in daily life. Instead of using the energy-consuming internal memory, visual information can easily be supplemented with input from our external visual world. As long as visual information is readily available in the external visual world, there is no need for a complete internal

---

[*] *Corresponding author*. Experimental Psychology, Helmholtz Institute, Utrecht University, Heidelberglaan 1, 3584 CS, Utrecht, the Netherlands.

E-mail address: rosylsomai@gmail.com (R.S. Somai).

https://doi.org/10.1016/j.cortex.2018.12.017

representation of the outside world, as the world can act as an external memory source.

Although the concept of the world as external memory has been proposed theoretically (O'Regan, 1992), there is currently little behavioral evidence. The most compelling evidence for the world acting as external memory originates from human gaze behavior. In one of the few studies on external memory usage, Ballard, Hayhoe, and Pelz (1995) used a copying task in which participants were instructed to remember, and copy a model consisting of an arrangement of many elements. The model, presented on one end of the screen, consisted of a pattern of colored blocks. Participants needed to copy the model by arranging a set of colored blocks in a workspace at the opposite end of the screen. Participants were instructed to copy the model as quickly and accurately as possible. Results show that participants made numerous eye movements between the workspace, model and blocks, often making one saccade per action performed. These results indicate that participants, when given little instructions, rely on the objects in the external world rather than storing a large quantity of information in VWM.

The study of Ballard et al. (1995) suggests that there is a lower cost associated with executing a saccade than using VWM, which makes executing saccades preferable over storing information in VWM. In that same study of Ballard et al. (1995), an additional experiment was performed in two participants in which the distance between the model and workspace was increased from 15° to 70°, which the authors interpreted as adding a cost to a saccade. Indeed, neural systems are often described as systems that actively weigh outcomes to minimize costs, whilst maximizing efficiency, as a trade-off (i.e., the free-energy principle; Friston, 2009, 2010; Sengupta, Stemmler, & Friston, 2013). In the unique combinations of experiments in a relatively natural setting in the study of Ballard et al. (1995) successfully observed a reluctance for the use of "expensive" memory and the possibility of adding cost to a saccade, indicating the presence of a trade-off between storing information in VWM and making saccades. In the current study, we want to explore the specifics of this trade-off by systematically varying the cost of a saccade by manipulating the availability of external visual information. If there is an adaptive trade-off between using the external visual world and VWM, the trade-off should be influenced by increasing the cost associated with using external information. To explore the influence of costs on the trade-off between using the external visual world and taxing VWM, we increased the cost of a saccade by increasing the amount of time between saccade offset and external availability of visual information. We hypothesize that observers use VWM more when external information accessed through eye-movements becomes less efficient to use. As a result of VWM usage, fewer saccades towards the model are required. Furthermore, we investigated the amount of time participants observed the model, based on the concept of *active vision* (Hayhoe, Bensinger, & Ballard, 1998), which shows the need for availability of sensory input to compensate for information missing in internal representations. Therefore, a shift towards VWM use would express itself in longer dwell times in order to gather enough sensory input to build a more complete internal representation than the representation needed without a delay in viewing.

2. Experiment 1

2.1. *Methods*

2.1.1. *Participants*

We tested 12 participants (three males, mean age = 27 years, SD = 8.06, age range = 19–44 years). The original effect was observed with 7 participants (Ballard et al., 1995), however to allow for a linear mixed model analysis, we decided to increase the number of participants to 12. All participants had normal or corrected-to-normal visual acuity. Participants were naïve to the purpose of the experiment. All participants signed an informed consent form after a short explanation of the procedure. Participants received monetary compensation of 10 euros for participation afterwards. The experiment was approved by the local Faculty Research Ethics Committee of Utrecht University. The data, scripts for both experiments and all analyses are registered and available online (Somai, 2018).

