The functional role of gaze scanpaths in long-term visual memory

Agata Elzbieta Bochynska

Master of Philosophy in Psychology

Cognitive Neuroscience

Department of Psychology

UNIVERSITY OF OSLO

May 2014

The functional role of gaze scanpaths in long-term visual memory

2014

The functional role of gaze scanpaths in long-term visual memory

Agata Elzbieta Bochynska

http://www.duo.uio.no

Trykk: Reprosentralen, Universitetet i Oslo

Acknowledgements

I would like to thank my supervisor Professor Bruno Laeng for giving me the opportunity to take part in this project and successfully accomplish the study. Thank you, Bruno, for the original idea of the study as well as the suggestions that were crucial for designing the experiments. In addition, thank you for your careful and strict guidance and for the inspiration that kept me working intensively over two years. I would also like to thank my fellow students that were always there with a good advice and technical support during the process of data collection and analysis. Lastly, I would like to thank my parents for their support and care during my studies abroad.

Abstract

Eye movements occur spontaneously during recall from the visual memory despite neither the object nor context remembered may be visible at the time of recall. Recent research provides support for the hypothesis that these eye movements could play a functional role in the retrieval of stored information by matching the original eye fixations from the time of encoding. However, studies have provided conflicting results in this field. To investigate this, we designed a novel experimental procedure where the eye scanpaths from encoding were forced to be either enacted or not during recognition of the image. Simple geometrical patterns resembling checkerboards were used as stimuli in the current experiment. We hypothesized that enactment of the original scanpath would facilitate memory retrieval. Observed results supported the predictions, which is in accordance with theories pointing to a functional role of eye movements in long-term memory.

Contents

Introduction

Researchers have observed that eye movements are rarely under voluntary control but they can give valuable insights in the activation of representations or the time course of cognitive processes (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). Additionally, early in the literature one can find reports that point to the relationship between eye movements and imagery recall (Hebb, 1968; Jacobson, 1932; Neisser, 1967; Totten, 1935). It has been also shown that spatial locations can be successfully used as cues in memory retrieval (Sinclair, Healy, & Bourne, 1997; Winograd & Church, 1988) and recent research focusing on the refixations to original locations from encoding consistently supports those findings (e.g. Johansson, Holsanova, Dewhurst, & Holmqvist, 2012; Laeng, Bloem, D'Ascenzo, & Tomassi, 2014; Spivey & Geng, 2001). Moreover, the interest in theories supporting embodiment of cognition or simulation of perception has been steadily increasing the past several years (Barsalou, 2008, 2010; Hesslow, 2011). Thus, if one would associate eye movements with a motor pattern, it would point to the relationship between visual memory and oculomotor coordinates (Richardson & Spivey, 2000). However, it is still unclear whether eye movements to relevant locations play a functional or epiphenomenal role in visual memory. Additionally, no conclusive evidence for the role of temporal order of fixations has been provided to date (see Foulsham & Kingstone, 2012). The present study offers additional evidence in favor of a functional role of oculomotor information in visual memory.

Background

Scanpath Theory

The idea of a special role of eye movements both in scrutinizing an object when present and in recalling it when absent has appeared early in the psychological literature, particularly in connection with the emergence of research on visual imagery. According to Neisser's (1967) suggestion, either eye movements or processes that trigger them play an essential and active role in constructing visual images. Hebb (1949, 1968) also pointed out that eye movements are crucial in organizing parts of the images into complete visual images.

Yarbus (1967) noticed that eye movements are executed towards objects or shapes that provide useful information. Additionally, eye movements might be repeated several times in cycles, refixating on relevant features. Following those thoughts, Noton and Stark (1971a, 1971b) suggested that such cycles could be executed at both encoding and recognition and that they could assist retrieval processes. They defined scanpath as a fixed path that was followed intermittently but repeatedly by participant's eyes while viewing a pattern.

Several years later, Brandt and Stark (1997) ideated an experimental procedure to investigate scanpaths during perceptual viewing and imagery conditions. Participants would look at several geometric diagrams that resembled checkerboards and subsequently imagined the seen patterns. Eye movements were recorded during both sessions and compared with the help of a string editing analysis. The analysis revealed a close correlation between eye scanpaths executed during encoding and imagery recall. Based on those observations, Brandt and Stark (1997) concluded that eye movements at imagery were not random but they reflected the content of the encoded scene. This finding pointed to the possible relationship between memory processes and motor activity of the eyes.

Brandt and Stark's (1997) results supported the hypotheses formulated earlier by Noton and Stark (1971a) known as Scanpath Theory. The theory is based on an information-processing approach to visual perception, which presupposes that sequences of sensory and motor activities form an internal representation of the image (Brandt & Stark, 1997). That is, scan patterns produced during encoding of the picture are an integral part of the memory for a concrete picture. However, the theory has not gained much interest among researchers and many remained skeptical to the idea that retrieval from visual memory would require the exact re-enactment of the eye scanpaths (see Henderson, 2003). Therefore, in the course of years many scholars have criticized this account.

The main reservation to the theory was that it originally based on relatively little evidence. Furthermore, researchers made unjustified assumptions about the neural processes underlying oculomotor activity and its relationship to memory retrieval (Humphrey & Underwood, 2008). As Humphrey and Underwood (2008) pointed out, occurrence of the scan patterns at encoding and retrieval did not necessarily mean that integrated perceptual and motor representations are involved. Therefore, the issue would require more research to draw such conclusions.

Originally, one of the main arguments against Scanpath Theory was that bottom-up influences like visual saliency of the image drive the sequences of the eye movements and therefore scan patterns at encoding and retrieval for the same image are similar. One could argue that the influence of visual saliency was overdriven but then the argument about semantically interesting parts of the image that attract attention and cause scanpath re-enactment could nevertheless hold (Humphrey & Underwood, 2008). However, many studies avoided this issue with focusing on imagery tasks (Brandt & Stark, 1997; Laeng & Teodorescu, 2002; Laeng et al., 2014) or on the observation of spontaneous eye movements to blank spaces (Ferreira, Apel, & Henderson, 2008). Nevertheless, Humphrey and Underwood (2008) observed that one could not be sure about the direction of causation in this cases, as scan pattern might match original eye movements from encoding as a result and not as a consequence of the better recall.

Thus, Scanpath Theory was one of the first theoretical frameworks where the core assumption was that the order of eye fixations is a part of a memory trace. However, no reliable evidence for the "strong" version of the theory has been provided so far. Moreover, none of the studies to date have incontrovertibly supported the hypothesis that re-enacting the original sequences of fixations could facilitate retrieval. However, several studies that took an attempt to test single hypotheses of the theory observed reliable results that could support them to some extent (Foulsham & Kingstone, 2012; Humphrey & Underwood, 2008, Laeng & Teodorescu, 2002). Additionally, the idea of the involvement of eye movements in the retrieval processes has been gaining interest in the recent years, providing new evidence for their role in visual memory and imagery.

Common neural processes underlying visual imagery and memory

Studies that investigate the role of eye movements at retrieval have tested both imagery and memory processes. However, joint conclusions for those two lines of research are justified as both processes are overlapping to a great extent and imagery can be considered as a perceptual form of episodic memory. Already in XIX century Sir Francis Galton noticed that visual imagery seemed to play an essential role for visual memory (Galton, 1883). Similarly, in his theory of 'visual buffer', Kosslyn (1994) connected memory and imagery in a cognitive processes underlying encoding and retrieval. According to him, mental images are generated in the buffer and their representations are then stored in long-term memory so as to re-create the image in working memory during retrieval.

In line with the above considerations, research indicated the existence of common brain regions underlying both processes (Farah, 1995; Kosslyn, Thompson, Kim, & Alpert, 1995). Kosslyn, Ganis and Thompson (2001) found out that frontal-parietal control regions were active during visual imagery and memory tasks. In another study, the same authors observed that memory and imagery processes were overlapping, mostly in fronto-parietal but also in temporal and occipital regions (Ganis, Thompson, Kosslyn, 2004). However, in a recent report, Slotnick, Thompson and Kosslyn (2012) noticed that the existing evidence has not provided sufficient support for the hypotheses about common neural correlates of visual memory and imagery. Therefore, they introduced a novel task paradigm where both processes were tested in a similar way while brain activity was measured in fMRI. Their main hypothesis was based on the assumption that if two processes activate the same brain regions then they should be to some extent functionally identical. The experiment consisted of familiarization phase, where participants viewed different objects (e.g. zebra, feather, rope), and subsequent memory and imagery phases after 1 to 7 days delay. In the memory phase participants' task was to respond "remember", "know" or "new", whereas in the imagery phase they were rating the vividness of recalled object by stating "high vividness", "moderate vividness" or "low vividness". Results revealed that several distinct brain regions were activated. The superior parietal lobule, inferior parietal lobule and the precuneus, as well as areas in fusiform gyrus together with extrastriate and striate cortex were active for memory but not imagery processes. However, both processes elicited significant activation in common areas like frontal-parietal control regions, specifically anterior frontal cortex, dorsolateral prefrontal cortex and intraparietal sulcus. According to Slotnick, Thompson and Kosslyn (2012) those results indicate that memory and imagery are overlapping to a great extent in terms of the neural activity. Nevertheless, non-overlapping activity revealed that they are also two different processes.

Thus, neuroimaging studies point to similar brain regions underlying visual imagery and visual memory. Consequently, as both processes share common neural paths and given that research has revealed significant motor activity of eye movements in visual imagery (e.g. Brandt & Stark, 1997; Laeng & Teodorescu, 2002; Underwood & Humphrey, 2008), one could expect eye movements to play an important role in visual memory as well.

Encoding and retrieval compatibility

The idea suggesting that gaze plays a functional role in memory presupposes that oculomotor processes present at the encoding can be reinstated during retrieval. Hence, it suggests that encoding and retrieval processes are compatible by using the same processes and procedures in both instances. Already in the 70's researchers observed that the match between encoding and retrieval facilitates retrieval of the information by bringing back the cues that were present at encoding (Kolers, 1973; Tulving & Thompson, 1973). Similarly, a transfer appropriate processing account in the cognitive sciences presupposes that repeating processes used at study would enhance memory performance during test (Blaxton, 1989; Kolers, 1973). Additionally, context-dependent memory research is based on similar assumptions (Bjork & Richardson, 1989; Godden & Baddeley, 1975; Winograd & Church, 1988). According to these accounts, the stronger the overlap between encoding and retrieval processes, the more successful retrieval from memory. However, even though there is a vast evidence for this phenomenon (Godden & Baddeley, 1975; Morris, Brandsford, & Franks, 1977), Holm and Mäntylä (2007) point out that none of those studies or theories provides an explanation of how the cues or the processes are brought back at recognition, whether it is an incidental reinstatement or reconstruction driven by the agent. They referred to Jacoby and Craik (1979) as one of the few researchers that proposed a more specific mechanism of encoding and retrieval compatibility. Jacoby and Craik (1979) suggested that in the absence of a direct access to the representation, a reconstructive process is undertaken starting with a general context through subsequent reconstructive operations until the recognition is achieved. Based on this reasoning one could conclude that an agent's goal is to aim for the highest encoding-retrieval compatibility in order to facilitate the memory (Holm and Mäntylä, 2007).

Neuroimaging studies provide a support for those hypotheses through growing evidence for common neural systems activated during encoding and retrieval (Nyberg, Habib, McIntosh, & Tulving, 2000; Wheeler, Petersen, & Buckner, 2000). Remarkably, already in the XIX century, James (1890) suggested that processes from encoding could be reinstated at retrieval in the respective brain areas. In their recent review, Danker and Anderson (2010) pointed to a vast number of neuroimaging and patient studies that report a reliable similarity of the brain states at perception and recall or study and test. This regularity concerns different sensory modalities and valences of emotional stimuli as well as reactivation of strategies used at encoding. Additionally, studies show that reactivation of the encoding processes is stronger when a bigger amount of information is being retrieved whereas incorrectly remembered information leads to reduced reactivation (Danker & Anderson, 2010).

Similarly, Kent and Lamberts (2008) reviewed studies investigating the relationship between encoding and retrieval and interpreted retrieval in terms of mental simulation. They referred to the evidence from neuroimaging, behavioral studies on eye movements and the relation between time courses of processes at encoding and retrieval. According to the authors, findings from all mentioned areas support their hypothesis that encoding-retrieval compatibility is based on the simulation processes. Interestingly, these authors point to the eye movements as behavioral indicators of such simulation or reinstatement in visual memory and imagery and that the role of those movements is functional in retrieval (e.g. Laeng & Teodorescu, 2002; Johansson, Holsanova, & Holmqvist, 2006). Consequently, the theory about the reinstatement of encoding processes in retrieval seems in accordance with simulation theories and supports the hypothesis about an important role of eye movements in memory retrieval.

Simulation of perception and grounded cognition

A reinstatement of the encoding processes in retrieval suggests that both the sensory and motor activity reappear during recall. Accordingly, there is a growing body of evidence for the presence of motor perception in imagery and memory. Early studies have shown that imagery engages activity in the muscles that would be used in the imagined activity (Jacobson, 1932) and duration of imagined action is highly similar to the original time required for the action itself (e.g. Decety, Jeannerod, & Prablanc, 1989). In line with the view that motor processing is strongly connected to imagery, studies have shown that some motor activity can disrupt generation of a mental image. Quinn (1994) reported disruption of the recall due to the concurrent arm movement. Similarly, clockwise mental rotation can be disrupted by the perception of a counterclockwise rotation of the same object (Wexler, Kosslyn, & Berthoz, 1998). Regarding involvement of the eye movements, several studies have shown that precluding those movements or executing them in irrelevant ways has an effect on the quality of visual memory retrieval (e.g. Barrowcliff, Gray, Freeman, & Macculloch, 2004; Gunter & Bodner, 2008; Postle, Idzikowski, Della Sala, Logie, & Baddeley, 2006). Additionally, in a study with amnestic patients with damage in the temporal lobe and hippocampus regions, authors observed refixations to the correct locations of the images in the absence of explicit memory for them (Laeng et al., 2007). Taken together, generation of a mental image or retrieval from memory can be affected by motor activity of the body, which is in accordance with the idea that eye movements have a positive impact on recall from visual memory.

