UNIVERSITY OF GOTHENBURG

DEPARTMENT OF APPLIED IT

THE ROLE OF EYE GAZE IN VISUOSPATIAL MEMORY RECALL

- an eye-tracking study

David da Silva Jonasson

Marcus Tofftin

Thesis:	15 hp
Program:	Cognitive Science
Level:	Bachelor's thesis
Year:	2019
Supervisor:	Robert Lowe and Ben Clarke
Examiner:	Leonard Ngaosuvan
Report nr:	2019:064

Abstract

Research on the connection between eye movements and visuospatial memory has shown that eye movements have a functional role besides visual perception. Because of its predictable patterns during memory recollection, this phenomenon has been called "look-to-nothing", and it has been hypothesized that the eye gaze is used as an aid to memory. Although several studies have shown and argued for its effect, the evidence for which factors contribute to the mechanism has been inconclusive. In the present study, we investigate if different types of recollection methods induce more or less congruent eye movements, and if these eye movements can predict memory performance. A design of cued recall (simple questions) compared to semi-free recall (describing questions) was used, as participants were recalling information when looking at a blank screen. Complex naturalistic images were used as stimuli. In these conditions, correlational effects between memory performance and eye gaze congruence were also explored, in addition to participants' self-reported vividness rating of the different conditions. Our results provide novel evidence in this field of research, suggesting that eye movements show more congruence in a cued recall condition compared to free recall. The study found no correlation between memory performance and congruent eye movements.

Keywords

Eye movements, memory, visuospatial memory, attention, eye-tracking, mental imagery

Sammanfattning

Forskning om sambandet mellan ögonrörelser och visuospatialt minne har visat att ögonrörelser har en funktionell roll utöver visuell perception. På grund av dess förutsägbara mönster vid återkallning av minnen har detta fenomenet kallats "look-to-nothing", och det har hypotiserats att ögonrörelserna används som ett hjälpmedel till minnet. Även om flera studier har visat och argumenterat för dess effekt är bevisen för vilka faktorer som bidrar till mekanismen ofullständiga. I den här studien undersöks om olika typer av återkallningsmetoder inducerar mer eller mindre kongruenta ögonrörelser och om dessa ögonrörelser kan förutsäga minnesprestation. En design med cued recall (enkla frågor) och semi-free recall (beskrivande frågor) användes. Komplexa naturalistiska bilder användes som stimuli. Under dessa förhållanden undersöktes också korrelationseffekter mellan minnesprestation, kongruenta ögonrörelser och även deltagarnas självrapporterade tydlighet i visualisering under de olika förhållandena. Våra resultat ger nya bevis inom detta forskningsområde, vilka indikerar på att ögonrörelser är mer kongruenta i ett cued-recall-tillstånd jämfört med semi-free-recall. Studien fann ingen korrelation mellan minnesprestation och kongruenta ögonrörelser.

Nyckelord

Ögonrörelser, minne, visuospatialt minne, uppmärksamhet, eye-tracking, ögonspårning, mental imagery

Foreword

Bakgrundslitteratur är vald, läst och diskuterad tillsammans. Den tekniska utrustningen och dess dokumentation har vi tillsammans fått lära oss. Alla experiment var utförda tillsammans då båda två behövde vara närvarande vid experimenttillfället och vi hade en uppdelning i vem som gjorde vad under experimentet. Vi har spenderat nästan all tid tillsammans på universitetet och även om vi inte alltid har arbetat på exakt samma saker så har arbetsfördelningen varit jämlik och ingen av oss är missnöjd.

” You don’t want to become so open-minded that the wind can whistle between your ears”

-Terrence McKenna

Table of Contents

1. Introduction

Accumulating evidence in cognitive science suggests that humans engage in mental simulations for several cognitive processes (Kent & Lamberts, 2008). The evidence also shows that there is a strong relationship between retrieval of information and perception- that the retrieval mechanisms reactivate the processes that were active during encoding, not only on a neurological level but for sensory-motor processes as well. For visual stimuli, the experimental results have shown that there is a strong correlation between the oculomotor patterns executed during encoding and the patterns shown when recalling a stimulus and that these patterns can predict memory performance (Laeng & Teodorescu, 2002). This re-enactment of oculomotor patterns during recollection suggests that humans use their gaze as an aid for memory function as experiments have shown that forcing a participant to have a fixed gaze, both when recalling a stimulus and when encoding a stimulus, drastically impairs memory performance (Johansson, Holsanova, Dewhurst & Holmqvist, 2012). Additionally, findings in neuroscience have shown that mental imagery (the act of imagining a visual image) and visual perception engage in closely related neural mechanisms (Ganis, Thompson & Kosslyn, 2004).

Neisser (1967) has argued that the act of creating an internal mental image might be similar to the act of perceiving, in that humans need to move their gaze to complement the entire image. This would mean that mental imagery is a synchronized activity between the visuospatial memory and the eye movement patterns. Another account comes from Hebb (1968), who argued that if mental imagery is a simulation of the perceptual process then eye movements are a key part of the process and oculomotor function in recall is fundamental. Likewise, Mast and Kosslyn (2002) have argued that eye movements are stored as spatial indexes which are then used to organize parts of the image. They expand this by suggesting that eye movement may be used to activate sequences of memories and that eye movements could be stored in memory combined with the encoded image. This would indicate that the eye gaze plays an important role in mental imagery. A similar approach by Barsalou (1999) suggests that mental imagery is a simulation of the neural activity that was present when encoding information. This means that the mental imagery does not contain the *exact* information encoded but that sensory, motor and mental states are simulated.