2.1.2. *Apparatus and stimuli*

The experiment was programmed in Python 2.7, using the PyGaze library for eye-tracking (Dalmaijer, Mathôt, & Van der Stigchel, 2014). Gaze data was collected with the Eyelink 1000 (SR Research Ltd., Canada). The left eye was recorded in all participants, at 1000 Hz. The task was displayed on an ASUS PG278 LCD monitor (27 inch, 60.1 by 34.0 cm) with a refresh rate set at 120 Hz and a spatial resolution of 2560 × 1440 pixels. Simulated viewing distance was set at 70 cm. The eye tracker data files were processed with Python 2.7, statistical analyses were performed in R 3.4.1 (R Development Core Team, 2008). The experiment was conducted in a dimly lit room. Participants rested their head in a desk-mounted chin- and head-rest and were seated approximately 70 cm from the computer screen. The stimuli consisted out of five geometric shapes and each shape had a different color filling, inspired by the original design from Ballard et al. (1995) in which colored blocks were used. Each shape was presented in a square of approximately 1.85° by 1.85°. The distance between the model grid and response grid made it impossible to discriminate elements within the model grid while fixating on the response grid. The large distance between the model area and the response area assures that crowding drastically impedes the recognition of detail of peripheral information (e.g., Strasburger, Rentschler & Jüttner, 2011). Therefore, observers were required to make a saccade to properly identify the elements of the model grid. The figures in the current study had three additional variants to decrease distinctiveness: mirrored vertically, mirrored horizontally and flipped 180°. The specific lay-out of the experiment and an overview of the experiment are shown in Fig. 1.

2.1.3. *Procedure*

The experiment consisted of one practice trial and three blocks of 35 trials. Participants were instructed to examine the grid filled with stimuli (the model grid) on the left, and reconstruct the model grid using the building blocks on the right and placing them in the empty grid on the right side (the response grid). Participants were told that there were no time limitations, and that it was not possible to correct after placing

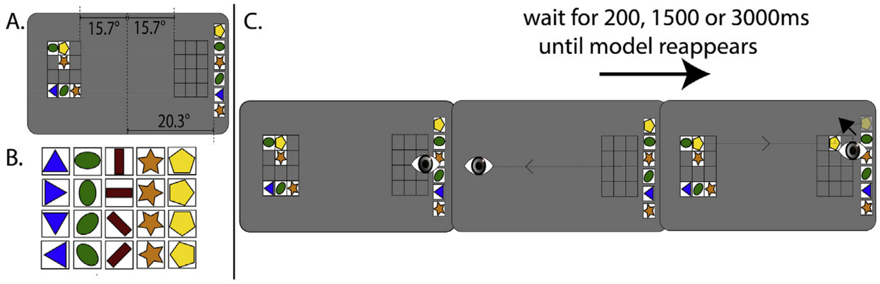

Fig. 1 – A. Overview of the experiment. Left is the pattern of blocks with a geometric shape, referred to as the model grid. This area is referred to as the model-area. On the right are the response grid and the building blocks. Orientation of the building blocks was correct and therefore only needed to be dragged into the correct position. B. Stimuli set. C. Overview of the task. At the start of the experiment the model grid was visible, but if an eye movement was made across the meridian of the screen, the model grid was occluded. The time during which the model grid was occluded is referred to as the Delay and is either 200 msec (baseline), 1500 msec or 3000 msec. After the Delay the model grid reappeared. The participant used the mouse to click and drag the building blocks into the response grid.

a figure in the response grid. Furthermore, participants were made aware that the model grid could be occluded and replaced by a fixation cross during an eye movement. Participants were instructed to continue fixating to make the model grid reappear. No further instructions concerning viewing behavior were given. After participants received full instructions on the screen, calibration took place.