Evidence from neuroimaging supports this hypothesis since imagining and playing a concrete piece of music on a piano (Meister et al., 2004) as well as imagery of various movements and performing them (Decety et al., 1994; Lotze et al., 1999) engage activity in similar brain areas. Additionally, some studies revealed the engagement of motor areas in the brain during the task of mental transformations of Shepard-Metzler cubes‘ drawings (Deutsch, Bourbon, Papanicolau, & Eisenberg, 1988). Furthermore, patient studies showed that lesions in parietal cortex can lead to the difficulties with inhibition of the overt movements during imagery and at the same time supported the hypotheses that imagery and movement enact similar neural processes (Schwoebel, Boronat, & Coslett, 2002). Consequently, acknowledging common processes underlying visual memory and visual imagery as well as evidence supporting the simulation of perception theory, one could reason that visual memory is also driven through similar neural paths as perception. Thus, memory retrieval can be assisted by the motor activity like eye movements.

In a recent review paper, Hesslow (2011) reviews various studies supporting simulation theory and argues that this approach provides an explanation of the occurrence of the “inner self” as well as relationships between cognitive and motor functions. Additionally, he defines ‘simulation theory’ and its principles. That is, in the light of this theory simulation is based on purely associative mechanisms and involves the same neural mechanisms that those originally engaged in movement or perception. The theory consists of three main components: simulation of behavior, simulation of perception and anticipation. The last component, anticipation, involves making use of the knowledge about the sensory consequences of action to predict the consequences of behavior. However, it is not only experienced actions that can be used. Simulations can be based on the experience to anticipate a new consequence so that one is able to imagine a new and not known object or action (Hesslow, 2011). In terms of the explanatory power of the theory, Hesslow (2011) indicates the recall from long-term memory as an example where a reactivation of the sensory activity from encoding clearly takes place. Kent and Lamberts (2008) suggest that all recall from memory is a mental simulation.

Similarly, Barsalou (2008) argued for the concept of grounded cognition, where bodily states, modal simulations and action are essential parts of cognitive processes, in contrast with the standard theories of cognition, where memory systems are separated from perception or action. Barsalou’s concept points to the importance of simulation as re-enactment of bodily states that were present during gathering of the information. Encoded information is integrated across modalities in the brain and in the retrieval all modalities are reactivated while bodily states reappear. According to Barsalou (2008), the concept of grounded cognition could explain the relationship between eye movements and visual memory or imagery. Re-enactment of eye movements in imagery or recognition phase would correspond then to the simulation of bodily states. In this way, visual cognition would be “grounded” in eye movements.

Similar accounts for the simulation processes underlying cognition have appeared under different names. Ballard, Hayhoe, Pook and Roe (1997) pointed to the specific level of embodiment in cognition when “the constraints of the physical system determine the nature of cognitive operations” (p. 723) and claimed that motor systems can affect different cognitive processes. Similarly, Thomas (1999) pointed to the active nature of perception and suggested a “perceptual activity theory”. The extent of those considerations and the supporting evidence points to a crucial role of motor processes in cognition.

Eye fixations and visual saliency

However, reinstatement of the encoding processes or simulation of perception are not the only reasons for which particular eye movements may occur during retrieval. The locations of eye fixations can be determined by meaningful information in the picture and by its perceptual saliency. Saliency can be defined as a quality of an item to be visually distinctive relative to the items surrounding it (Humphrey & Underwood, 2008). This can influence attention shifts manifested by the eye movements that would be drawn to the most salient regions of the image while looking at a scene or an object (Itti & Koch, 2000; Koch & Ullman, 1985). Salience of the image is calculated as a map of local visual contracts in the picture, separately for color, orientation and intensity (Itti, Koch, & Niebur, 1998). The contrasts are summed up for each location and thus can be represented topographically in one salience map. Fixations are more likely to occur in the positions with higher local salience value (Itti & Koch, 2000). However, there is conflicting evidence for an influence of the visual saliency on eye movements. Some researchers find a strong evidence for attracting fixations by saliency in the picture (Parkhurst, Law, & Nebur, 2002; Underwood & Foulsham, 2006; Underwood, Foulsham, van Loon, Humphreys, & Bloyce, 2006) while the others observe minimal correlation between the fixations and salient features of the image (Tatler, 2007; Tatler, Baddeley, & Gilchrist, 2005; Foulsham et al., 2012). Additionally, more factors than just visual salience could drive fixations to particular locations, e.g. top-down control of vision and a preference for new (as compared to repeated) information (Valuch, Becker, & Ansorge, 2013).

Henderson (2003) defined gaze control as a "process of directing fixation through a scene in real time in the service of ongoing perceptual, cognitive and behavioral activity" (p. 498). This control could be either stimulus-based (bottom-up) or knowledge-driven (top-down). Stimulus-based control indicates that eyes are fixating on the salient areas in the picture or scene whereas knowledge-driven control is based on the stored in memory or mental representations information that could guide eye movements to particular locations. According to Henderson (2003), a top-down approach assumes that eye movements are driven by internal information particularly in active tasks with meaningful stimuli.

Recently, several different studies have provided a reliable evidence for the essential role of top-down gaze control in visual memory and imagery (Foulsham & Underwood, 2008; Humphrey & Underwood, 2009; Underwood, Foulsham, & Humphrey, 2009). Foulsham and Underwood (2008) investigated scanpaths' spatial similarity at encoding and retrieval and used colorful images of scenes as stimuli. They observed higher than chance within-subjects adherence of the order of fixations in encoding and recognition. Additionally, authors compared the model of saliency maps of the images with spatial distribution of the fixations and observed that visual saliency could not account for the similarity between the scanpaths. Furthermore, Foulsham with colleagues (2013) observed recently similar results while using a different method of scanpath comparison (MultiMatch) for the data collected by Foulsham and Underwood (2008). Consequently, these recent studies support an account for knowledge-driven or top-down gaze control.

"Looking at nothing"

Top-down control of the gaze is an essential assumption for the hypothesis of a functional role of eye movements. That is, if only the salient features in the picture guided gaze, eye movements would not act as a part of a memory trace. In fact, the occurrence of spontaneous eye movements to blank spaces during recall is one of the strongest arguments against stimulus-driven gaze control. However, there is an ongoing debate about eye movements in such "looking at nothing" situation (Foulsham & Kingstone, 2012). Some researchers (Richardson, Altmann, Spivey, & Hoover, 2009; Richardson & Spivey, 2000) propose that spontaneous eye movements to blank spaces during recall don't have an impact on memory performance whereas the others (Ferreira, Apel, & Henderson, 2008; Johansson, Holsanova, & Holmqvist, 2006) observe similarity between eye movements at encoding and during retrieval when there is nothing to look at. Laeng et al. (2014) pointed out that this phenomenon might seem puzzling due to the seeming waste of bodily energy and useless cognitive effort for producing eye movements when there is nothing to look at. According to the researchers, a solution for this puzzle could be to acknowledge that those movements are not irrelevant but that they play a useful role in visual memory and imagery (Laeng et al., 2014).

Spontaneous eye movements to blank spaces or in the darkness indicate that not only saliency or meaningful objects in the picture attract attention and drive fixations. Ferreira, Apel and Henderson (2008) described this phenomenon as “looking at nothing” and defined as fixating at the locations previously occupied by significant visual information. Remarkably, this spatial indexing mechanism makes eyes move also to the locations where removed item was previously present, while subjects are not aware what is missing (Henderson & Ferreira, 2004). Moreover, studies show that looking at blank but relevant regions, where the target was previously encoded, could improve memory for that object (Hollingworth, 2006; Johansson & Johansson, 2014).

Accordingly, Ferreira, Apel and Henderson (2008) pointed out that looking at nothing could facilitate the retrieval of both visual and linguistic information. These authors argued for integrated representations in visual cognition that are formed during scene viewing. The main assumption underlying this hypothesis is that mental, internal representations are formed from the external information about the world. It indicates that visual information is encoded from where the subject was fixating his gaze. Therefore, when listening to spoken language that describes particular scene or object, eyes would move to the spatial locations where they could be placed. However, those refixations would not appear in order to encode the external visual information but to re-create its internal representation. Thus, this account supports the functionality of the eye movements in memory retrieval. Additionally, such reasoning stays in accordance with the recent findings of Johansson and Johansson (2014; see also Johansson, Holsanova, & Holmqvist, 2006; Johansson Holsanova, Dewhurst, & Holmqvist, 2012) where the spoken instructions evoked refixations to the locations where the stimulus had been originally encoded. Consequently, “looking at nothing” situation supports the importance of eye movement’s role in visual imagery and memory.

Pupillary changes and cognitive effort

Recent research investigating the role of eye movements in visual memory and imagery have tested the effect of precluding such movements through central fixation on the memory retrieval (e.g. Johansson et al., 2012; Laeng & Teodorescu, 2002; Mäntylä and Holm, 2006). The method seems crucial to show that if gaze is perturbed during recall so it is the memory, suggesting a causal link between the two. However, this forced fixation manipulation has been criticized based on the possibility of an additional cognitive effort that maintaining central fixation could cause during the task (Martarelli & Mast, 2013). Although several researchers have argued that maintaining fixated gaze is not very likely to increase working memory load (Johansson et al., 2012; Laeng et al., 2014) and that it is in fact a common procedure in psychological experiments on attention and perception, no external measurement of changes in attentional effort have been empirically provided during memory studies comparing forcing fixation versus freedom to moving the eyes. Crucially, pupillary responses – which are obtained together with gaze information in current infrared eye-trackers, have been indicated as a reliable and valid index of cognitive load or attentional effort (Kahneman, 1973; Kahneman & Beatty, 1967). Similarly, several researchers suggested that pupillary changes could index the level of intensity with which the system is operating (Just & Carpenter, 1993). Beatty and Kahneman (1967) observed that the increase of the number of digits to be remembered and thus the increase of the working memory load was positively correlated with the increase in pupil diameter. Subsequent studies consistently supported the hypothesis about the relationship between pupillary responses and cognitive effort (see Laeng, Sirois, & Gredebäck, 2012 for a recent review). Thus, pupillary changes can be used as a reliable measure of the additional load in the working memory.

Previous findings

Functional or epiphenomenal role of eye movements?

Recent research investigating spontaneous eye movements in visual imagery has consistently reported the congruency between scan patterns during encoding and retrieval (e.g. Altmann, 2004; Brandt & Stark, 1997; Humphrey & Underwood, 2008; Johansson, Holsanova, & Holmqvist, 2006; Knoeferle & Crocker, 2007; Laeng & Teodorescu, 2002;

Spivey & Geng, 2001). On the other hand, the explanation for this phenomenon is not consistent any more and different studies provide conflicting results supporting either functional or epiphenomenal role of eye movements in visual imagery and memory.

One of the most relevant studies investigating the relationship between eye movements and recall is by Laeng and Teodorescu (2002), which revealed a significant effect of scan patterns on the quality of retrieval. Researchers used simple stimuli presenting irregular checkerboard or color images of fish that were firstly encoded and subsequently retrieved in an imagery task. Participants were divided into groups where their eye movements were respectively manipulated. In the first experiment subjects were viewing grid patterns resembling checkerboards (similar to the patterns used in the study by Brandt & Stark, 1997) for 20 seconds and subsequently were asked to imagine the patterns. Participants were divided into two groups: Free Viewing group was allowed to freely inspect the patterns during encoding whereas Central Fixation group maintained the gaze in the middle of the screen while encoding the patterns. Both groups were allowed to move their eyes during imagery task and the ratings regarding vividness of imagery were collected. Additionally, a spatial memory test was pursued after viewing and imagery of each pattern. Data analyses revealed that scanpaths from perceptual phase were strongly correlated with eye movements for the same object in the imagery phase. These results support the findings by Brandt and Stark (1997) as well as Noton and Stark's (1971a, 1971b) hypotheses. Interestingly, in Laeng and Teodorescu's (2002) study, subjects' scanpaths were also reliably similar in the Central Fixations group, which means that central encoding led to the tendency of voluntary fixation of the gaze in the middle of the screen during imagery. Thus, such oculomotor behavior would point to a mechanism based on the generation rather than inspection of the picture during imagery, as participants freely chose to suppress spontaneous eye movements. Additionally, a relation between scanpaths similarity and recall (i.e. memory accuracy) was observed and spatial memory scores were equally correct.

In the same study, Laeng and Teodorescu (2002) conducted a second experiment in order to replicate results from the first one as well as further investigate the relationship between eye scanpath and recall. Therefore, a third group of subjects was introduced. In this condition participants were asked to freely investigate the images during encoding but remain central fixation during imagery phase. The main reason for adding a novel condition was to investigate whether the image generation process would be disrupted when oculomotor behavior from encoding was precluded. Furthermore, new stimuli were used and presented a color images of a different tropical fish localized in one of the four corners of the screen. The participants' task was to firstly memorize the images (either freely moving their eyes or fixating in the center) and subsequently imagine seen pictures and respond to the questions about particular features of the fish as a measure of memory accuracy. Results revealed that findings from the first experiment were successfully replicated as scanpaths observed at encoding were re-enacted during recall. Additionally, third group of subjects (Free Perception & Fixed Imagery) showed reliable lower accuracy rates compared to the other groups. Thus, precluding eye movements during recall but not during perception had an effect on the quality of memory.