On the other hand, it has been argued by Phylysyn (2000, 2001, 2002, 2003) that humans use their gaze as a spatial index to gather information from the external world and that no internal mental imagery exists. While Phylysyn is unconvinced of the phenomenon of mental imagery, more recent research opposes this view as Johansson, Holsanova and Holmqvist (2006) have shown that corresponding eye movements occur even in complete darkness, thus indicating some kind of internal processing of this information. The evidence for some kind of mental imagery and its correlation to eye movements has also been put forward in multiple experiments (Johansson et al., 2012; Johansson & Johansson, 2013) including for recollection in complete darkness (Johansson, et al., 2006) which indicates that eye movements have a role in memory recollection. These experiments were conducted with a recollection method in which the participant was to recall an image that was earlier encoded in their own words. The results showed congruent eye movements to spatial locations both when viewing an image, listening to a scene description, and for recall in darkness. These studies have also shown that when participants are to retell a scene that was earlier encoded, they move their gaze to a congruent position to describe an object *after* that object was mentioned (Johansson et al., 2006). Holsanova (2001) has suggested in earlier research that this may indicate that the gaze is used as a support to memory function for making an object more salient. This function or behavior has been called "look-to-nothing" and adds to the theorizing that the eyes have a function in addition to sensory visual input.

Many of the studies investigating the functional role of eye gaze on memory have used a method of manipulating the participants' gaze either when they are encoding information or when they are recalling information (e.g., Laeng & Teodorescu, 2002; Johansson et al., 2012). This meant one condition in which the participant was forced to have their gaze fixed when performing a task, and this was compared to a free gaze condition. The results showed that there is a hindering effect on memory recall when the gaze is fixed. Furthermore, the results indicated a corresponding relationship between the gaze pattern from encoding a stimulus and the pattern that is present when recalling a stimulus, looking at a blank screen. This led the researchers to the conclusion that the functional role of eye gaze is not for memory retrieval exclusively, but that it is a process of mental simulation when perceiving visuospatial scenes (Johansson et al., 2012). An opposing argument against the hypothesis for the functional role of eye gaze as a memory aid comes from the discussion of cognitive load (e.g., Johansson et al., 2012). Is keeping the eyes fixed either when encoding or recalling information attentionally taxing? Could this draining of additional cognitive resources be an explanation for why these results have been found?

Multiple studies with different approaches have addressed this issue. For example, Micic, Ehrlichman and Chen (2010) investigated the performance of the N-back task (a demanding cognitive task) and a verbal fluency task in a condition in which the participants maintained a fixed gaze and another condition in which the participants were free to use their gaze normally. They found no significant difference between the conditions. Another study by

Postle, Idzikowski, Della Salla, Logie and Baddeley (2006) addressed this issue by investigating the connection between visuospatial memory and eye gaze in experiments using both a condition of fixed gaze recall and a condition with free gaze recall. When manipulating the gaze of the participants, this study found no difference in the memory performance. Additionally, Johansson and Johansson (2013) conducted a memory experiment with eye gaze manipulation and designed it in a way that the eye movement constraints in the conditions were identical; they were to look inside a congruent or incongruent square when recalling, which was connected to encoded stimuli. Participants demonstrated a benefit of both accuracy and response time when looking inside a congruent square. The researchers concluded that the results found could not be merely due to cognitive load because of the identical task in both conditions, and that eye gaze does have a functional role in memory recall. On the other hand, recent research from Scholz, Klichowicz and Krems (2017) has shown that memory performance, when having a fixed gaze, may come from a covert shift of attention and they argue that eye movements may just be an aid for when the task is difficult. If only a few visual objects are to be memorized then a covert shift of attention may suffice, and eye movements would not be needed as an aid for retrieval.

Other conflicting arguments and results come from the investigation of eye movements and working memory, where studies have shown that eye movements disrupt spatial working memory (e.g., Theeuwes, Belopolsky & Olivers, 2009). They conclude that the process of eye movement and the process of spatial working memory are executed by the same attentional mechanism and when exercising them at the same time, working memory performance declines. Support for this argument in earlier research comes from studies that have reported that non-visuospatial tasks such as performing verbal tasks or solving arithmetic are more likely to induce eye movements than visuospatial tasks (e.g., Bergstrom & Hiscock, 1988).

Furthermore, earlier studies investigating the functional role of eye gaze have used relatively simple artificial stimuli for the participant to encode and later to recall (e.g., Laeng & Teodorescu, 2002). Andersson, Ferreira and Henderson (2011) have argued for the importance of using complex stimuli over simple artificial stimuli because of the difference in how visual information is processed in complex pictures compared to simple artificial ones. For instance, models of visuospatial memory suggest that the capacity for simple visual stimuli is limited to three or four items (Baddeley, Anderson & Eysenck, 2015), and therefore it is possible that using simple stimuli might not induce these eye movements because of the inherent capacity of short-term memory. Moreover, Johansson et al. (2012) have reported that "eye movements during mental imagery are less likely to appear for recall tasks that are relatively easy (e.g., questions about color, shape and location for single objects in a scene)" (p. 1290). A complex picture is also preferable in mental imagery research, both as Neisser has argued that a more vivid picture is more likely to induce mental imagery (Sheehan & Neisser, 1969), and for the ecological validity of the research. In addition to the measures of eye gaze patterns and memory recall, Johansson et al. (2006) used a measure for the participants' self-reported vividness of the mental imagery, called vividness rating. Although no correlation between memory and vividness rating was found in their study, the vividness rating is a widely used measure for subjective mental imagery (e.g., Cui, Jeter, Yang, Montague & Eagleman, 2007; Marks, 1973; Sheehan et al., 1969).