Each trial started with a drift correction. When a trial started, the model grid was not occluded (Fig. 1b). In occluded state, a fixation cross was shown on the place where the center of the model grid would be. The model grid was occluded when a gaze sample was detected on the left side of the screen after one was detected on the right side of the screen. If gaze samples were detected on the left side of the screen for a preset amount of time (Delay; 200 msec, 1500 msec, or 3000 msec) the model grid would reappear. The Delay was set at one of the three values during an entire block of trials. The order of the block of trials was counterbalanced across participants. The task automatically continued to the feedback screen when all six items were placed in the response grid by clicking and dragging the building blocks. After each trial participants received feedback on the number of correctly placed items.

2.1.4. *Data analysis*

In total, 6 trials were excluded from further analysis due to inconsistencies in the raw data that suggest a faulty calibration on that specific trial. After exclusion 1266 trials remained for analysis (104 trials per participant on average).

We calculated the number of saccades towards the model and average dwell time per trial. We determined the number of saccades offline. In the study of Ballard et al. (1995) the actions between the response grid and items that can be selected were also measured, but in our study, we are solely interested in the effect of the availability of the visual input. Consequently, this means we could divide the display in two halves in which left is the visual input and right is the workspace. The number of saccades was defined as the number of times in which the gaze position crossed the midline from right to left. To only measure saccades that retrieved visual input and exclude inefficient saccades that returned to the response area without any new visual input, a saccade was only counted if the dwell duration within the model area surpassed the Delay. Next, we calculated average dwell time per trial. The average dwell time was defined as the number of samples on the left side of the screen (at 1 sample per msec) in which the model was not occluded. Effectively, this means the exclusion of datapoints with a gaze positioned on the left side of the screen during the occlusion. This was done to avoid influence from inefficient saccades, which are saccades that did not lead to visual input of the model. Lastly, we calculated the total duration it took to complete the task to explore to what trade-off the behavior is adaptive. We call this the completion time and this was calculated by subtracting the total time in which the model was occluded from the total time of the trial.

We expected participants to execute less saccades to the model as Delay increased as well as increased dwell time on the left (model) side of the screen as Delay increased. We also expect the trade-off between saccades and taxing VWM to be time-efficient, which would mean that completion time should not be affected by Delay. To test this, we constructed three linear mixed models, one model with dwell time as dependent variable, one model with number of saccades per trial as dependent variable and a last model with completion time as dependent variable. For each linear mixed model, we report the parameter estimate (β), standard error (SE), t value and p value. All linear mixed models included Delay as a categorical factor (200 msec, 1500 msec, or 3000 msec), and the intercept and slope in performance per participant as a random effect. We included Delay as a categorical factor, rather than a numerical factor, as to not assume a linear effect between the different Delays. To compare all Delay times with each other, we ran the linear mixed model twice, first we a Delay of 200 msec and subsequently with a Delay of 1500 msec as the reference. The threshold for significance was set at $\alpha = .05$. Lastly, the normality of the residuals of the random effects was visually inspected. All models showed no violations of normality of the residuals.

2.2. Results Experiment 1

In the current study, we investigated the effect of occlusion on gaze behavior, and the effect of stimulus availability on the use of working memory resources. We hypothesized that participants would rely more on an internal representation when external information becomes (increasingly) cumbersome to use. We expected that, in a matching task where the model and the to-be-matched area are seperated by a saccade, participants executed less saccades and dwelled on the model for longer if the model was occluded after each saccade.

2.2.1. *Number of saccades towards the model*

We constructed a linear mixed model to examine how the number of saccades towards the model area was affected by the participant and the time the model was occluded after a saccade (Delay). The results are shown in Fig. 2 (left panel). We find that participants executed significantly less saccades in trials where the Delay was 1500 msec, $\beta = -1.99$, $SE = .45$, $t = -4.40$, $p < .01$, and 3000 msec, $\beta = -2.33$, $SE = .44$, $t = -5.33$, $p < .001$, as compared to trials in which the Delay was 200 msec. Observers did not make less saccades towards the model in trials where the Delay was 3000 msec when compared to 1500 msec. Therefore, in line with our hypothesis, observers executed more saccades towards the model as the delay between saccade offset and presentation of the model decreased.