The functional account, according to Laeng and Teodorescu (2002), is based on the hypothesis that encoding of each fixation on the perceptual level works later as an index to the particular location in the image. This form of spatial coding presupposes that visual information is stored together with the motor commands. On the contrary, the epiphenomenal account interprets eye movements during imagery as passively subordinated to the covert shifts of attention. In this way eye movements are reflecting the imagery process but are irrelevant to it. Based on the results from the study, Laeng and Teodorescu (2002) concluded that their findings pointed to the functional role of eye movements in visual imagery as the same oculomotor behavior during encoding and retrieval had a positive effect on recall. This stays in accordance with the effect of eye movements' manipulations on the recall scores observed earlier by Janssen and Nadine (1974).

Nevertheless, other researchers report results that stand in contrast with the above findings. Several studies reported the lack of the effect of eye movements on the memory retrieval processes and therefore acknowledged their role in recall as epiphenomenal (Hoover & Richardson, 2008; Richardson & Kirkham, 2004; Richardson & Spivey, 2000; Richardson et al., 2009; Spivey & Geng, 2001). The study by Spivey and Geng (2001) investigated eye movements during imagery and memory for disappearing objects. In the first experiment, participants' task was to look at a blank screen and listen to the spoken descriptions of particular scenes. Results revealed a reliable tendency of directing the eyes to the locations suggested by auditorily presented scene description. Similarly, in the second experiment, where participants looked at simple shapes and were asked about particular features of the vanished objects, the eye movements were again directed to the relevant blank regions when the information was recalled. Thus, Spivey and Geng (2001) acknowledged that those movements are clearly not random behaviors. However, accuracy did not differ depending on the proportion of fixations directed to the relevant blank regions. Therefore, in this study, the spontaneous eye movements “to nothing” did not facilitate memory retrieval and could not be recognized as functional.

The paradigm developed by Spivey and Geng (2001) was used in Richardson and Spivey’s (2000) research investigating similar issues. Authors designed several experiments where they tested for the relationship between spontaneous eye movements to blank regions and retrieval processes. Participants were presented with auditory information in different visual locations and video files were used as stimuli. Richardson and Spivey’s (2000) results revealed that spontaneous eye movements were pursued to the empty but relevant locations. However, those movements were not associated with better accuracy in the task and therefore could not be interpreted as functional. Additionally, when eye movements were differently manipulated for two groups of subjects in the encoding phase, researchers did not observe any reliable differences between those two groups. Consequently, authors of the study proposed an interpretation in terms of spatial indexing, where cognitive system was giving a spatial tag to the particular representation (Richardson & Spivey, 2000).

Conflicting results lead to ongoing discussion in the literature and various attempts of explaining this issue. Johansson et al. (2012) (see also Johansson & Johansson, 2014) pointed to several limitations of the previous research, particularly the usage of simple and artificial stimuli and a low level of difficulty in the recall tasks as well as unclear instructions. As authors underlined, those issues suggest the importance of the experimental procedure as they could strongly influence the results. Furthermore, the studies that showed contrasting results (Laeng & Teodorescu, 2002; Spivey & Geng, 2001) used distinct experimental designs and stimuli as some experiments were based on the spoken descriptions of the scenes while the others involved only visual images (Johansson et al., 2012). Taken together, a reliable evidence for the functional role of eye movements in visual memory has not been provided yet.

Refixations to the same locations and memory facilitation

Already Geoffrey Loftus (1972) had observed that the recognition of a scene was positively correlated with the number of fixations made during encoding. Some later studies showed similar results when investigating object’s recall (Irwin & Zelinsky, 2002). However, recently a growing number of studies examined the relationship between those fixations made to the original locations from encoding and the quality of memory retrieval. In their eye-tracking study, Mäntylä and Holm (2006) used the remember/know paradigm in order to decompose recognition into separate measures and observed impairment of explicit memory (but not familiarity-based recognition) when subjects' eye movements were precluded. The remember/know paradigm (Tulving, 1985) assumes that recognition occurs when specific experiences in which the information was originally encoded are brought to mind whereas knowing emerges from feeling of familiarity. Mäntylä and Holm's (2006) results revealed a significant reduction of "remember" responses when eye movements were constrained either during encoding or during retrieval whereas this preclusion did not affect "know" answers. According to these authors, the finding pointed to the distinct role of eye fixations in explicit and familiarity-based memory. Furthermore, in a follow up experiment, Mäntylä and Holm (2006) observed that explicit memory was also associated with more consistent fixations between encoding and retrieval phases. Taken together, eye movements could play a functional role only in explicit memory and a marginal role in familiarity-based recognition.

Following this thought, Holm and Mäntylä (2007) designed a new experiment where they investigated the relationship between eye movements and explicit memory as well as familiarity. They also argued that one of the reasons for the lack of the reliable evidence for such a relationship could be that researchers disregarded those two separate components of recognition. In their study, Holm and Mäntylä (2007) defined eye movement consistency in terms of proportions of fixations at retrieval falling within 2 degrees of visual angle from the center of any fixation from the encoding of a stimulus. They created three different experiments where they showed images of landscape paintings. All three experiments supported the hypothesis that encoding-retrieval consistency is significantly related to the recognition performance. Holm and Mäntylä (2007) also observed a positive correlation between the number of refixations to the locations that subjects looked at during learning sessions, and the level of explicit memory. Thus, they acknowledged explicit memory as function of perceptual reconstructions.

In a different experiment that likewise tested the role of refixations, Johansson, Holsanova and Holmqvist (2006) investigated whether seeing a picture or listening to spoken descriptions of the scene while looking at the blank screen will evoke similar patterns of eye movements during encoding and retrieval. Their results showed that when subjects were recalling the image from memory, they were executing eye movements to highly similar locations as during encoding. In a follow-up of this study, Johansson, Holsanova, Dewhurst and Holmqvist (2012) designed several new experiments with eye movements' manipulations to investigate whether 1) eye movements during recall are reinstatements of those during encoding and 2) eye movements during recall have an impact on memory retrieval. Four experiments were set up with diverse combinations of free viewing and forced fixation conditions both in encoding and retrieval phases. They used both spoken scene descriptions and visual stimuli in different combinations. The results showed that forced, central fixation in the encoding phase had no effect on eye movements during recognition but hindered memory. On the other hand, central fixation in the retrieval phase impaired the memory for the scene regardless of the encoding modality. Therefore, authors concluded that eye movements during retrieval were not reinstatements of the original scanpath from the encoding phase, but that they have a functional role as they affect the memory retrieval.

However, Martarelli and Mast (2013) argued that maintaining central fixation could act like a distraction and increase working memory load. Thus, such an additional cognitive effort during central fixation could have had an impact on other on-going cognitive processing, like memory. Therefore, Johansson and Johansson (2014) used recently an experimental procedure similar to the one designed by Martarelli and Mast (2013), which helped to avoid the issue of increasing cognitive effort. Additionally, unlike in the previous experiments, authors investigated the memory for the spatial arrangement of the objects and not the single objects' properties. Thus, they tested the idea that eye movements work as 'spatial indexes' for the memory. Spatial indexing (Richardson & Kirkham, 2004) is understood here as linking internal representations to perceived locations in order to reduce the effort of working memory (Ballard, Hayhoe, Pook, & Roe, 1997). As Johansson and Johansson (2014) underlined in the article, the role of eye movements in visuospatial memory for object features could be different from their influence on spatial arrangement between the objects. Therefore, they created a new experiment based on the previous experimental designs from Johansson, Holsanova and Holmqvist (2006) and Johansson Holsanova, Dewhurst and Holmqvist (2012). The study consisted of encoding and retrieval phases and used visual stimuli. In the encoding phase participants were asked to look at 24 objects distributed in the four quadrants of the screen (6 objects in each quadrant). Firstly, they looked at 6 objects from each quadrant simultaneously for 30 seconds and named them and their positions (half of the objects were facing right and half of them were facing left). Subsequently, they looked at all 24 objects in four quadrants for 60 seconds. In the recall phase participants heard statements regarding intraobject features, i.e. the orientation of the object (left/right), and interobjects features, i.e. the spatial location of the object, under four eye movements' conditions: free viewing, central fixation, congruent viewing (looking at the square where recalled object was originally placed), incongruent viewing (looking at the square where recalled object was not placed). Participants' task was to say "yes" if the statement was true and "no" if it was false. Authors observed spontaneous eye movements to the quadrants where recalled object was originally placed as well as the impact of the central fixation on memory. However, only the memory for the interobject features was hindered as shown by the longer response times. Additionally, both accuracy and response time results indicated memory (in both inter- and intraobject statements) facilitation in the congruent viewing, i.e. when participants were asked to look at the square corresponding with the original location of the object. These results provide a new evidence for the facilitatory role of gaze position, but not for the re-enactment of the scanpath. However, it underlines the importance of the memory for spatial position of the objects. According to the results from Johansson and Johansson's (2014) experiment, eye movements may play a functional role only with regard to the interobject's features (like spatial location) and not intrinsic object features (like the orientation of the object). This finding could explain why some of the previous studies that focused on the retrieval of intraobject features did not observe a significant link between eye movements and recognition performance.

Furthermore, Johansson and Johansson's (2014) findings are in accordance with the study by Postle et al. (2006), which showed that memory for locations was impaired by saccadic distraction task but memory for shape was intact. It supports the vast number of studies investigating dorsal and ventral cortical pathways that relatively independently drive visual memory for objects ('what' information: color, shape, texture) and location memory ('where' information: localization, spatial relationships, spatial transformation) (Gazzaniga, Ivry, & Mangun, 2008; Goodale & Milner, 1992). Additionally, Farah, Hammond, Levine and Calvanio (1988) observed this cortical distinction for visual imagery, where dorsal stream was driving spatial imagery and ventral stream – object imagery.

Recently, the number of studies focusing on the eye fixations with regard to visual memory has been growing steadily. Valuch, Becker and Ansorge (2013) created two experiments with images of natural scenes where they investigated the role of repeated fixations in recognition compared to free viewing task. Results of the first experiment pointed to the bias towards locations that had been fixated when the picture was seen for the first time. Additionally, authors observed stronger bias in the group that got the memory task than in the group that got the free viewing task, i.e. was simply looking at the images. The number of fixations on the repeated parts of the image was significantly higher than on the novel parts of the image. According to the authors of the study, those results show that fixating on the same locations or objects during learning and later recognition can support visual memory for scenes. Interestingly, authors manipulated the stimuli so that some of them showed the same scene from a different perspective and concluded that fixations to the same locations can support the memory across changes of the scene perspective. In a follow up experiment in their study, Valuch, Becker and Ansorge (2013) observed that regions fixated in the learning session triggered faster and more accurate recognition compared to regions of the same pictures but not fixated during encoding. Other studies found the beneficial role of fixating on the previously studied locations during delay before visuospatial working memory retrieval (Olsen, Chiew, Buchsbaum, & Ryan, 2014; Tremblay, Saint-Aubin, & Jalbert, 2006). Interestingly, Olsen et al. (2014) observed the relationship where the greater fixation pattern similarity between encoding and delay, the higher was memory performance in the task. Additionally, in a different study from their research group (Chan, Kamino, Binns, & Ryan, 2011), researchers found out a link between the number of fixations and the level of face recognition performance. According to Olsen et al. (2014), this can indicate a different role of eye movements in memory for spatial locations, where rehearsing fixations occurs as functional strategy to build and maintain the image in visuospatial memory, and for objects, where better encoding demonstrated by the number of fixations determines higher level of recognition performance. However, these studies did not investigate the temporal order of the fixations, which could be an essential component in the reactivation of the memory.

Recently, Laeng et al. (2014) investigated refixations in visual imagery in three separate experiments. Researchers used geometric patterns as well as images of animals as stimuli in the study. In the first experiment, where participants viewed and subsequently imagined equilateral triangles of different orientation, Laeng et al. (2014) observed that participants’ gaze dwelled in the locations corresponding to the shapes or their parts being imagined. Results from the second and third experiment revealed a positive relationship between the similarity of the scan patterns at encoding and retrieval and the quality of long-term memory performance. Additionally, precluding eye movements to the original and relevant for retrieval locations resulted in the loss in accuracy. Similar results were observed by Martarelli and Mast (2013) who used pictures of animals as well as images of imaginary creatures as stimuli. Participants’ task was to memorize the pictures and answer the questions regarding the content of the images either immediately after encoding or one week later. Results indicated significantly longer fixations to the relevant areas where the information was encoded. However, Martarelli and Mast (2013) did not observe any effect of those refixations on the accuracy rates.

Taken together, reviewed studies point to the crucial role of eye fixations to the original locations of objects or parts of a scene in both short-term and long-term memory. According to those findings, repeated eye movements to locations that were fixated during encoding, i.e. perceptual reinstatement, can facilitate memory retrieval. Additionally, recent studies show that the role of eye movements in visual memory might differ in terms of memorizing spatial relations between objects or objects themselves. However, none of those studies investigated the role of the temporal sequence of eye movements in encoding and retrieval.