1.1. Theory

The somewhat conflicting findings and their interpretation from earlier studies led us to hypothesize that a key factor for progressing knowledge in the area is the investigation of mental imagery. We view it as a natural step in this area of research to investigate further using complex pictures as stimuli with different recollection methods to see what it is that induce these congruent eye movements and if they can predict memory performance.

We hypothesize that when a participant is to describe a scene or an object from memory using their own words (semi-free recall), this will induce more congruent eye movements than when they are answering simple questions (cued recall). This is because of the observation that cued recall often is easier than free recall and does not rely as heavily on retrieval strategies to recall items (Baddeley et al., 2015). We also hypothesize that there is a positive correlation between congruent eye movement to a recollected object and the memory performance. The other part of the inquiry is vividness rating, where we hypothesize that self-reported vividness is higher in the semi-free recall condition compared to the cued recall condition.

In the present study, we set out to investigate the eye movement congruency "to nothing" (recall on a blank screen) by using different recollection methods as an independent variable. The independent variable recall type will consist of two conditions, semi-free recall and cued recall. Our dependent variables measured are eye gaze congruency, memory performance and vividness rating (see Appendix A for operational definitions).

2. Method

2.2. Participants

Thirty-three students of Gothenburg university and individuals who signed up online (19 female, 14 male) participated in the study (mean age = 28.0 years, *SD* = 8.4). Prior written informed consent was obtained from all participants. All reported normal or corrected-to-normal vision and normal color vision.

2.3. Apparatus and stimuli

Stimuli were presented using PsychoPy3 (version 3.0.5) on a 550 x 490 mm monitor (resolution = 1680 x 1050 pixels). Eye movements were measured using Pupil Labs Mobile Eye Tracking Headset that recorded binocularly at 200 Hz. The data were recorded with Pupil Labs software using manual 17-point calibration (average measured accuracy = 0.48°, *SD* = 0.13°), 2D detection and mapping mode, and a fixation detection duration of 80.0 ms with a maximum dispersion of 3.00°. Fixations were detected using a dispersion-based algorithm. Visual stimuli consisted of two pictures (see appendix B). Both were naturalistic and complex pictures with a similar number of salient objects selected online, using a definition from Johansson et al. (2012). The pictures were chosen to be novel and unknown to avoid prior knowledge and associations; avoiding famous monuments, figures and pop culture. The stimuli also needed to contain an approximately equal number of salient objects in all four quadrants of the picture to support our surfaces. Standardized image sets such as Bank of Standardized Stimuli (BOSS) did not offer pictures that conformed to our criteria and thus images were chosen from Google Images. Questions were presented verbally from an experiment leader (see Appendix C for experiment questions).

2.4. Design and procedure

Each participant did both conditions once and balancing of conditions was achieved using all possible orders of conditions and stimuli, and randomized order of question blocks (amounts to 24 orders), making our design an incomplete repeated measures design. Data were obtained in two runs, each of which comprised an encoding phase and a recall phase. The recall phase consisted of two conditions, a semi-free recall condition and a cued recall condition. The semi-free recall condition consisted of four questions in which the participant was asked to describe an area or an object in the picture in as much detail as possible and tell the experiment leader when he/she was done. The cued recall condition consisted of 17 questions. Questions in both conditions were designed in such a way as to not give any indication about direction (e.g., not "to the left in the picture"). For participants who were not comfortable in English, we made a replica of our experiment in Swedish, as we learned from piloting that some participants reported that they were "thinking in" Swedish and therefore that made the recollection phase more challenging. To design for comparable conditions, piloting gave us the data to equate the number of questions used in the cued recall condition, and the number of questions used in the semi-free recall condition. Looking at the number of fixations the average participant executed per condition, we calibrated the questions to have an approximately equal number of fixations in both conditions, and for both stimuli. To measure the number of correct answers in the semi-free recall condition we needed a method to transcribe and analyze the answer into measurable units. As the participant was asked to freely describe a certain part of the picture our analysis needed to decide whether their expressed descriptions were correct according to the picture or not. This was done by separating each detail of the participant's description into a measurable unit. The transcripts were segmented as in the examples that follow:

"The woman had *dark hair* and *a red shirt.*"

"The sign had *white text* with *a green background* and said *Tennessee.*"

This would be counted as two and three memory units, respectively.

In addition to measuring memory performance, we used a technique to measure the proportions of fixations on a congruent surface area. This was to measure whether our conditions resulted in a difference in the use of eye gaze to aid recollection. To accomplish this, the screen was split into four equally large surfaces on which we could measure gaze fixations. If the participant fixated their gaze on the "correct" (congruent) area during recall on the white screen it was counted as a congruent fixation. This was used in comparison to incongruent fixations during the same frame interval (question onset to answer offset), creating a percentage of congruent fixations. Participants were told in the introduction that we were investigating pupil dilation in relation to mental workload, and thus were not primed on what to do with their gaze (see appendix D for pre-experiment instructions). The participant was seated 80 cm from the monitor and answered the questions in both conditions verbally. The experiment leader was positioned slightly behind the participant, out of view.