2.2.2. *Average dwell time*

Next, we constructed a linear mixed model to examine how average dwell time on the model side of the display was affected by Delay and the participant. Our results show that as Delay increased from 200 msec to 3000, $\beta = 3743$ msec, $SE = 774$ msec, $t = 4.84$, $p < .001$, participants dwelled longer on the model side of the display. Participants did not dwell significantly longer in the 3000 msec Delay condition as compared to the 1500 msec condition, although there appears to be a trend, $\beta = 3556$ msec, $SE = 1758$ msec, $t = 2.02$, $p < .1$. The results are shown in Fig. 2 (right panel). As hypothesized, we find that participants dwell time increased as the time between saccade landing and stimulus presentation increased.

2.2.3. *Average completion time*

Next, we constructed a linear mixed model to examine how average completion time on the model side of the display was affected by Delay and the participant. Our results show that as Delay increased from 200 msec to 3000, $\beta = 5576$ msec, $SE = 1021$ msec, $t = 5.46$, $p < .001$, participants took longer to complete the task. Participants did not need significantly more time to complete the task in the 3000 msec Delay condition as compared to the 1500 msec condition, although there appears to be a trend, $\beta = 4533$ msec, $SE = 2293$ msec, $t = 1.98$, $p < .1$. The results are shown in Fig. 2 (right panel). Contrary to the hypothesis, we find that participants completion time increased as the time between saccade landing and stimulus presentation increased.

2.3. Discussion Experiment 1

We hypothesized that the trade-off between VWM and gaze behavior is adaptive. Our results show that this time costs are reflected by an increased VWM usage: we observed a decrease in the number of saccades towards the model and an increase in average dwell time per trial in trials in which Delay time was increased. Moreover, the completion time also increased significantly in the longest Delay condition, which indicates that the extra time spend viewing the model does not lead to increased speed in the execution of the other parts of the task (such as the correct selection of the building blocks). Thus, it appears that the increased use of VWM instead of external visual information is not a time efficient adaptation of behavior. This finding is in line with the explanation of Ballard et al. (1995) stating that serialized eye movements are used to avoid memory costs.

At a first glance, it appears that the results of Experiment 1 are consistent with the hypothesis that there is a trade-off between VWM and gaze behavior. However, we used easily verbalized stimuli as memoranda. By using easy to verbalize stimuli, we allowed for the use of the phonological loop in which phonological information can be stored.

To explore whether the observed effects were really due to the retention of visual information, we performed a subsequent experiment with stimuli that were difficult to verbalize.