Experimental tests of scanpath theory

Research on refixations facilitating recognition have mainly focused on eye movements to the same locations, at encoding and retrieval, but have typically disregarded the temporal order of the gaze scan patterns. However, a few studies made an attempt to test the sequential order of the fixations and its impact on recall. Early studies that investigated principles of scanpath theory provide conflicting findings in terms of the relationship between scanpaths' similarities and recognition performance (Locher & Nodine, 1974; Whiteside, 1978). Nevertheless, recent findings give more insights into this relation. In their study investigating fixation sequences, Humphrey and Underwood (2008) compared scan patterns at imagery and recognition both immediately after stimuli presentations and after 2 days. Images used in the experiment were distinctively individual pictures in color. Half of the pictures (i.e. 30 images) were used in the encoding session and labeled as 'old' whereas the other half was shown together with the 'old' pictures in the retrieval session and labeled as 'new'. Thirty participants were divided into two groups: Imagery First and Recognition First. Imagery First group's task was to imagine every picture immediately after it was presented on the screen for 3000 ms and perform a recognition memory task after seeing and imagining all set of stimuli for the first time. Subsequently, they needed to come back after 2 days for the imagery task again. In the delayed imagery task participants saw short, written cues (e.g. "the American Football game") instead of complete image and were asked to imagine a picture associated with particular cue. The Recognition First group saw all 30 'old' pictures in a raw, took a short break and was asked to do a recognition memory task. When the recognition task was completed, participants took a short break and performed the imagery task. This group came again after 2 days for the imagery task as well.

Humphrey and Underwood (2008) analyzed accuracy and eye-tracking data, including fixation durations and saccade amplitudes. Additionally, they made a comparison of the individual scan patters from different conditions (encoding, imagery and recognition) using string editing technique (Brandt & Stark, 1997; Choi, Mosley, & Stark, 1995; Foulsham & Underwood, 2008; Privitera, Stark, & Zangemeister, 2007). The technique segregates the stimuli into labeled regions and in this way it turns a particular sequence of fixations into a string of characters. Subsequently, the similarity between two scan patterns is computed through estimating the minimal number of steps required to transform one string into another (Humphrey & Underwood, 2008).

Such a string analysis revealed that the highest significance of similarity occurred between scan patterns at encoding and recognition as well as between imagery and delayed imagery, i.e. at so-called pure processes comparisons with similar visual input (image or blank screen) in every condition. However, even in mixed processes comparisons, like encoding and imagery, although significance was quite low it was still reliably above chance. According to Humphrey and Underwood (2008), those results support the view that not only bottom-up features of the image but also top-down control drive eye movements at encoding and retrieval, as they don't depend on any visual input. Moreover, bottom-up influence on the eye movements alone cannot explain scanpaths' similarity at encoding and retrieval. In general, Humphrey and Underwood's (2008) study showed that the more similar encoding and retrieval processes, the more similar would the scanpaths be. Additionally, scan patterns were stable over time as shown through the significant similarity of the strings in imagery and delayed imagery condition. Similarly, in a different study with the use of images of natural scenes and people, Humphrey and Underwood (2010) showed that fixation order and pattern was highly similar at encoding and retrieval. However, they additionally observed the lack of the relationship between accuracy at recognition and string similarity what led to the conclusion that re-enactment of eye movements alone did not suffice for the memory facilitation. Several other studies (Harding & Bloj, 2010; Humphrey & Underwood, 2009; Underwood, Foulsham, & Humphrey, 2009) used the same method of string analysis and observed a significant number of scanpath repetitions across multiple viewings, supporting the idiosyncratic character of scanpaths as well.

An interesting study providing new insights into Scanpath Theory was recently conducted by Foulsham et al. (2012), which investigated several issues connected to the theory, in particular to what degree scanpaths are repeatable. This was a follow-up experiment to the study by Foulsham and Underwood (2008) and used its empirical data. Foulsham with colleagues (2012) used a MultiMatch model to align pairs of scanpaths and compare corresponding elements. The scanpath was defined as "the ordered sequence of fixations (…) made in particular image viewing trial" (Foulsham et al., 2012, p. 6). The comparison of the scanpaths between- and within-subjects for the same and different images revealed significantly more similar scanpaths for the same pictures than for different ones as well as significantly more similar scanpaths within subjects in all dimensions (with the exception of the length of saccades). The highest scanpath similarity was observed in the within-participant, within-image comparison. It showed also higher significance comparing to within-participant, between-image similarity and authors interpreted it as a finding that stayed in accordance with Noton and Stark's observations (1971a, 1971b) about the idiosyncratic nature of the scanpaths, i.e. that individuals show repeatable sequences of eye fixations over a specific image. Moreover, as Foulsham et al. (2012) underlined, this effect could not occur due to the visual saliency of the image or the general content of the image as the similarity was higher within subjects comparing to between subjects looking at the same image. However, when authors of the study investigated the relationship between the scanpath similarity and recognition performance, they did not observe reliable correlations. Consequently, Foulsham and colleagues (2012) concluded that their results pointed to the idiosyncratic character of the scanpaths and stayed in contrast with the bottom-up explanation of eye movements during retrieval but the encoding and storage of the scan patterns could not account for the similarity of the scanpaths, what stays in accordance with Humphrey and Underwood's study (2010).

It is important to note that described studies, which calculated scanpath similarities and compared them across different conditions, were simply correlational. Such correlations might have emerged for various reasons and the direction of causality (if it applies) remains unclear. Therefore, Foulsham and Kingstone (2012) proposed recently a different approach where they manipulated eye movements at encoding and retrieval in order to test the impact of scanpath similarity on memory. They designed 5 distinct experiments where scan patterns were manipulated in several different ways. The general research question regarding all 5 experiments was "whether congruency in the scanpaths made during encoding and recognition leads to improved recognition for previously seen ('old') images" (Foulsham & Kingstone, 2012, p. 4). In the first experiment, the sequence of fixations was constrained during encoding but free under recognition. Thus, particular regions of each picture (square "patches") were firstly shown to participants in a specific order and subsequently complete scenes were viewed in the recognition session. Foulsham and Kingstone (2012) hypothesized that the forced scanpath from encoding would be also adapted during retrieval and that this adaptation would be more likely to occur in the correct trials. Results showed that during memory retrieval subjects fixated longer on those regions, which they had seen during encoding and that looking at those regions was associated with correct recognition. Additionally, authors of the study observed that temporal sequence of the patches mattered but only in terms of the first seen regions (which was interpreted as a primacy effect), i.e. patches seen earlier were better remembered than those seen last.

To investigate the reverse effect of their design, Foulsham and Kingstone (2012) created a second experiment where participants were allowed to freely inspect the images during encoding but their eye movements were constrained at recognition. For each participant, the first five fixations from encoding were stored after each trial. During recognition either those regions were shown in an original temporal sequence (respectively to the order of participant's fixations) or patches presenting random regions. Data analysis revealed better accuracy in the trials with the regions fixated during encoding comparing to the trials with the random patches. However, this difference could be due to the meaningless content of the random regions and therefore, Foulsham and Kingstone (2012) conducted a third experiment where patches at memory retrieval were shown either in the original sequence of fixations ('self' condition) or in the sequence made by the same observer but from a different image ('other image' condition). Comparison revealed a recognition advantage for the trials in 'self' condition, i.e. when the original fixations were recreated in the right order. In the fourth experiment authors of the study conducted another comparison, this time between 'self' condition and 'other person' condition. In the new condition each participant was shown patches corresponding to other participant's content and order of fixations. In this way Foulsham and Kingstone (2012) investigated the idiosyncratic character of the scanpaths. Results pointed to the lack of the significant difference between trials in 'self' and 'other person' conditions. Authors of the study interpreted it as evidence against a "strong" interpretation of the scanpath theory. Additionally, it stays in contrast with the other findings pointing to the idiosyncratic character of the scan patterns (Foulsham et al., 2012).

Lastly, in the fifth experiment, Foulsham and Kingstone (2012) compared memory performance in the 'self' and 'self shuffled' condition. The new condition involved presenting patches that were fixated in the encoding session but their order was randomized ('shuffled') at recognition. However, also this comparison did not reveal reliable differences between conditions. Authors of the study interpreted this finding as further evidence for the lack of an observable effect of scanpaths on long-term visual memory.

In a follow up study consisting of three short experiments, authors observed that participants were able to recognize their own fixations with accuracy rate significantly above chance level (Foulsham & Kingstone, 2013). Subjects differentiated best between their own and random fixation locations. However, they managed as well (although with significantly lower accuracy rates) to recognize own fixations when compared with other image's or other person's distribution of fixations. As Foulsham and Kingstone (2013) pointed out, differentiating between own and other person's scanpath might be most difficult task, as people tend to look at similar locations in the image. However, authors tested only for the spatial distribution of the fixations and not for the temporal sequence of the fixations. Therefore, it is difficult to apply an interpretation in terms of the scanpath theory. Furthermore, the decisive evidence for the "scanpath memory" has not been given yet and the existing one remains weak (Johansson, Holsanova, & Holmqvist, 2006). Consequently, the final questions whether eye movements have functional or epiphenomenal role in memory retrieval and to which degree temporal order matters have not been answered yet.

The present study

The aim of the study

The main aim of the study was to investigate the role of eye movements in long-term memory. The additional aim was to design a new experimental procedure that would adequately test for the role of eye movements and their effect on memory retrieval. Most of the studies investigating principles of the scanpath theory tried to find a proper method for measuring similarities between scanpaths at encoding and recognition. Many studies compared scan patterns in a correlational way (e.g. Brandt & Stark, 1997; Humphrey & Underwood, 2008) and only few tried to experimentally test scanpath theory manipulating eye movements in various ways (Foulsham & Kingstone, 2012). Additionally, a vast amount of research used spoken descriptions as auditory cues in imagery and memory tasks (Johansson, Holsanova, & Holmqvist, 2006; Johansson et al., 2012; Johansson & Johansson, 2014; Richardson & Spivey, 2000; Spivey & Geng, 2001). In the current study a different approach was applied. Eye movements were manipulated in a way that either the same or different scanpath was forced during recognition. This procedure allowed us to investigate the relationship between scanpaths and memory performance at retrieval. The purpose of this particular design was to investigate whether the enactment of the specific sequences of eye movements previously shown would facilitate the retrieval from long-term memory. Additionally, our purpose was to contribute to the ongoing discussion about the validity of scanpath theory and functionality of eye movements in encoding and retrieval.

Research questions

Given the mixed evidence in the literature but anyway reliable findings for the crucial role of eye movements in memory retrieval, we hypothesized that the re-enactment of the scanpaths from encoding does facilitate memory retrieval. Based on that, two main research questions were stated for the current research: “Do eye movements play a functional role in the retrieval of the long-term visual memory for grid patterns?” and “Can long-term visual memory performance be facilitated through the enactment of the eye scanpath?”.

On the grounds of the research questions, three hypotheses were formulated. Firstly, in the present study, we hypothesized that memory performance in the recognition session will be significantly better when presented with the same scanpath, i.e. the same sequence of eye fixations. According to the main hypothesis, it was expected that a) accuracy rates would be higher and response times would be faster when subjects would see the pattern in the recognition session in the same sequence as their original scanpath from the learning session, b) this difference would occur only in the free viewing condition, where participants were allowed to move their eyes and not in the forced fixation condition. Secondly, it was hypothesized that an external measure of spatial memory, like the Corsi Block-Tapping Task (Lezak, Howieson, & Loring, 2004, p. 355), would correlate with the results in the grid’s memory task. Lastly, pupillary responses were expected to reflect the level of cognitive effort in the task. Hence, we hypothesized that if maintaining central fixation results in additional cognitive load in working memory, then we will observe a significant difference in pupillary changes between forced fixation and free viewing conditions. In other words, pupil diameters should be larger in the former than in the latter condition.

Methods

The pilot study

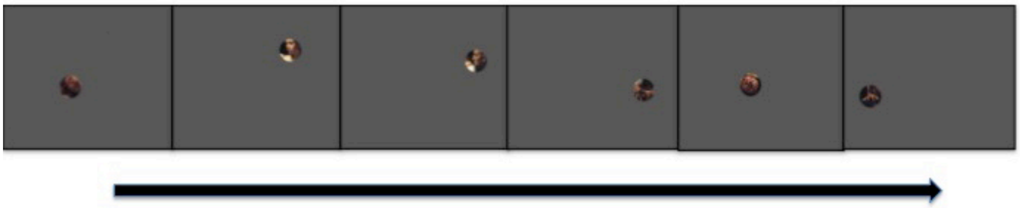
Considering the exploratory character of the research, a thorough pilot study was conducted to develop the proper method. Firstly, images of historical paintings were chosen as stimuli with regard to their complexity and proportionality of the details spread within the space of each picture. The participants' task was to try to memorize the pictures and then attempt to recognize them after some of the details that they had fixated while memorizing were shown again in isolation, either in the same sequence in which they were observed or in a different sequence. However, this test did not show the expected results although there was a trend in the expected direction. We reasoned that these stimuli might have been too easy to memorize and recognize. Therefore, more challenging stimuli were designed and the procedure for the recognition session changed accordingly. Specifically, images of grids (resembling irregular checkerboards) were used and after minor adjustments in the short pilot phase, they were also applied in the main study.

Pilot experiment 1: Paintings

Participants. Thirty-eight subjects were tested in two parts of the pilot study with the paintings. Twenty-five subjects were tested in the first part and subsequently, 13 subjects were tested after particular changes were applied to the experimental design. Participants were at the time of the testing current students at the University of Oslo.