The experiment started with an instruction text on the screen in which the participant was told to study the picture as thoroughly as possible and to keep their eyes pointed towards the monitor, after which the participant was to prompt the experiment leader when he/she was ready. In the encoding phase, the participant studied a picture for 30 seconds, after which the screen turned white and the participant was given questions about the picture. There was no time limit in the recall phase (see Fig 1. For experiment setup). After both recall conditions the participant filled out a survey with the vividness rating for each condition (see Appendix E for the survey).

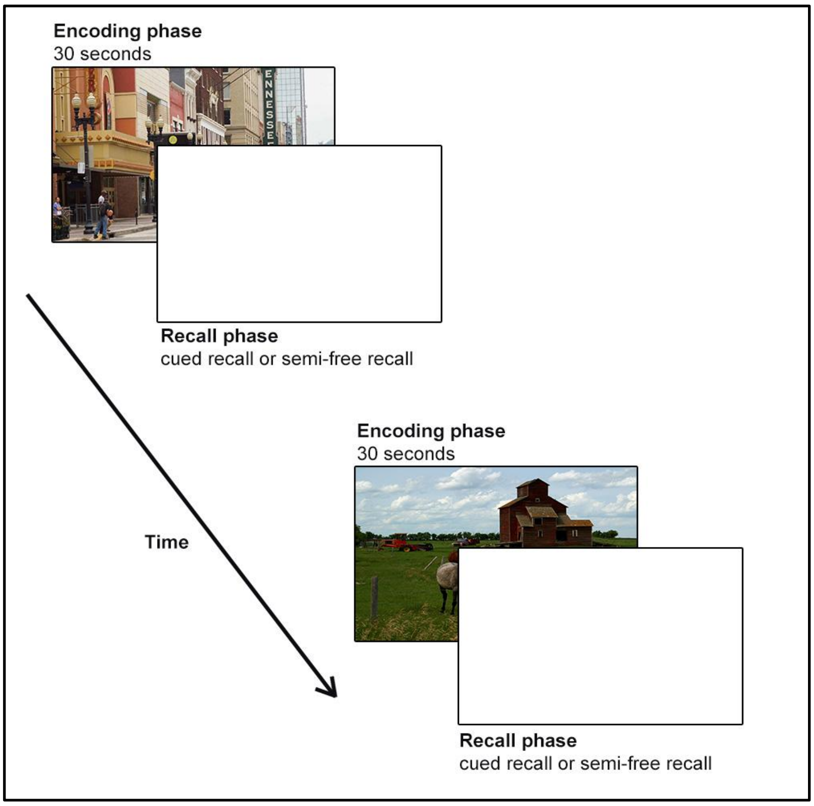


Figure 1. Experiment timeline.

2.5. Data analysis

Fifteen participants were excluded from the analysis because of corrupt data from software issues. Thus, 18 participants (10 female, 8 male) were included in the analysis (mean age = 27.8 years, *SD* = 9.2). The minimum number of participants required was determined to be 15 by an a priori power analysis (Gpower; Faul, Erdfelder, Buchner & Lang, 2009), assuming a large effect size (Cohen, 1988). An analysis of the final data showed that the conditions were normally distributed. The gaze congruency was calculated from intervals (from the onset of a question to the end of the answer) of the participants' answers by calculating the proportion of congruent fixations on that particular surface. This was done over all answers for both conditions and the means of all the participants were calculated. For the cued-recall condition, the percentages of correct answers were calculated. For the semi-free recall condition, the number of correct memory units was calculated. Vividness rating was collected using a 5-point

Likert scale for each condition. The means of each condition over all participants were calculated.

3. Results

One outlier was assessed by boxplot and was excluded from the analysis.

Gaze fixation data from the different conditions were compared to assess the difference in fixation congruency during recall. A paired samples *t*-test indicated that there was a statistically significant difference in the semi-free recall condition (M = 17.8, SD = 8.2) compared to the cued recall condition (M = 24.5, SD = 8.2), $t(16) = 3.29$, $p = .005$, $d = 0.82$, 95% CI = [-11.01, -2.29]. Levene's test did not indicate unequal variances ($F = 0.01$, $p = .913$).

A Pearson correlation coefficient was calculated to assess the relationship between congruent eye movement and memory performance. For the semi-free condition, no correlation was found between the variables, Pearson's $r(17) = .316$, $p = .216$. For the cued recall condition, no correlation was found between the variables, Pearson's $r(17) = -.183$, $p = .482$.

For our third hypothesis, a Mann-Whitney test indicated that the vividness rating was not greater in the semi-free recall condition (Mdn = 4) compared to the cued recall condition (Mdn = 3), $U = 91.5$, $p = .067$.

4. Discussion

In the present study we set out to investigate which kind of recollection method induces eye movements "to nothing", and we also set out to explore the correlation between congruent eye movements, memory performance and self-reported vividness rating.