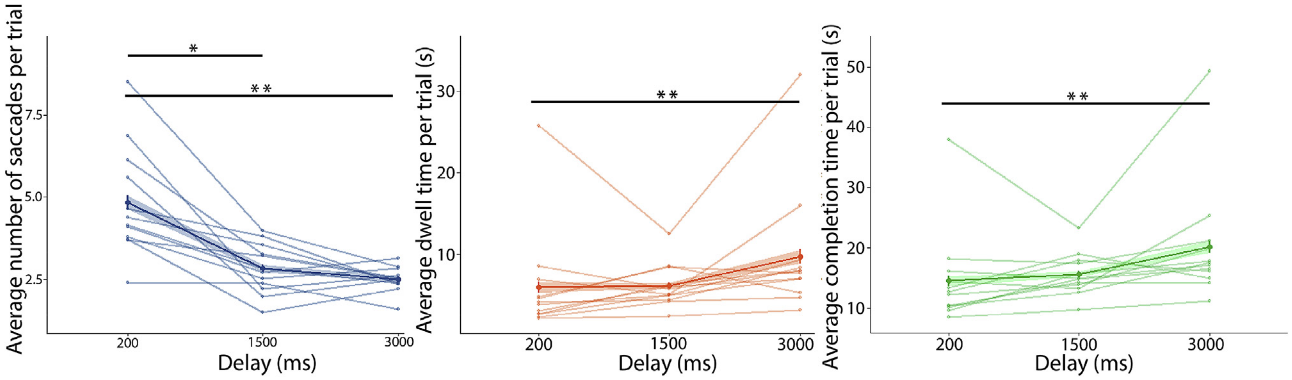

Fig. 2 – Linear mixed models for number of saccades per trial (left panel), average dwell time per trial (middle panel) and average completion time (right panel) over increasing delays. In both figures the open/transparent points represent participant averages, the closed points the linear mixed model averages. The shaded region and error bars show 95% confidence interval on the parameter estimates of the linear mixed model. **$p < .001$, *$p < .01$.

3. Experiment 2

This experiment is similar to the first experiment, but with nonsense shapes. Stepping away from easily verbalized stimuli will mean that our findings concerns the use of VWM, instead of other components of working memory.

3.1. *Methods*

The methods and data analysis of Experiment 2 were similar to the first experiment. The main difference was the stimuli that we used, which were adopted from a study on the familiarity and recognition of nonsense shapes by Arnoult (1956; see Fig. 3). We adopted these shapes in the current study based on the large amount of false recognition and recognition failures in the original study, indicating the difficulty participants experienced in remembering these stimuli. Again, the figures had three additional variants to decrease distinctiveness: mirrored vertically, mirrored horizontally and flipped 180°. We tested 12 participants (six males, mean age = 23 years, SD = 3.8, age range = 19–31 years).

3.1.1. *Data analysis*

Trials in which the next trial did not automatically start after completion were excluded from further analysis, as these trials may show a different pattern of gaze behavior. We also excluded trials in which the occluder may have not been presented correctly, due to inaccuracies inherent in online gaze detection. After exclusion, 1259 trials remained for analysis (13 trials removed in total, 104 trials per participant on average). The normality of the residuals of the random effects was visually inspected. All models showed no violations of normality of the residuals.

3.2. *Results Experiment 2*

3.2.1. *Number of saccades towards the model*

We constructed a linear mixed model to examine how the number of saccades towards the model area was affected by the participant and the time the model was occluded after a saccade (Delay). The results are shown in Fig. 4 (left panel). We observed that participants executed significantly less saccades in trials where the Delay was 1500 msec, $\beta = -2.44$, $SE = .48$, $t = -5.13$, $p < .001$, and 3000 msec, $\beta = -3.33$, $SE = .50$, $t = -6.67$, $p < .001$, as compared to trials in which the Delay was 200 msec. Observers also executed less saccades towards the model in trials where the Delay was 3000 msec when compared to 1500 msec, $\beta = -.89$, $SE = .25$, $t = -3.54$, $p < .01$. Therefore, in line with our hypothesis, observers executed more saccades towards the model as the delay between saccade offset and presentation of the model decreased.

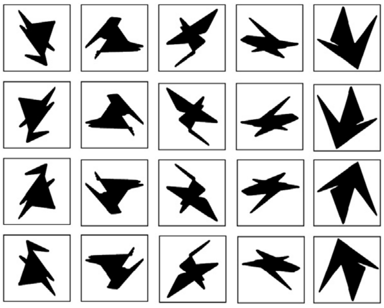

Fig. 3 – Stimulus set adopted from Arnoult (1956).

3.2.2. *Average dwell time*

Next, we constructed a linear mixed model to examine how average dwell time on the model side of the display was affected by Delay and the participant. Our results show that as Delay increased from 200 msec to 1500 msec, $\beta = 2404$ msec, $SE = 872$ msec, $t = 2.76$, $p < .05$, and 200 msec to 3000, $\beta = 4393$ msec, $SE = 856$ msec, $t = 5.13$, $p < .001$, participants dwelled longer on the model side of the display. Participants also dwelled significantly longer in the 3000 msec Delay condition as compared to the 1500 msec condition, $\beta = 1989$ msec, $SE = 731$ msec, $t = 2.72$, $p < .05$. The results are shown in Fig. 4 (middle panel). As hypothesized, we find that participants dwell time increased as the time between saccade landing and stimulus presentation increased.