Stimuli and procedure. Images of historical paintings from the XVI and XVII century were collected from "Wikimedia Commons", an online database of freely usable media files, and used as stimuli (see Appendices, Appendix 1 for the complete set of stimuli). Twenty pictures were chosen and paired in terms of the approximate number of the characters, the size of the characters (with the focus on their faces), the colors in the picture and its luminance (see Figure 1). Ten images were used in the learning session and for the purpose of the analysis labeled as 'old' whereas their pairs were labeled as 'novel' and used together with the 'old' pictures in the recognition session. Participants did not see or know about the labels. Therefore, learning session consisted of 10 pictures of paintings that participants were asked to memorize and a recognition session of all 20 pictures that were set up in the random, but fixed order for every participant. Before each picture appeared a baseline image that was adjusted to the luminance of the test picture. This baseline consisted of a blank gray field that included a fixation cross in one of the corners. The stimuli for the recognition session were additionally manipulated, so that most of the image was covered in blank grey color (adjusted to the luminance of the painting) and only the particular details were uncovered one after one in a circular window (see Figure 2). The manipulation was done by means of PowerPoint presentation software and the test stimuli were saved as video files (.wmv). Crucially, the sequence of the details in the video was dependent on the sequences of participants' fixations in the learning session. Therefore, each recognition session was individually adjusted for every participant and no sequence was exactly the same from a participant to another. The details were shown either in the same sequence as participants looked at them in the learning session, or in the reversed sequence. Thus, either the same or reversed scanpath was evoked. Additionally, participants were divided into 2 groups for the

Figure 1. An example of the pair of paintings.



A. Sequence of the video

B. Corresponding complete picture

Figure 2. A) An example of the sequence of details viewed in the recognition session and B) Corresponding complete picture.

recognition session. One group was allowed to freely look at appearing details of the paintings ('free viewing' group) whereas the second group was asked to maintain a central fixation in the middle of the screen while the details were appearing ('forced fixation' group).

The participants' task in the recognition session was to respond with the 'B' key if they recognized a painting or the 'N' key if they did not recognize the painting. Testing took place in the cognitive laboratories at the Psychology Department, University of Oslo. Eye-tracking data was recorded by Remote Eye-Tracking Device from Senso-Motoric Instruments (SMI, Berlin, Germany). The gaze position was sampled at a rate of 60 Hz. Each participant performed a standard calibration procedure before both learning and recognition session. The delay between the sessions was set to 48 hours.

Changes and adjustments. After testing the first 25 participants, on the basis of the experimenter's observations and on the participants' feedback, a new sequence was implemented. Reversed sequence was replaced by shuffled (mixed) sequence, where the order of participant's fixations from the learning session was random. In this way, enacted scanpath was either the same as in the learning session or completely different instead of just reversed. An additional group of participants ($N = 13$) was tested after adjustments. The rest of the procedure was kept identical to the first one.

Data analysis and results. A 2x3 ANOVA analysis was conducted with Condition (free viewing or forced fixation) as a between-subject factor and Sequence (same, reversed or shuffled, novel) as a within-subject factor on the mean accuracy percentages first for the 25 participants with 'reversed' sequence and subsequently for the 13 participants with 'shuffled' sequence. A similar but separate ANOVAs were also performed on the mean response times. These analyses did not reveal any significant effects. However, some difficulties with the stimuli and procedure were noticed. Participants tended to respond before the sequence was completed on the basis of one or two details seen at the beginning of the trial and therefore not on the sequence, i.e. enacted scanpath. The average accuracy for all participants equaled 67.12% with standard deviation of 46.99. Therefore, enabling response while the details were still appearing on the screen caused differences across participants in terms of the seen number of the details. Thus, the experiment failed to test for the effect of the scanpath's enactment. Therefore, both the stimuli and the experimental procedure were modified and a second pilot study was conducted.

Pilot experiment 2: Checkerboards

Stimuli and procedure: The images of paintings were replaced by the black and white images resembling checkerboards so as to eliminate the presence of meaningful information details like faces, objects, colors and other details. Further, the following adjustment was based on the assumption that, by memorizing geometrical grid patterns, participants would need to rely on the spatial locations of particular black squares and not on the shape or color of the squares themselves. The procedure was kept similar to the foregoing with the use of paintings. Forty pictures were used in the learning session and each image of the grid was shown for 5000 ms on the screen. Viewing conditions in the recognition session (forced fixation and free viewing) were organized in two blocks and presented in the fixed pseudo-random order to every subject (within-subject design). Within each condition, three sequences were used: ‘same’, ‘different’ (i.e., shuffled) and ‘novel pattern’. Black squares were appearing one after another on the white grid, forcing the same or different sequence of fixations as the sequence at encoding. The presentation time of each black square decreased compared to the presentation time of the paintings’ details in order to reflect better the original time of the main fixations at encoding. For the recognition session, 80 images were presented with 40 ‘old’ patterns (in the ‘same’ or ‘different’ sequence) and 40 ‘novel’ patterns. The participant’s task in the recognition session was to press ‘M’ key on the keyboard if the pattern was recognized and ‘Z’ key if it was not recognized.

Changes and adjustments: Preliminary results from first 5 subjects revealed too high level of difficulty of the task based on the mean accuracy results that were at chance (50%). Therefore, the number of the stimuli was reduced leaving 32 images in the learning session and 64 in the recognition session. Additionally, the viewing time in the learning session was increased from 5000 to 8000 ms, compared to the previous pilot study. Moreover, at encoding, participants were asked to imagine the pattern each time the image of the grid disappeared and the blank, grey screen was presented. Imagination phase was added to the task to keep participants concentrated and memorize better. After imagery phase participants were deciding when to see the next image by pressing a spacebar on the keyboard. The rest of the procedure was kept identical to the preceding. Subsequently, additional 4 subjects were tested to control the new procedure that was also used in the main experiment.

The current study

Participants

Twenty-eight participants were tested in 56 testing sessions. Data from 4 participants was excluded due to the technical issues, problems with vision or very poor accuracy that pointed to the lack of a proper engagement in the learning (or encoding) session. The rest of 24 subjects were right-handed students with normal or corrected to normal vision. The mean age of the participants was 26.73 years (SD = 7). Among these 24 participants, there were 15 females and 9 males. Three of the participants recently completed higher education whereas the rest was current students.

Recruitment for the experiment took place at the Institute of Psychology, University of Oslo, as well as on the Internet by means of social media. Subjects were asked if they are interested in participation and in the case of a positive answer, a written detailed description of the experiment and requirements was provided. The experimenter stressed that only subjects with normal or corrected to normal vision could be enrolled. Additionally, all participants were aware that no payment was provided for the participation in the experiment. Immediately before the first experimental session, participants signed a Consent Form (see Appendices, Appendix 2).

Stimuli

Sixty-four images of black and white grids resembling checkerboards were used in the study (see Appendices, Appendix 3 for the complete set of stimuli). Each image was a 5 x 5 grid where four black squares (RGB values: 0, 0, 0) formed a random pattern and the rest of the squares remained white (RGB values: 255, 255, 255) (see Figure 3). All patterns were designed and produced by the author for the purposes of the current study. Thirty-two images were randomly chosen to use in both the learning and the recognition session and labeled for the purpose of the analysis as ‘old’. Participants did not see the labels and did not know about them. Those were the images that participants were to memorize. The other half of the images were labeled as ‘novel’ and used together with the ‘old’ pictures in the recognition session as foils.

The images of the grids had a constant luminance to which a gray baseline (RGB values: 190, 190, 190) with a white fixation cross (RGB values: 255, 255, 255) was adjusted. The fixation cross (1° of visual angle) appeared randomly but equally in one of the four corners of the screen. After each image of a grid, a blank gray picture appeared (RGB values: 142, 142, 142). It had the same RGB values as the gray background used in the images of the grids.

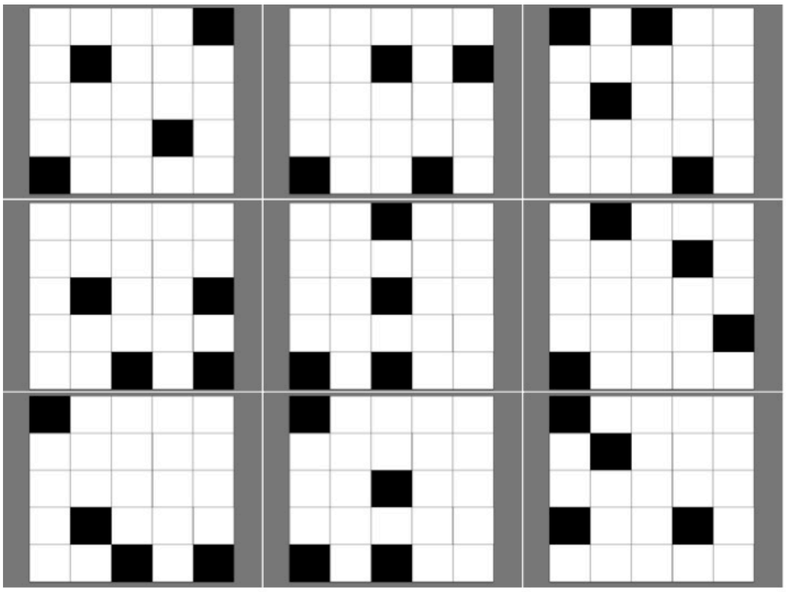
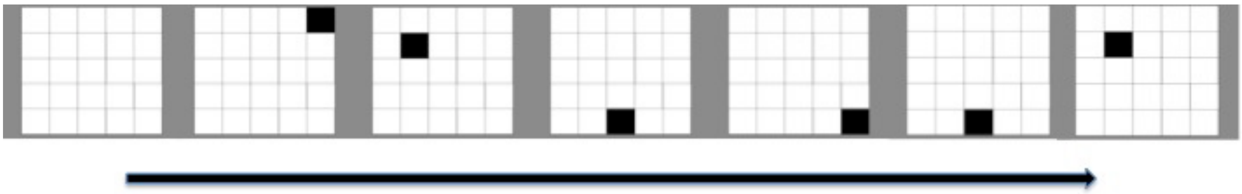


Figure 3. Examples of the stimuli used in the current experiment.

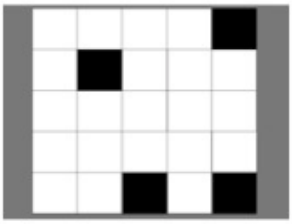
All pictures had resolution of 1680 x 1050 pixels. The size of the grid on the screen was 30.5° x 30.5° of visual angle. The size of a single square within the grid was 6.24° x 6.24° of visual angle (see Appendices, Appendix 4).

The stimulus set-up differed in the learning and recognition session. In the learning session, when participants were to memorize the patterns, the complete static images were shown. In the recognition session the patterns were shown as videos, where each of the black squares appeared on a white grid immediately after the previous one disappeared. The videos were created in Microsoft Office PowerPoint software, where a 7-slides presentation (1 slide of a white grid and 6 slides of single black squares on the grid) was saved as a movie file (.wmv). In this way each pattern was shown in a particular sequence, which was based on the participants' original fixations from the learning session (see Figure 4). The timing for each slide with a single black square was set to 500 ms and the total time of the video was 3000 ms. The resolution of the video was kept the same as the resolution of the images: 1680 x 1050 pixels. The size of the squares and the whole grid did not differ from their sizes in the images.

The order of the black squares appearing in the videos was directly dependent on the participant's fixations from the learning session, when the picture was memorized. For each participant the sequence of the main 6 fixations was extracted by the means of SMI BeGaze® analysis software separately for each picture. It means that every picture with an individual scanpath was extracted and saved as a movie file (.avi). Following the order of fixations in the extracted video, one by one, a PowerPoint presentation was made, i.e. the slides with single black squares were put in the right sequence and saved as a video file that was later used as a stimuli in a recognition session.



A. Sequence of the video



B. Corresponding complete pattern

Figure 4. A) An example of the sequence of the video from the recognition session and B) Corresponding complete pattern.

Apparatus

The data was collected by means of a Remote Eye-Tracking Device from Senso-Motoric Instruments (SMI, Berlin, Germany). This eye-tracking device has an automatic compensation for head movements at a 70 cm distance and in a range of 40 x 20 cm. However, a chinrest was additionally used in this study to keep the participant's head as stable as possible. The distance of the chinrest from the screen was 60 cm. The distance of the screen from participant's cornea when the head was placed in the chinrest was 55 cm. The lightning in the room was kept stable throughout the whole experiment.

The frequency of the samples of the eye position was 120 Hz in the learning session and 60 Hz in the recognition session. This system is accurate within 0.03° of visual angle regarding the spatial resolution. Additionally, a fixation is detected automatically when the gaze dwells in minimum 80 ms within a region of maximum 100 pixels, following an algorithm applied in the system.

Stimuli were presented in a fixed pseudo-random order to the participants by the use of iView 3.0® Experiment Center' software. They were shown on a color, flat Dell LCD monitor. The size of the monitor was 47 cm. The resolution of the screen was set to 1680 x 1050 pixels.

Experimental design

The experimental design of the current study was a 2x3 mixed factorial design with viewing Condition (free viewing or forced fixation) and the Sequence of original scanpaths' fixations (same sequence, different sequence or novel pattern) as independent variables. Additionally, it was a within-subject design, where each participant saw the equal number of known patterns appearing in the same or different sequence and novel patterns in both free viewing and forced fixation condition.

To minimize the effect of the patterns' order on memory, the pictures in both learning and recognition sessions were shown to the participants either in the original or reversed order in counterbalanced conditions (original-original, original-reversed, reversed-original, reversed-reversed). The number of participants in each of these conditions was the same.

Procedure

The testing sessions took place in the Cognitive Laboratories at the Institute of Psychology, University of Oslo. iView R.E.D. tracking system and iView 3.0® Experiment Center' software were used for data collection and stimulus presentation. Each participant was asked to read and sign a Consent Form (see Appendices, Appendix 2) before the experiment started. Additionally, participant's biographical data including age, sex and handedness (by self-report) were collected. Subjects were informed that they needed to participate in two experimental sessions with around 48 hours delay. Participants were told that the purpose of the experiment was to investigate long-term visual memory for geometrical patterns.