Our results did not confirm our first hypothesis, that eye gaze is more congruent in the semi-free condition. There was a statistically significant difference in eye gaze congruency, the cued recall condition showing higher congruency in comparison to the semi-free recall condition, the opposite of our hypothesis. If this result is to be viewed through the lens of earlier research, in which eye gaze is an aid to recall, we can speculate that in the semi-free recall, the participant is recalling only "what they know" and thus might not need the aid of eye gaze to the same extent. In the cued recall condition, the participant is forced to answer questions about information that they did not necessarily encode consciously, and the eye gaze could function as an aid for recollecting that information. We hypothesized based on earlier research and arguments (Johansson et al., 2012; Scholz et al., 2017) that simplistic tasks such as recalling color, location or shape of objects would not produce as many eye movements (nor congruent eye movements), and that a more complex recall task, such as describing the picture or parts of it from memory, would show higher eye gaze congruency. Even though our results did not show support to this argument, without further research it is too early to reject it. It is possible that this reversed effect can be attributed to the cued recall condition having produced a more difficult recollection scenario, as we had no way of measuring the complexity of our independent variable of recall type. To our knowledge, the difference in the difficulty of free recall and cued recall has not been studied using complex pictures and recollection of this type may function differently with visuospatial information (as opposed to semantic information).

Our second hypothesis, that memory performance correlates positively with eye congruency, was not supported by the results. Neither the semi-free recall condition or the cued recall condition found any correlational effects. We did notice that numerous participants during the semi-free recall had their gaze fixed on one point of the screen while recalling information. If it is the case that eye movements disrupt spatial working memory, as argued by Theeuwes et al. (2009), the behavior might be explained by this. Another possibility as suggested by Johansson, Holsanova and Holmqvist (2005) is that participants shrink their mental image to the extent that they can "scan" it with their inner attention without shifting their gaze noticeably. It would then be possible for participants to perform well in the memory task without showing large (and/or congruent) eye movements. It may also be the case that participants use semantic memory strategies when encoding information, and thus recollection can be made without simulating the mental image to the same extent and thereby not inducing eye movements "to nothing". Additionally, a post-hoc analysis indicated that minimum sample size requirements for the Pearson correlation coefficient were not met, as calculated by Gpower (Gpower; Faul et al., 2009).

Examining our third hypothesis, that the vividness rating is higher in the semi-free recall condition, the results did not show a statistically significant result. One explanation for this might be the inherent difficulty in self-reporting the vividness of a mental image; the measurement may be unreliable as it is introspective. Our hypothesis was based around the connection in the existing literature between eye movements and mental imagery, and that our conditions would induce a difference in the vividness of mental imagery. It may also be the case that our conditions did not produce a significant difference in mental imagery.

4.1. Limitations of this study

The eye-tracker used in the present experiment was placed as a pair of glasses on the participant's head, which may interfere with attention and self-awareness of eye movements which was reported by a couple of participants. Many earlier studies have used a desk sitting eye-tracker that possibly does not impact the attention of the participant to the same extent. A few participants also reported post-experiment that the instructions to "keep the eyes pointed towards the monitor" made them focus too heavily on just keeping their eyes inside the monitor frame. It is possible this restriction made them not want to use their eyes normally and thus may have affected their behavior. This is another problem that could be avoided by not using a head-mounted eye-tracker.

4.2. Future research

For future research, we suggest that other methods for measuring eye gaze congruency could be investigated. We chose to measure proportions of fixations during a time interval; a different appropriate measurement is to measure saccades in the correct direction during the same interval, as one earlier study has done (Johansson et al., 2006). Analyzing the data in the present study it is quite evident that participants will sometimes saccade to an appropriate location when being asked a question, and just before answering a question. Due to technical limitations, we had no way to test this hypothesis. We also suggest to further categorize and separate the questions in the cued recall condition, as one could argue that there was some overlap in question type between the cued recall condition and the semi-free recall condition in the present study. The present study used a paradigm of measuring short-term memory recall; recall immediately after encoding, Mast et al. (2002) have suggested that the sensorimotor trace of eye gaze might only be stored in short-term memory and thus a next step in this area would be to investigate the eye gaze connection to memory over a longer period of time.

To investigate our suggested conclusion to the results, that when using complex pictures, a cued recall condition might be more difficult, we propose a further exploration of recollection with complex pictures as stimuli. Additionally, we suggest research on the probabilities of eye gaze to salient objects in a complex picture, to better be able to draw conclusions of this kind.

5. Conclusion

Although our hypotheses were not supported by this experimental design, we believe that the present study contributed a natural next step in the field to the existing literature and that it presented multiple options for future research. The argument for eye gaze as an aid for memory was not supported in this study, which may suggest that it is under certain conditions that the effect arises. The results did however show a novel finding in greater eye gaze congruency in the cued recall condition which helps in approaching one of the goals of this study- to isolate the factors that induce these eye movements “to nothing”. We believe further research about what causes this difference in recall type may bring us closer to an answer in this regard.

6. References

Andersson, R., Ferreira, F., & Henderson, J. M. (2011). I see what you're saying: The integration of complex speech and scenes during language comprehension. *Acta Psychologica.* doi:10.1016/j.actpsy.2011.01.007

Baddeley, A. D., Anderson, M. C., & Eysenck, M. W. (2015). *Memory*. Hove: Psychology Press.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22,* 577–660.

Bergstrom, K. J., & Hiscock, M. (1988). Factors influencing ocular motility during the performance of cognitive tasks. Canadian Journal of Psychology, 42, 1-23.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: L. Erlbaum Associates.

Cui, X., Jeter, C. B., Yang, D., Montague, P. R., & Eagleman, D. M. (2007). Vividness of mental imagery: Individual variability can be measured objectively. *Vision Research,47*(4), 474-478. doi:10.1016/j.visres.2006.11.013

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160.

Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2004). Brain areas underlying visual mental imagery and visual perception: An fMRI study. *Cognitive Brain Research, 20,* 226–241.