3.2.3. *Average completion time*

Next, we constructed a linear mixed model to examine how average completion time on the model side of the display was affected by Delay and the participant. Our results show that as Delay increased from 200 msec to 1500 msec, $\beta = 4062$ msec, $SE = 1254$ msec, $t = 3.24$, $p < .01$, and 200 msec to 3000, $\beta = 6808$ msec, $SE = 1162$ msec, $t = 5.86$, $p < .001$, participants took longer to complete the task. Participants also needed significantly more time to complete the task in the 3000 msec Delay condition as compared to the 1500 msec condition, $\beta = 2746$ msec, $SE = 1125$ msec, $t = 2.44$, $p < .05$. The results are shown in Fig. 4 (right panel). Contrary to the hypothesis, we find that participants completion time increased as the time between saccade landing and stimulus presentation increased.

3.2.4. *Exploring individual strategy*

We performed a post-hoc correlation between the number of saccades towards the model and average dwell time per participant to explore the presence of different strategies in Experiment 2. We expected that dwell time would correlate negatively with the number of saccades, as longer fixations reflect stronger intake into VWM. Seven out of 12 participants show a significant correlation in the expected direction, and 2 showed the opposite direction. The correlation plots per participant can be found in the Supplementary materials. These results indicate that there are indeed individual differences, but that the majority of the participants follow a similar strategy. The strategy chosen by the majority of the participants shows that we make fewer saccades towards visual information that was previously associated with long fixations, in line with the idea that longer fixations reflect stronger intake into VWM.

4. General discussion

Here, we manipulated the external availability of a complex stimulus (model) by occluding the model for a short period after a saccade towards the model. By increasing the time

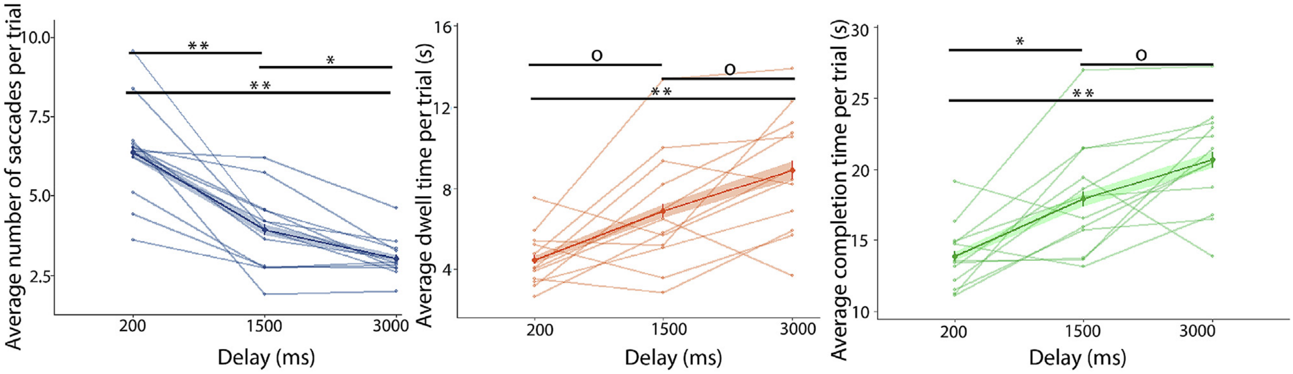

Fig. 4 – Linear mixed models for number of saccades per trial (left panel), average dwell time per trial (middle panel) and average completion time (right panel) over increasing delays. In both figures the open/transparent points represent participant averages, the closed points the linear mixed model averages. The shaded region and error bars show 95% confidence interval on the parameter estimates of the linear mixed model. $^{**}p < .001$, $^{*}p < .01$, $^{\circ}p < .05$.