Learning session. The experiment started with the learning session where the task was to memorize and imagine 32 pictures of grids resembling checkerboards. Participant was asked to sit in front of the stimulus monitor while instructions were given. Right before the experimental procedure started, participant placed his/her head on a chinrest and was instructed to sit possibly stable and keep his eyes open during both viewing and imagery phases. A standard calibration procedure was performed before the images were displayed.

During the calibration procedure, participants were asked to follow carefully a red dot in a white circle that moved to four locations on the screen, forming a 2 x 2 matrix. When the calibration was successful (i.e. deviation was lower or equal 0.5), detailed instructions appeared on the screen. Participant decided himself to start the experimental trials by pressing a spacebar on the keyboard. Learning session consisted of 32 trials. Each trial started with presentation of the fixation cross (triggering, i.e. staying on the screen until participant's gaze dwelled on the cross for 500 ms) followed by the image of a grid (8000 ms) and a grey, blank screen when participant was imagining seen pattern (see Figure 5). The time for the imagery phase was not limited. Participants were asked to try to imagine the just-seen grid as clearly as possible and then go further to the next picture by pressing a spacebar on the keyboard. The average time of the learning session was 9 minutes. The order of the trials in this session was reversed for a group of participants (*N* = 12, females = 6).

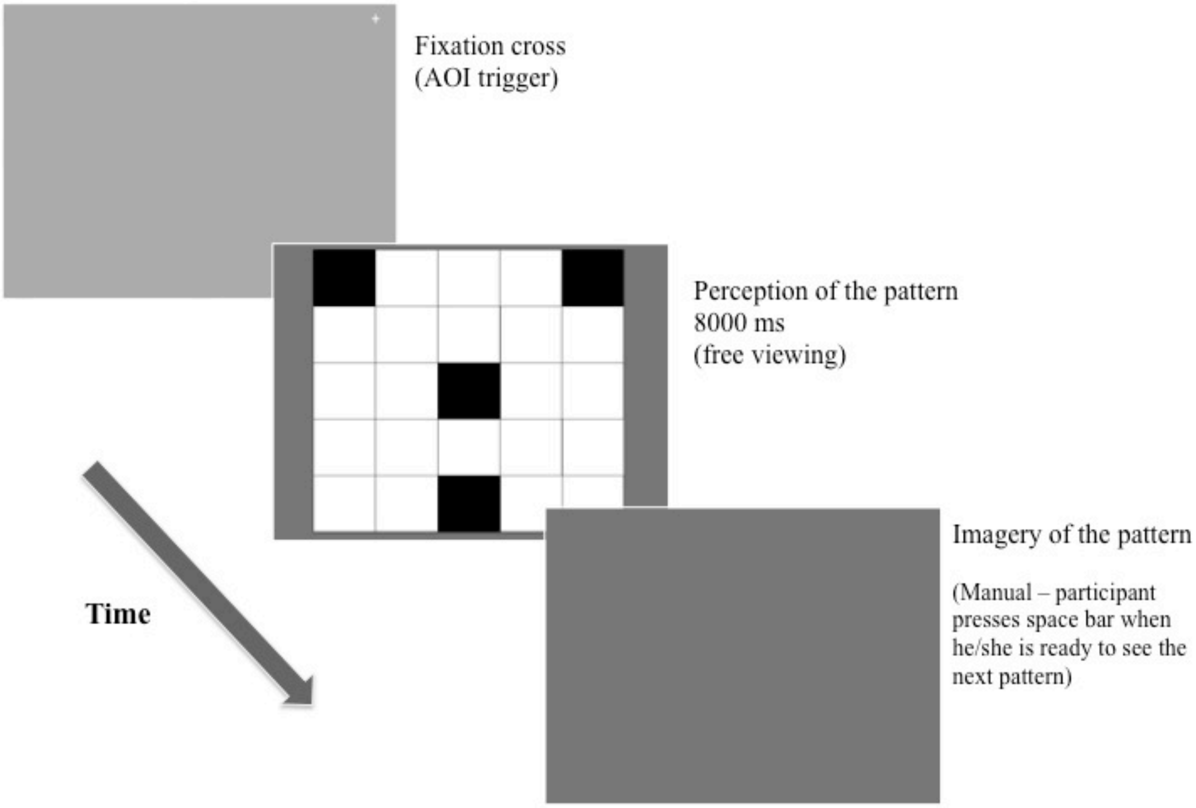


Figure 5. Learning session procedure with perception and imagery phase.

Recognition session. The second testing session, called the recognition session, took place within a 48 hours delay from the learning session. The following time-frame was set arbitrarily but studies on visual long-term memory capacity show that pictures' details can be remembered over several days (Brady et al., 2008).

Each participant came back to the same laboratory and was again asked to sit in front of the stimulus monitor and place his/her head on the chin rest while the instructions were given. The same calibration procedure as in the learning session was applied. Calibration was again followed by the detailed instructions on the stimulus screen. When the participant was ready, he/she pressed a spacebar on the keyboard and the practice trial begun. After the training, participant decided again by pressing the spacebar whether he/she had a sufficient understanding of the task and would like to start the first trial. Each trial presentation was triggered by pressing the spacebar on the keyboard.

This recognition session consisted of 64 trials where 32 pictures from the learning session ('old') were shown again mixed with the other 32 pictures that were new to participant ('novel'). Old and novel patterns were presented in a fixed, randomized order. In the one half of the old pictures (16 images) the pattern appeared according to the original sequence of participant's fixations from the learning session ('same' sequence); in the other half (16 images) the sequence was randomly mixed ('different' sequence). Thus, participant's original scanpath was either enacted or not through forcing the same or different sequence of fixations. Additionally, the whole session was divided into 2 blocks – free viewing and forced fixation. In the free viewing block (16 'old' pictures with equal number of the 'same' and 'different' sequence of patterns and 16 'novel' pictures) participants were allowed to freely follow appearing squares with their eyes. In the forced fixation condition (16 'old' pictures with equal number of the 'same' and 'different' sequence of patterns and 16 'novel' pictures) participants were asked to fixate their gaze in the middle of the image each time there was a picture of the grid on the screen and follow appearing squares with covert attention. Half of the participants started with the free viewing condition, the other half with the forced fixation condition. Additionally, the order of the trials in the blocks was reversed for the part of the subjects that had got reversed order in the learning session (*N* = 6, females = 4) and part of the subjects that had not got reversed order in the learning session (*N* = 6, females = 5).

In the recognition session each trial consisted of a fixation cross presentation (1500 ms), a video presenting a pattern in a particular sequence (3000 ms) and a gray, blank screen with no limited time (see Figure 6). The participant was asked to respond each time when the image of the grid disappeared and the gray, blank screen appeared, by pressing the 'M' key on the keyboard when he/she thought that the pattern had been shown already in the learning session or 'Z' key when he/she thought that the pattern had not been seen earlier. The time for the response was unlimited but the participants were informed that both accuracy and the time of the response matters in the task. No feedback was given to the participant regarding accuracy of the responses. The average total time of the recognition session was 10 minutes.

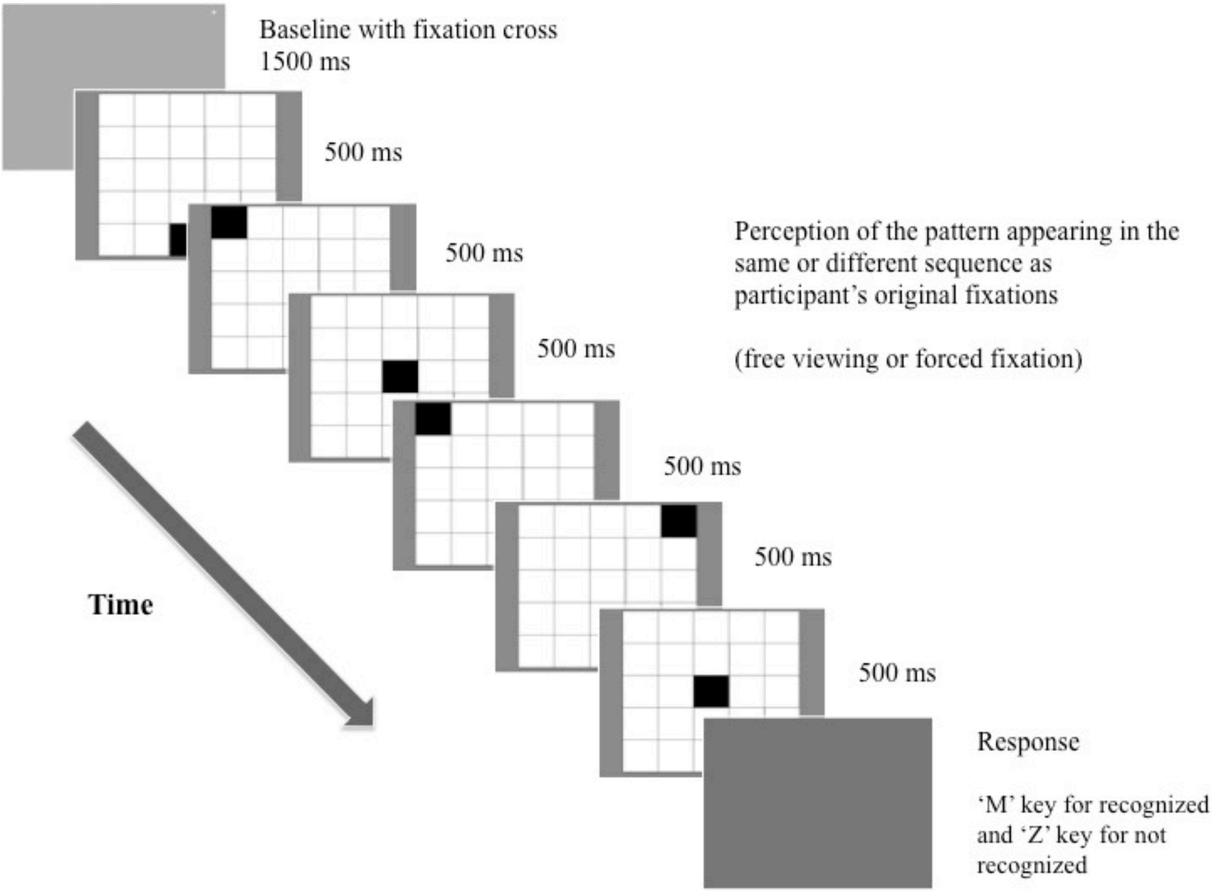


Figure 6. Recognition session procedure with perception of the pattern and response.

The Corsi Block-Tapping Task. When the recognition session was over, participants were subsequently asked to take part in additional task – The Corsi Block-Tapping Task (see Lezak, Howieson, & Loring, 2004, p. 355). The test was applied as an additional measure of individual differences in visuospatial memory. It consisted of nine black cubes (30x30x30 mm) on a black board (225x205 mm, see Appendices, Appendix 5A for detailed measurements). The procedure applied in the current experiment closely followed the guidelines of the standard Corsi Block-Tapping Task (Kessels et al., 2000). Participants were asked to sit in front of the examiner and watch her tap particular number of cubes in particular sequence and then immediately repeat the tapping in the same serial order. The numbers on the cubes were visible only for the examiner to follow the standardized sequences, and not for participants (see Figure 7). Subjects were informed that the number of tapped cubes would increase each time. They were not informed that there were always two consecutive sequences within the same number of cubicles (i.e. of equal length) and that when they made a mistake at the first one, the second one was performed anyway. When they made a mistake in both sequences of equal length, the test was terminated. For a full list of block sequences, see Appendices, Appendix 5B.

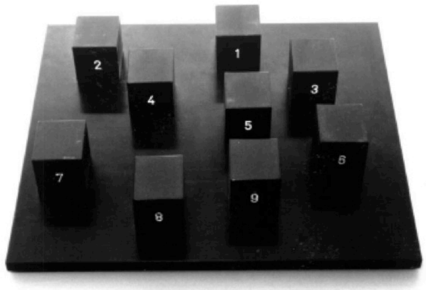


Figure 7. The black board with the numbered cubes used in the Corsi Block-Tapping Task (source: Kessels et al., 2000).

Results

The accuracy and response times data as well as pupillometry data were extracted by the means of SMI BeGaze® analysis software. The standard statistical software (Microsoft Excel, IBM SPSS Statistics and StatView) was used for computing and analyzing the data.

Behavioral results

The grids' memory task

Computing accuracy data. Participants' responses were extracted by the means of SMI BeGaze® software and organized in Excel. For each correct response ('M' key press for the 'old' pattern or 'Z' key press for the 'novel' pattern) a value of 100 points was ascribed and for each incorrect response ('Z' key press for the 'old' pattern or 'M' key press for the 'novel' pattern) a value of 0 points was ascribed. Subsequently, a mean percentage score for accuracy was obtained for each subject for 3 different sequences (same, different and novel) in 2 conditions (free viewing and forced fixation).

Analyses of the accuracy data. A repeated-measures 2x3 ANOVA for the mean percentages of accuracy was conducted with Condition (forced fixation and free viewing) and the Sequence (same, different or novel) as within-subjects factors. Analysis revealed a main effect of the Sequence in the task, *F*(2,22) = 4.824, *p* = .012. Further analysis with the paired samples t-tests revealed significant difference between 'same' and 'different' sequences in free viewing condition, *t*(1,23) = -2.395, *p* = .025, but not within forced fixation condition (see Figure 8). The average accuracy results in free viewing condition for the patterns with the 'same' sequence were significantly above chance, *t*(1,23) = 2.635, *p* = .015, whereas those in the trials with the 'different' sequence were at chance level, *t*(1,23) = .239, *p* = .813. Paired samples t-tests did not show any significant differences between forced fixation and free viewing conditions separately in 'novel', 'same' and 'different' sequences.