Hebb, D. O. (1968). Concerning imagery. *Psychological Review, 75,* 466–477

Holsanova, J. (2001). *Picture viewing and picture description: Two windows on the mind.* Unpublished doctoral dissertation, Lund University, Sweden.

Johansson, R., Holsanova, J., & Holmqvist, K. (2005). *What do eye movements reveal about mental imagery? Evidence from visual and verbal elicitations.* Proceedings of the Cognitive Science Conference, Stresa, Italy.

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. doi:10.1207/s15516709cog0000_86

Johansson, R., Holsanova, J., Dewhurst, R., & Holmqvist, K. (2012). Eye movements during scene recollection have a functional role, but they are not reinstatements of those produced during encoding. *Journal of Experimental Psychology: Human Perception and Performance,38*(5), 1289-1314. doi:10.1037/a0026585

Johansson, R., & Johansson, M. (2013). Look Here, Eye Movements Play a Functional Role in Memory Retrieval. *Psychological Science,25*(1), 236-242. doi:10.1177/0956797613498260

Kent, C., & Lamberts, K. (2008). The encoding–retrieval relationship: Retrieval as mental simulation. *Trends in Cognitive Sciences,12*(3), 92-98. doi:10.1016/j.tics.2007.12.004

Laeng, B., & Teodorescu, D. (2002). Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cognitive Science,26*(2), 207-231. doi:10.1207/s15516709cog26023

Marks, D. (1973). VISUAL IMAGERY DIFFERENCES IN THE RECALL OF PICTURES. *British Journal of Psychology, 64*(1), 17-24.

Mast, F. W., & Kosslyn, S. M. (2002). Eye movements during visual mental imagery. *TRENDS in Cognitive Science, 6,* 271–272.

Micic, D., Ehrlichman, H., & Chen, R. (2010). Why do we move our eyes while trying to remember? The relationship between non-visual gaze patterns and memory. *Brain and Cognition, 74,* 210–224. doi:10.1016/ j.bandc.2010.07.014

Neisser, U. (1967). *Cognitive psychology.* New York: Appleton-Century-Crofts.

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

Pylyshyn, Z. W. (2000). Situating vision in the world. *TRENDS in Cognitive Science, 4,* 197–207.

Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition, 80,* 127–158.

Pylyshyn, Z. W. (2002). Mental imagery: In search of a theory. *Behavioral and Brain Sciences, 25,* 157–238.

Pylyshyn, Z. W. (2003). Return of the mental image: Are there really pictures in the brain? *TRENDS in Cognitive Science, 7,* 113–118.

Scholz, Agnes, et al. "Covert Shifts of Attention Can Account for the Functional Role of 'Eye Movements to Nothing.'" *Memory & Cognition*, vol. 46, no. 2, 2017, pp. 230–243., doi:10.3758/s13421-017-0760-x.

Sheehan, P. W., & Neisser, U. (1969). Some variables affecting the vividness of imagery recall. *British Journal of Psychology, 60,* 71–80.

Theeuwes, Jan, et al. "Interactions between Working Memory, Attention and Eye Movement-s." *Acta Psychologica*, vol. 132, no. 2, 2009, pp. 106-114., doi:10.1016/j.actpsy.2009-01.005.

7. Appendices

Appendix A

Operational definitions

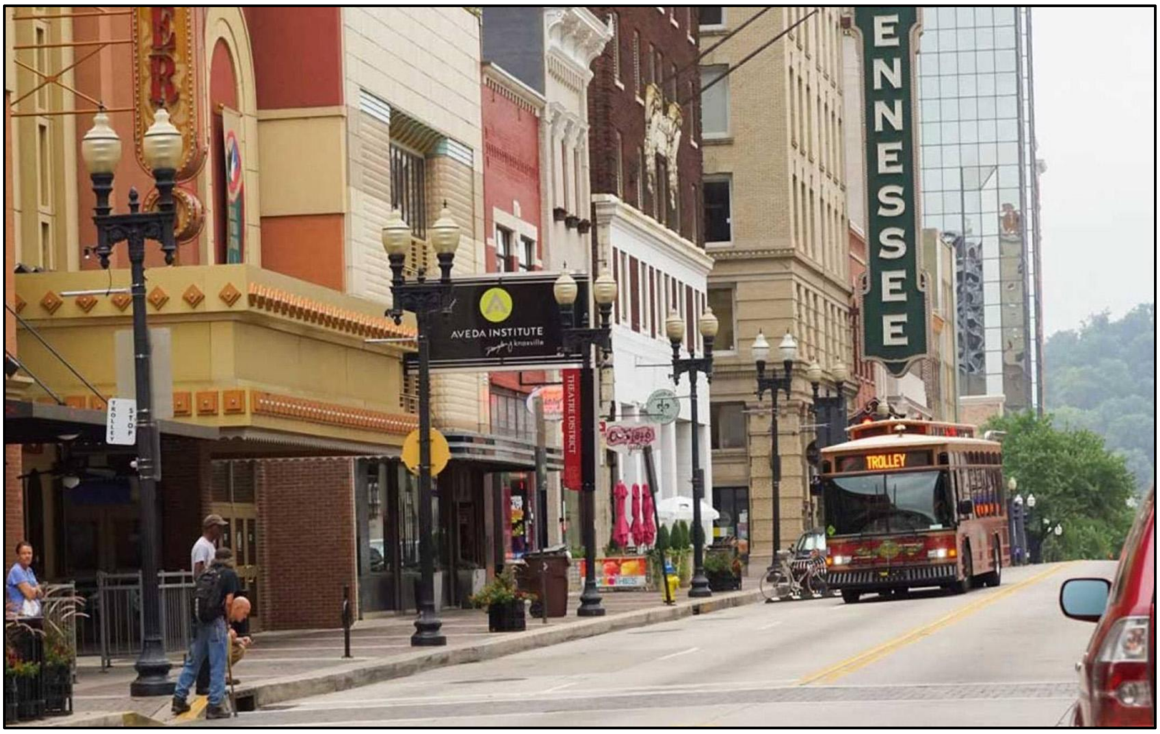
Congruent eye movement: when a participant's gaze in the recall phase matches the correct area of the screen where a specific object was encoded. Measured in percent (congruent / incongruent eye gaze fixation).