between saccade offset and the availability of visual information, the execution of a saccade was associated with a time cost. Our results show that the increased costs are reflected in an increased VWM usage: we observed a decrease in the number of saccades towards the model and an increase in average dwell time per trial. This behavior is not a time-efficient adaptation, as the absolute time needed to complete the task (i.e., without the individual occlusion delays) became longer as the delay increased. This indicates that our findings are not the result of a simple strategy employed to become faster to complete the task, but can be explained by the participants' incentive to decrease the costs associated with the access of visual information. These findings provide further behavioral evidence for a trade-off between resampling visual information, as a cost-efficient memory buffer, versus storing visual information internally in VWM.

Prior research has indicated that observers show a preference for making additional saccades to sample information over storing multiple objects in VWM (Ballard et al., 1995). One possible mechanism for the preference of making additional saccades is that resampling information is more cost-efficient than storing information internally. Our results provide evidence that the trade-off between executing eye-movements and building an internal representation of our world is indeed based on an adaptive mechanism governed by cost-efficiency. Moreover, these results indicate that the default mode of executing many eye-movements is more cost-efficient than storing information in memory.

Our findings are in accordance with concepts such as *active vision* (Hayhoe et al., 1998), which take the sampling of visual information in order to avoid complex internal representations into account. The concept of *active vision* is directly linked to the idea that eye movements can be represented "explicitly as a variable in short term memory" (Hayhoe et al., 1998) indicating that eye movements and VWM are directly linked. There is now a growing body of evidence of this link (Van der Stigchel & Hollingworth, 2018), mainly coming from studies using visual search and corrective saccades. For example, it was found that the content of VWM partly determines target selection during visual search and also drives the correction of saccade endpoint errors (Hollingworth & Luck, 2009; Hollingworth, Richard, & Luck, 2008). Furthermore, executing saccade tasks and maintaining information in VWM are known to share the same resources (Schut, Van der Stoep, Postma, & Van der Stigchel, 2017). The role of eye movements in the functioning of VWM therefore goes beyond simply the resampling of visual information. Even when the external world does not provide the relevant visual information, studies still find that participants make eye movements to spatial location were the relevant visual information was previously (e.g., Richardson & Spivey, 2000). Ferreira, Apel, and Henderson (2008) are suggesting that 'looking at nothing' is not just a strategy, but it reflects the underlying representation of visual memoranda.

The idea that the external world plays an active role in cognitive processes has been mentioned several times (e.g., Clark & Chalmers, 1998), but more recently by Rowlands (2009) in his *extended mind thesis (EMT)*. Theories like the EMT suggest that there is more than a dependence on the external world, but that this dependency is an essential prerequisite for the functioning of a cognitive process. Our findings provide evidence for this theoretical framework using clear behavioral measures, by showing that the effect of a change in the external world – introducing a time delay – modulates the internal cognitive process of VWM utilization.

In conclusion, our study provides strong behavioral evidence for a trade-off between using the external world as a memory buffer, versus building complex internal representations. Prior research has shown that VWM and eye movements are directly linked and our results provide evidence that the link between executing eye-movements and building an internal representation of our world is based on an adaptive mechanism governed by cost-efficiency. Lastly, these results indicate that the default mode of executing many eye-movements is more cost-efficient than storing information in memory.

5. Author contributions

S. Van der Stigchel developed the study concept. All authors contributed to the study design. Programming of the task was done by R.S. Somai and M. J. Schut. Testing and data collection were performed by R. S. Somai. R. S. Somai and M. J. Schut performed the data analysis and interpretation under the supervision of S. Van der Stigchel. R. S. Somai drafted the manuscript, and M. J. Schut and S. Van der Stigchel provided critical revisions. All authors approved the final version of the manuscript for submission.

Acknowledgements

This research was supported by the Netherlands Organisation for Scientific Research [VIDI Grant awarded to Stefan Van der Stigchel; grant number 452-13-008]. The authors declare no competing financial interest.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cortex.2018.12.017.