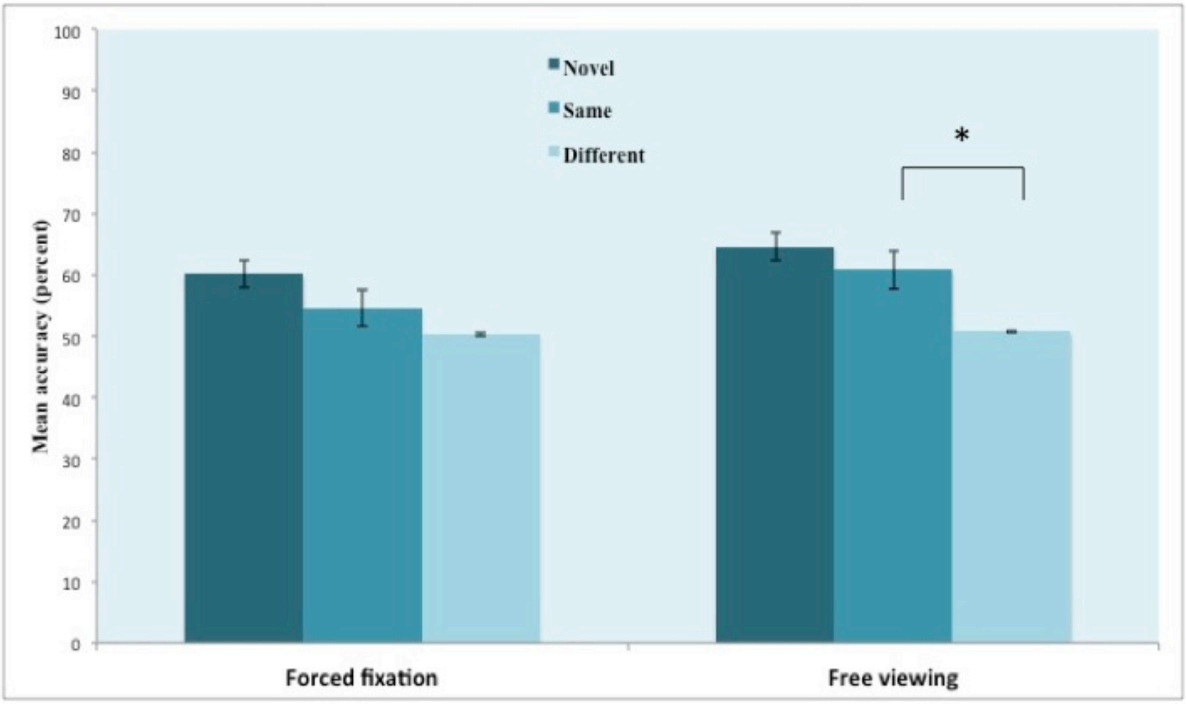


Figure 8. Mean accuracy results in forced fixation and free viewing conditions with the significant difference between 'same' and 'different' sequence of the scanpath within one of the conditions.

Computing Response Times data. The exact times of the key-presses were extracted by the means of SMI BeGaze® software and organized in Excel. Response times (RTs) for the incorrect responses were excluded from the analysis. The data was controlled for possible outlying RTs. The times that exceeded 3 standard deviations from the mean were excluded from the analysis. The average values of the remaining response times were calculated for each subject for 3 different sequences (same, different and novel) in 2 conditions (free viewing and forced fixation).

Analyses of the Response Times data. Similarly as for the accuracy data, a repeated-measures 2x3 ANOVA for the mean RTs was performed with Condition (forced fixation and free viewing) and the Sequence (same, different or novel) as within-subjects factors. Analysis revealed again a main effect of the Sequence in the task, *F*(2,22) = 5.707, *p* = .01. Likewise, further analysis with the paired samples t-test was conducted for the mean response times and revealed significant difference between 'same' and 'different' sequence in free viewing condition, *t*(1,23) = 2.109, *p* = .046, but, similarly as for the accuracy data, not in the forced fixation condition (see Figure 9). Paired samples t-test analysis revealed also a significant difference between 'novel' and 'same' sequences within free viewing condition, where participant's responses were faster for known than for novel stimuli, *t*(1, 23) = 3.954, *p* = .001. Additionally, free viewing and forced fixation conditions did not differ within 'novel', 'same' and 'different' sequences.

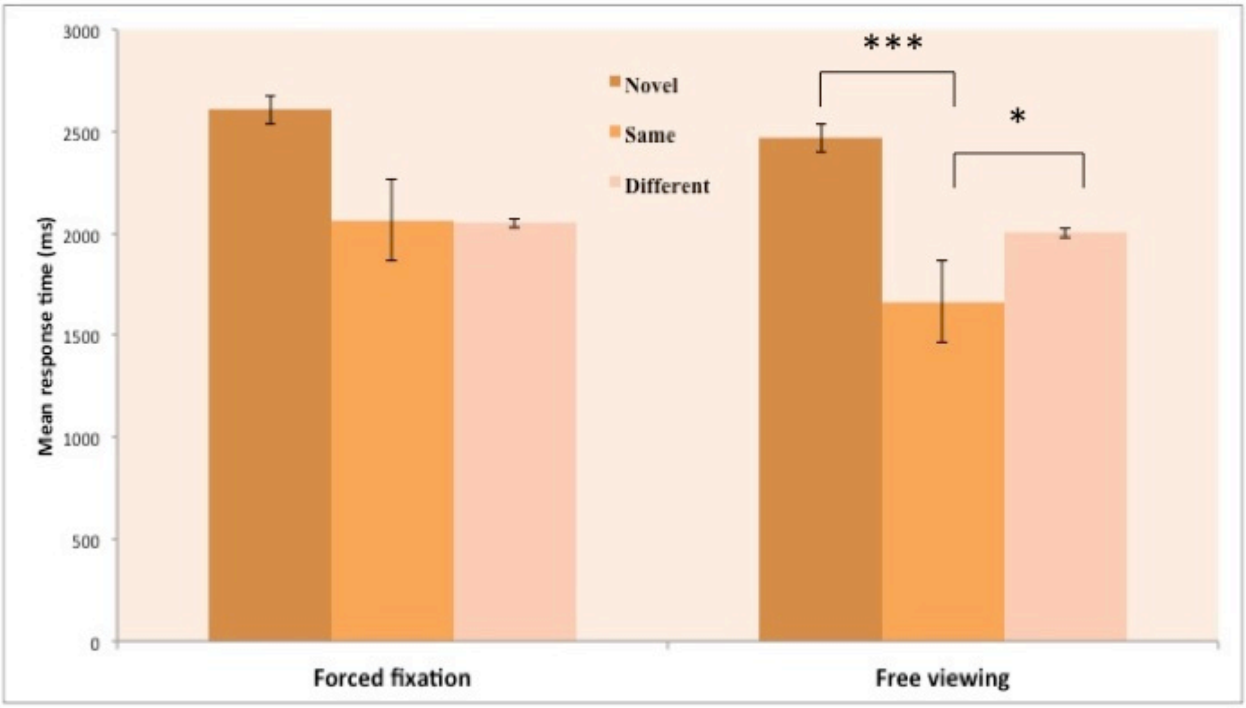


Figure 9. Mean RTs results in forced fixation and free viewing conditions with the significant difference between 'same' and 'different' as well as 'same' and 'novel' sequences within one of the conditions.

The Corsi Block-Tapping Task

Computing the Total Scores. The results for the Corsi Block-Tapping task were calculated according to standardization of the task (Kessels et al., 2000). Firstly, the basic score (Block Span) was computed for each participant. It was the last correctly repeated sequence before the test was terminated, i.e. if the participant made a mistake in both sequences with 7 cubes, his Block Span equaled 6. Subsequently, a Total Score was computed as the product of the number of correct trials and the Block Span, i.e. if participant's Block Span equaled 6 and he/she performed correctly on all the trials until the second one with the lengths of 6, then his Total Score equaled 60 (6 x 10 trials). See Figure 10 for the distribution of the Total Scores across participants.

Linear regression analysis. A simple linear regression analysis was performed on Total Scores from the Corsi Block-Tapping task and mean accuracy percentages as well as mean RTs from the grids' memory task. These analyses did not show significant correlations between the results from the Corsi Block-Tapping task and accuracy or RT data.

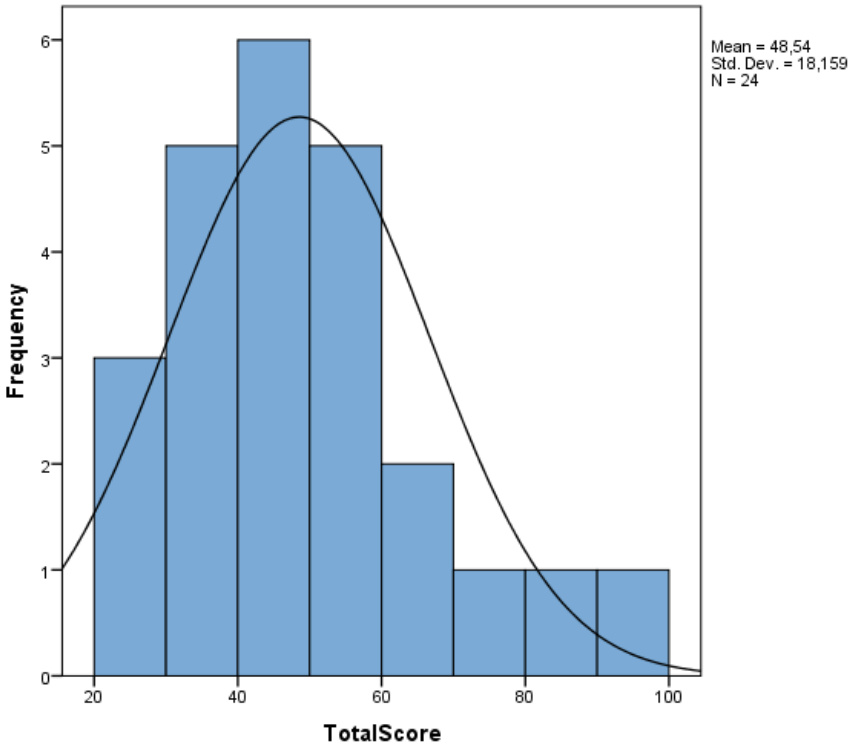


Figure 10. The distribution of the Total Scores in the Corsi Block-Tapping Task (N=24).

Gender differences across the tasks

A post-hoc comparison of the mean results for males and females was conducted for accuracy and RT data as well as for Total Scores from the Corsi Block-Tapping task. Repeated-measures 2x3 ANOVA with a Condition (forced fixation and free viewing) and Sequence (same, different or novel) as within-subject factors and Gender as a between-subject factor did not show any significant effects of gender on the mean percentages of the accuracy in the task with the grid patterns. Similarly, a repeated-measures 2x3 ANOVA with Condition (forced fixation and free viewing) and Sequence (same, different or novel) as within-subject factors and Gender as a between-subject factor on the mean RTs did not reveal reliable effects of either factor on females' and males' performances. Additionally, one-way ANOVA was conducted on the Total Scores from the Corsi Block-Tapping task with Gender as a between-subject factor. The mean score for females equaled 50.08 (SD = 20.02) and for males 50.56 (SD = 17.14). The analysis did not reveal significant gender differences in this task, *F*(1,23) = .004, *p* = .95.

Pupillary results

Computing pupillary changes. The pupillary data (in pixels) from the recognition session was extracted by the means of SMI BeGaze® software and organized in SPSS with separate values for the baseline images and videos of the patterns. The values were aggregated to obtain a one average value for every baseline image and video with the pattern. Subsequently, a difference between baselines and videos' values was computed as a new variable. Thus, pupillary changes were calculated for each trial in the task.

Analysis of the pupillary changes. We conducted a repeated-measures 2x3 ANOVA with Condition (forced fixation and free viewing) and Sequence (same, different or novel) as within-subject factors on pupillary changes. The analysis revealed a significant effect of Sequence on pupillary responses, *F*(2,22) = 3.486, *p* = .048. Pairwise comparisons (by t-tests) indicated that the main effect appeared due to the results in the forced fixation condition. We reasoned that the changes in pupil size in the forced fixation condition might have been affected by the frequency of the black squares appearing in the middle of the screen where the gaze was fixated. That is, pupillary changes could differ significantly in the trials where the gaze was fixated only on a white square comparing to the trials where the gaze was fixated on the square that turned black once or more. This hypothesis was based on the fact that each square covered around 6° of visual angle and the amplitude of pupillary light reflexes within the region of 15° of visual angle is greatest at the center region which is of about 6° (Mizukawa, 2009). Therefore, we computed the frequency values of the black squares appearing in the middle of the grid for all trials from the forced fixation condition (see Table 1). Calculations revealed that the middle square of the grid turned black 0, 1, 2 or 3 times during stimulus presentations, depending on the pattern at the image and on the participants' original scanpaths.

Frequency values	**0**		**1**		**2**		**3**	
Pupillary changes *(in pixels)*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
	-0.204	0.423	0.131	1.328	0.112	0.601	0.006	0.422

Table 1. Mean pupillary changes for the trials with different frequencies of central appearance of black squares (forced fixation condition).