Memory performance: in the cued recall condition, measured as proportion of correct answers (percentage). Measured in memory units in the semi-free condition as quantity of correct units.

Memory unit: a measurable unit transcribed from a participant's description of a scene. Every distinct statement made about the picture is considered a memory unit which was judged be true or false.

Appendix B

Stimulus pictures



Appendix C1

Experiment questions (English)

Farm picture questions

1.
What kind of vehicles were in the picture?
What color was the tractor?
What color were the rims of the tractor?
Which way would you say the tractor was pointing?

2.
What color was the building that was seen in the picture?
How many windows were visible?
What kind of building was it?
How was the condition of the building?

3.
How many horses was there to the left of the little girl?
Was color was the horse closest to the camera?
Which way was it facing?
Was it wearing a saddle?

4.
What hair color did the woman that was seen in the picture have?
How long hair would you say she had?
What hair color did the little girl have?
What type of clothing was the girl wearing?
What was the little girl doing?

City picture questions

1.
How many lights did the lamppost closest to the camera have?
What did the black sign closest to the camera say?
What color was the logotype on it?
Was there a bird sitting on the sign?

2.
What did the large sign behind the bus say?
What color was the text of the sign?
What color was the background of the sign?
What type of building was it attached to?

4.
What color was the bus in the picture?
Which way was the bus facing?
What did the display on the bus say?
At the café by the bus there were three parasols, what color were they?

Appendix C1 (continued)

Experiment questions (English)

3.
How many were there in the group of people that was seen?
Did anyone wear a hat?
Was anyone bald?
What kind of pants were the man with the backpack wearing?
What color was the shirt of the person closest to the building?

City picture describing

In the following questions we want you to describe as well and as detailed as possible. Tell us when you are done.

3.
What can you tell us about the group of people seen in the picture?
2.
What can you tell us about the large sign behind the bus and the building it was attached to?
1.
What can you tell us about the building the group of people were standing outside?
4.
What can you tell us about the bus and its immediate surroundings?

Farm picture describing

In the following questions we want you to describe as well and as detailed as possible. Tell us when you are done.

1.
What can you tell us about the vehicles in the picture?
2.
What can you tell us about the building in the picture?
3.
What can you tell us about the horse that was closest to the camera?
4.
What can you tell us about the woman and the little girl in the picture?

Appendix C2

Experiment questions (Swedish)

Farm picture question (svenska)

1.

Vilken typ av fordon fanns det i bilden?
Vilken färg var traktorn?
Vilken färg var fälgarna på traktorn?
Vilket håll skulle du säga att traktorn stod riktad åt?

2.

Vilken färg hade byggnaden som sågs i bild?
Hur många fönster var synliga?
Vilken typ av byggnad var det?
Hur var skicket på byggnaden?

3.

Hur många hästar fanns det till vänster om den lilla flickan?
Vilken färg var hästen närmast kameran?
Vilket håll tittade den hästen åt?
Bar den hästen en sadel?

4.

Vilken hårfärg hade kvinnan som syntes i bild?
Hur långt hår skulle du säga att hon hade?
Vilken hårfärg hade den lilla flickan?
Vilket plagg hade hon på sig?
Vad gjorde den lilla flickan?

City picture question (svenska)

1.

Hur många lampor fanns det på lyktstolpen närmast i bild?
Vad stod det på den svarta skylten som var närmast i bild?
Vilken färg var logotypen på den?
Satt där en fågel på skylten?

2.

Vad stod det på den stora skylten som fanns bakom bussen?
Vilken färg hade texten på skylten?
Vilken färg var bakgrunden på skylten?
Vilken typ av byggnad satt den på?

Appendix C2 (continued)

Experiment questions (Swedish)

4.
Vilken färg hade bussen i bild?
Vilket håll var bussen riktad åt?
Vad stod det i fronten på bussen?
Utanför cafeet vid bussen fanns tre parasoller, vilken färg hade dem?

3.
Hur många var det i gruppen av människor i bild?
Bar någon av dem en hatt?
Var någon av dem flintskallig?
Vilken typ av byxor hade mannen som bar ryggsäck?
Vilken färg på tröjan hade personen som var närmast byggnaden?

City picture describing (svenska)

Följande frågor vill vi att du beskriver så väl och så detaljrikt som du kan. Säg till när du är klar.

3.
Vad kan du berätta om gruppen av människor som fanns i bild?
2.
Kan du beskriva den stora skylten bakom bussen och byggnaden som den satt på?
1.
Vad kan du berätta om byggnaden som gruppen av människor stod utanför?
4.
Vad kan du berätta om bussen och dess närliggande miljö?