REFERENCES

Arnoult, M. D. (1956). Familiarity and recognition of nonsense shapes. *Journal of Experimental Psychology, 51*(4), 269–276.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417–423.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. III, pp. 47–89). New York: Academic Press.

Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience, 7*(1), 66–80. https://doi.org/10.1162/jocn.1995.7.1.66.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, 7–19.

Dalmaijer, E. S., Mathôt, S., & Van der Stigchel, S. (2014). PyGaze: An open-source, cross-platform toolbox for minimal-effort programming of eyetracking experiments. *Behavior Research Methods*, 913–921. https://doi.org/10.3758/s13428-013-0422-2.

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.

Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends in Cognitive Sciences, 13*(7), 293–301. https://doi.org/10.1016/j.tics.2009.04.005.

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience, 11*(2), 127–138. https://doi.org/10.1038/nrn2787.

Gameiro, R. R., Kaspar, K., König, S. U., Nordholt, S., & König, P. (2017). Exploration and exploitation in natural viewing behavior. *Scientific Reports, 7*(1), 2311. https://doi.org/10.1038/s41598-017-02526-1.

Hayhoe, M. M., Bensinger, D. G., & Ballard, D. H. (1998). Task constraints in visual working memory. *Vision Research, 38*(1), 125–137. https://doi.org/10.1016/S0042-6989(97)00116-8.

Hollingworth, A., & Luck, S. J. (2009). The role of visual working memory (VWM) in the control of gaze during visual search. *Attention Perception and Psychophysics, 71*(4), 936–949. https://doi.org/10.3758/APP.

Hollingworth, A., Richard, A. M., & Luck, S. J. (2008). Understanding the function of visual short-term memory: Transsaccadic memory, object correspondence, and gaze correction. *Journal of Experimental Psychology General, 137*(1), 163–181. https://doi.org/10.1037/0096-3445.137.1.163. Understanding.

Luck, S. J., & Vogel, E. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*(6657), 279–281. https://doi.org/10.1038/36846.

Mack, A., & Rock, I. (1998). *Inattentional blindness*. Cambridge, MA: MIT Press.

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience, 17*(3), 347–356. https://doi.org/10.1038/nn.3655.

O'Regan, J. K. (1992). Solving the “real” mysteries of visual perception: the world as an outside memory. *Canadian Journal of Psychology, 46*(3), 461–488. https://doi.org/10.1037/h0084327.

R Development Core Team. (2008). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Richardson, D. C., & Spivey, M. J. (2000). Representation, space and Hollywood Squares: looking at things that aren't there anymore. *Canadian Journal of Psychology, 76*, 269–295.

Rowlands, M. (2009). The extended mind. *Zygon, 44*(3), 628–641.

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.

Sengupta, B., Stemmler, M. B., & Friston, K. J. (2013). Information and efficiency in the nervous system — A synthesis. *PLoS Computational Biology, 9*(7), 1–12. https://doi.org/10.1371/journal.pcbi.1003157.

Simons, D. J., & Ambinder, M. S. (2005). Change blindness: Theory and consequences. *Current Directions in Psychological Science, 14*(1), 44–48. https://doi.org/10.1111/j.0963-7214.2005.00332.x.

Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences, 1*(7), 261–267. https://doi.org/10.1016/S1364-6613(97)01080-2.

Somai, R. S. (2018). *P01: The world as external memory*. https://doi.org/10.17605/OSF.IO/rcd8S.

Strasburger, H., Rentschler, I., & Jüttner, M. (2011). Peripheral vision and pattern recognition: A review. *Journal of Vision, 11*(5), 13–13.

Van der Stigchel, S., & Hollingworth, A. (2018). Visuospatial working memory as a fundamental component of the eye movement system. *Current Directions in Psychological Science, 27*(2), 136–143.