Farm picture describing (svenska)

Följande frågor vill vi att du beskriver så väl och så detaljrikt som du kan. Säg till när du är klar.

1.
Vad kan du berätta om fordonen i bild?
2.
Kan du beskriva byggnaden som fanns i bild?
3.
Vad kan du berätta om hästen som var närmast i bild?
4.
Vad kan du berätta om kvinnan och den lilla flickan som sågs i bild?

Appendix D

Pre-experiment instructions (English and Swedish)

You are about to participate in an experiment that investigates pupil dilation in relation to mental workload. We will be using eye-tracking technology to film your eyes and also use audio recording to record your answers.

In the experiment you will first be shown a picture for 30 seconds, the screen will then turn white and you are to answer questions about the picture. Take all the time you need to answer. During the experiment please try to sit still and keep your eyes pointed at the monitor.

After the experiment you will be asked to fill out a short questionnaire.

Further instructions follow in the experiment. Please ask your experimenter if anything is unclear at this moment.

Du kommer nu att få deltaga i ett experiment som undersöker utvidgning av pupillerna i relation till mental ansträngning. Vi kommer att använda eye-tracking(ögonspårning)-teknik för att filma dina ögon och även att använda ljudinspelning för att spela in dina svar.

I experimentet kommer du att få se en bild i 30 sekunder, skärmen kommer sedan att bli vit och du kommer att få svara på frågor om bilden. Ta all den tid som du behöver för att svara. Under experimentets gång så ber vi dig sitta still och att hålla dina ögon riktade mot skärmen.

Efter experimentet så kommer du att få svara på en kort enkät.

Vidare instruktioner följer i experimentet. Fråga gärna din experimentledare nu om någonting är oklart.

Appendix E

Survey (English and Swedish)

Rate the vividness of your visualization during the answer phase in the experiment on a scale from 1 (not very vivid) to 5 (extremely vivid). Please rate the farm picture here. *

	1	2	3	4	5	
Not very vivid	○	○	○	○	○	Extremely vivid

Rate the vividness of your visualization during the answer phase in the experiment on a scale from 1 (not very vivid) to 5 (extremely vivid). Please rate the city picture here. *

	1	2	3	4	5	
Not very vivid	○	○	○	○	○	Extremely vivid

Hur tydlig var din visualisering när du svarade på frågorna i experimentet på en skala från 1 (inte så tydlig) till 5 (extremt tydlig). Var vänlig svara på bondgårdsbilden här. *

	1	2	3	4	5	
Inte så tydlig	○	○	○	○	○	Extremt tydlig

Hur tydlig var din visualisering när du svarade på frågorna i experimentet på en skala från 1 (inte så tydlig) till 5 (extremt tydlig). Var vänlig svara på stadsbilden här. *

	1	2	3	4	5	
Inte så tydlig	○	○	○	○	○	Extremt tydlig

Appendix F

Raw data

Participant	Angular Accura	Gender	Age	Cond. B(semi-fr	Memory units	Cond. A(cued)	Correct answers	Vividness B	Vividness A
Fp4	0.7	Male	38	15/143 (10.5%)	7	42/164 (25.6%)	13/17(76.5%)	3	4
Fp10	0.57	Female	30	71/241 (29.5%)	22	52/164 (31.7%)	9/17 (52.9%)	4	2
Fp11	0.31	Female	25	61/202 (30.2%)	8	81/196 (41.3%)	6/17 (35.3%)	3	4
Fp12	0.32	Male	61	31/161 (19.3%)	17	24/154 (15.6%)	9/17 (52.9%)	3	3
Fp13	0.34	Male	18	4/114 (3.5%)	15	29/106 (27.4%)	16/17 (94.1%)	3	4
Fp16	0.56	Male	25	14/172 (8.1%)	10	23/155 (14.8%)	8/17 (47.1%)	1	3
Fp17	0.38	Male	25	47/190 (24.7%)	12	68/202 (33.7%)	14/17 (82.4%)	3	4
Fp18	0.47	Male	29	24/143 (16.8%)	14	21/133 (15.8%)	11/17 (64.7%)	2	4
Fp19	0.46	Female	27	109/173 (63.0%)	20	49/123 (39.8%)	8/17 (47.1%)	4	2
Fp21	0.27	Female	26	19/122 (15.6%)	7	20/133 (15.0%)	14/17 (82.4%)	2	4
Fp22	0.54	Female	23	31/144 (21.5%)	13	33/157 (21.0%)	13/17 (76.5%)	4	4
Fp23	0.47	Female	26	37/177 (20.9%)	22	48/148 (32.4%)	10/17 (58.8%)	4	3
Fp24	0.41	Female	25	38/115 (33.0%)	17	46/170 (27.1%)	6/17 (35.3%)	4	3
Fp25	0.55	Female	26	12/108 (11.1%)	14	17/155 (11.0%)	13/17 (76.5%)	2	4
Fp26	0.4	Female	23	19/213 (8.9%)	12	38/132 (28.8%)	15/17 (88.2%)	3	4
Fp29	0.66	Male	25	13/75 (17.3%)	7	31/121 (25.6%)	13/17 (76.5%)	3	4
Fp30	0.6	Female	22	17/101 (16.8%)	10	28/126 (22.2%)	10/17 (58.8%)	3	3
Fp31	0.6	Male	26	26/175 (14.8%)	13	47/171 (27.4%)	10/17 (58.8%)	3	2