A simple linear regression analysis was conducted with frequency values (0, 1, 2 or 3) as predictors and pupillary changes as dependent variables. The analysis showed, in line with our hypothesis, a significant correlation between the frequency and the pupillary changes in the forced fixation condition, *F*(1,23) = 9.802, *p* = .002. Subsequently, we disregarded all trials with the frequency values of 1, 2 or 3 and conducted again a repeated-measures 2x3 ANOVA with Condition (forced fixation and free viewing) and Sequence (same, different or novel) as within-subject factors on pupillary changes. This analysis did not reveal any significant effects of Condition, *F*(1,23) = 0.898, *p* = .353 or Sequence, *F*(2,22) = 2.868, *p* = .079 on pupillary changes. There was no significant interaction of Condition*Sequence, *F*(2,22) = 2.276, *p* = .126.

Discussion

The current study showed that retrieval from long-term memory was indeed facilitated when the original scanpath from encoding was enacted by the stimulus sequence. This facilitation was indicated by a significantly better accuracy and shorter response times in the case of the same sequence of fixations than in the mixed sequence. Moreover, such an effect was observed, in accordance with our hypothesis, only in the free viewing but not in the forced fixation condition. Thus, a facilitating effect of the scanpaths appears only when the eyes can move and not when they are restricted. This result supports the hypothesis about the functional role of eye movements (Laeng & Teodorescu, 2002). Additionally, it stays in accordance with the research focusing on the crucial role of refixations to the original locations of the objects from encoding (e.g. Foulsham et al., 2012; Holm & Mäntylä, 2007; Johansson Holsanova, Dewhurst, & Holmqvist, 2012; Johansson, Holsanova, & Holmqvist, 2006; Johansson & Johansson, 2014; Laeng et al., 2014; Mäntylä & Holm, 2006; Valuch, Becker & Ansorge, 2013).

Importantly, according to Kosslyn (1994) eye movements during recall of visual information could reflect the occurrence of specific jumps in the visual buffer that let the subject inspect the image. In this way, the movement of the eyes would be subordinated to such shifts of inspection and not actively involved in the retrieval of the memory trace. However, it seems unlikely to be the case here as in the current experiment the sequence of the eye movements was forced and not spontaneous. Hence, the observed result of a facilitation of memory by the same sequence of fixations is in line with an account where eye movements occur to construct the whole images from the elemental "part images" and are not just epiphenomenal. At the same time, it supports Neisser's (1967) and Hebb's (1968) accounts that laid the foundations of the Scanpath Theory proposed by Noton and Stark (1971a, 1971b) and Brandt and Stark (1997).

While most of the previous studies compared scanpaths at encoding and retrieval in a correlational way, the current experiment manipulated (or "perturbated") eye movements in such a manner so as to expose the causal direction of the relationship between gaze scanpaths and recognition. Thus, the results of the current experiment are in contrast with Foulsham and Kingstone's (2012) findings that indicate a lack of memory facilitation in the case of repeating scanpath sequence. This difference can be explained on methodological grounds. In the current experiment, participants were instructed to attend to all of the appearing squares and only after the last element had appeared to subsequently respond whether they had recognized the pattern or not. Thus, the complete scanpath was enacted before response. In contrast, participants in Foulsham and Kingstone's (2012) experiment decided after each element (or "patch") if they wanted to see the next one and, additionally, they were allowed to respond if they recognized the scene already after one or two patches. As each patch corresponded to one original fixation from the encoding session, the complete scanpath was rarely enacted, in fact it was discouraged. Taken together, participants in Foulsham and

Kingstone's (2012) study might have entirely based their recognition of a particular scene on the basis of single details and not the reinstatement of the scanpath.

Johansson et al. (2012) also pointed to the important issue of the task difficulty in their critique of the Laeng and Teodorescu's (2002) study. They stressed that simple, geometrical and not complex stimuli might not reveal the reliable effects of eye movements in visual memory. Therefore, they applied special, more complex stimuli than those used in the previous experiments (see Laeng & Teodorescu, 2002; Richardson & Spivey, 2000). In contrast with this critique, Laeng and Teodorescu (2002) did observe significant results that pointed to a functional role of eye movements. Furthermore, the current experiment also used what apparently would seem very "simple stimuli" made of grid patterns and nevertheless revealed the effect of eye movements on memory. On the other hand, based on the level of difficulty in the grid's memory task, it is clear that the simpleness of the elements is deceiving, and whole patterns were difficult to remember. According to Johansson et al. (2012), during demanding task eye movements may be more active. They put forward an account where the more difficult the visual task, the more eye movements are executed during recall. Hence, the more support from the motor system is required to bring back and form the inner image of previously seen object or a scene. Accordingly, the number of eye movements declines with practice during imagery recall and while looking at a blank screen (Scholz, Melholrn, Bocklisch, & Krems, 2011). Thus, eye movements would play a functional role in the retrieval only when involved in a difficult or particularly demanding task (Laeng et al., 2014). This reasoning explains why some studies that used simple stimuli or easy tasks did not observe the memory facilitation (e.g. Richardson & Spivey, 2000). However, it could not apply in case of the study by Laeng and Teodorescu (2002) where the effect was nevertheless observed.

The usage of the simple, geometrical stimuli in the current experiment was based on the assumption that the presence of meaningful information (like faces, concrete objects, colors) could interfere with the effect of scanpath's reenactment. Interestingly, Johansson and Johansson (2014) recently provided evidence for the eye movements' role in memory for the intrinsic features of the objects and the spatial relations between them. They argued that the impact of eye movements on memory could differ depending on the nature of representation. Thus, the effect of memory facilitation through re-enactment of the scanpath might be more visible when spatial relations between the objects are memorized and not their intrinsic features.

Furthermore, some researchers argued that precluding eye movements through forcing central fixation could result in additional cognitive effort and therefore result in accuracy loss (Martarelli & Mast, 2013). This critique applies here only partially as in the current experiment the significant difference between the scanpath's impact on memory was visible also within free viewing condition. However, the forced fixation condition was also essential in the observed results. Following the argument of Laeng and colleagues (2014), it is questionable that memory could be disrupted through central fixation of the gaze. They pointed out that fixating a gaze on an object is a natural human behavior and it is in fact redundant eye movements to nothing, observed in imagery that could overload cognition if they were truly unnecessary. Additionally, a vast number of psychological studies used central fixation and did not report any disruption of cognitive processes caused by that (e.g. Micic, Ehrlichman, & Chen, 2010; Postle et al., 2006). Accordingly, our pupillary results point to a lack of additional cognitive effort in the forced fixation condition. Based on Kahneman's (1973) suggestion and numerous research supporting his hypothesis, working memory load can be indexed by pupillary changes. The current experiment provides no evidence for differences between the pupil diameters in forced fixation and free viewing conditions. Thus, there is no direct empirical support for maintaining central fixation as being more effortful that moving the eyes.

Interestingly, according to Foulsham and Underwood (2008), Scanpath Theory in its original form might be too rigid, as the subjects must repeat exactly the same sequences of fixations whenever they would see again an object or a scene. In an everyday life when one is encountered with many complex natural scenes perceived from different angles of view this exact re-enactment of the scanpath could be a challenge (Sanocki, 2003). However, this argument remains unsupported by empirical evidence. Moreover, it assumes that a memory trace can be perfect whereas none of the theories that try to explain processes underlying memory would claim that.

Lastly, we did not observe a correlation between the results from the grid's memory task and the Corsi Block-Tapping Task that was used as an additional measure of visuospatial memory in the current experiment. However, the Corsi Block-Tapping Task was designed to test the short-term memory whereas the grid's memory task tested long-term memory processes. Based on the research that investigates the differences between spatial representations in both short-term and long-term memory, those two memory storages operate on distinct coordinates of the frames of reference for the spatial layout and are underlined by different neural processes (Giudice, Klatzky, Bennett, & Loomis, 2013). We would argue that those differences caused the lack of the correlation between the tests. Additionally, no gender differences were observed across the tasks. However, this result stays in accordance with the other studies investigating the role of eye movements in the memory retrieval (e.g. Holm and Mäntylä, 2007; Johansson et al., 2012; Laeng & Teodorescu, 2002) as none of them reported differences between females' and males' performances.

Limitations and implications for the further research

In the current experiment an arbitrary decision regarding the scanpath length was made. Each time, the first 6 fixations, disregarding the first one, were recreated when preparing the order of the appearing squares for the recognition session. One could argue that in this way, an incomplete scanpath was enacted. In their study with the MultiMatch model, Foulsham and colleagues (2012) observed that the average number of fixations that belonged to one scanpath equaled 10.5. In a different study by Humphrey and Underwood (2008) the average number of fixations while looking at a picture for 5000 ms was 11. However, such measurements can be biased by the presentation time of the stimulus and above a certain number of fixations, additional fixations may not add any information that is relevant to the quality of the memory. Additionally, findings from Holm and Mäntylä (2007) showed that memory guidance is involved early and that refixations measure affected recognition judgments within first three fixations.

Furthermore, Foulsham et al. (2012) observed the highest similarity of the scanpath duration in within-participant, between-image, comparisons. Thus, the length of the fixations is of an idiosyncratic nature and each participant shows scanpath of similar duration regardless of the image. This observation is supported by other findings, in the earlier studies, reporting consistency of average saccade amplitude and fixation duration within a particular subject across different tasks (Andrews & Coppola, 1999; Groner & Menz, 1985; Rayner, Li, Williams, Cave, & Well, 2007). As in the current study the scanpath at recognition had an equal length for all participants, individual differences could interfere with the memory retrieval. However, it was assumed that six first main fixations would be enough for the participants to recognize the pattern as it consisted of four elements (i.e. black squares). If the original scanpath was much longer than that, repeating the first part of it should have sufficed to trigger memory retrieval. Additionally, close-up observations of the pilot study revealed that the average number of main fixations before repeating eye movements to already fixated locations slightly varied around an average six fixations. However, regardless of the original duration of every single fixation during encoding, each black square during recognition session appeared on the screen for the same amount of time, which disregards the naturalistic duration of a participant's individual fixations. Further research could investigate whether adjusting individually for each participant the number of fixations as well as the duration of every fixation would have an impact on the results through, for example, increasing the power of the significance.

Furthermore, no differentiation between explicit memory and familiarity-based recognition was made in the present study. However, based on the evidence provided by Mäntylä and Holm (2006) and Holm and Mäntylä (2007), those two types of recognition memory could evoke distinct patterns of eye movements. Researchers showed that familiarity-based recognition was not associated with the refixations to the original locations from encoding in contrast to the explicit memory. Thus, one could reason that specifying "yes" answer into "remember" and "know" could reveal a difference of, for example, effect size making the impact of the same scanpaths more reliable in case of the "remember" comparing to "know" trials.

Referring to the discussion about the complexity of the stimuli and the difficulty of the task (Johansson et al., 2012), one could argue that more complex stimuli like the pictures of natural surroundings might have a different impact on the relationship between eye movements and visual memory. As Johansson et al. (2012) underlined, encoding procedure might be significantly different for artificial objects like figures or grids and more complex pictures. However, more complex stimuli of this kind might consist of various meaningful details that could interfere with the scanpath's effect on memory. Consequently, it is a challenge for the future research to avoid this interference and be still able to test scan patterns with the help of different stimuli than just simple geometrical patterns.

Conclusions

In the current study we found a significant increase in accuracy rates and decrease in response times when the original eye scanpath from encoding was enacted during retrieval. Importantly, this effect was observed only when participants were allowed to move their eyes according to a pre-determined path but not when they maintained a central fixation. These findings are in accordance with recent theories pointing to a functional role of eye movements in visual memory as well as with the theories arguing for the reinstatement of the motor activity from encoding during retrieval. Moreover, they provide new evidence for the facilitatory role of the original sequences of eye fixations and that sequences are stored in long-term memory. Furthermore, a novel and successful procedure for testing the gaze scanpaths was introduced. Hence, the current experiment gives new insights in the understanding of the important role of eye movements in long-term visual memory.

References

Andrews, T. J., & Coppola, D. M. (1999). Idiosyncratic characteristics of saccadic eye movements when viewing different visual environments. *Vision Research*, *39(17)*, 2947-2953. doi:10.1016/S0042-6989(99)00019-X

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

Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Roe, R. P. N. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, *20*, 723–742. doi:10.1017/S0140525X97001611

Barsalou, L. W. (2010). Grounded Cognition: Past, Present and Future. *Topics in Cognitive Science*, *2*, 716-724. doi:10.1111/j.1756-8765.2010.01115.x

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–645. doi: 10.1146/annurev.psych.59.103006.093639

Barrowcliff, A. L., Gray, N. S., Freeman, T. C. A., & Macculloch, M. J. (2004). Eye-movements reduce the vividness, emotional valence and electrodermal arousal associated with negative autobiographical memories. *The Journal of Forensic Psychiatry & Psychology*, *15*, 325–345. doi:10.1080/14789940410001673042

Bjork, R. A., & Richardson-Klavehn, A. (1989). On the puzzling relationship between environmental context and human memory. In C. Izawa (Eds.)*, Current issues in cognitive processes: the Tulane Flowerree Symposium on Cognition*, (pp. 313-344). Hillsdale, NJ: Lawrence Erlbaum.

Blaxton, T. A. (1989). Investigating dissociations among memory measures: Support for a transfer-appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *15*, 657-668. doi:10.1037/0278-7393.15.4.657

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences of the United States of America*, *105(38)*, 14325–14329. doi:10.1073/pnas.0803390105

Brandt, S. A., & Stark, L. W. (1997). Spontaneous eye movements during visual imagery reflect the content of the visual scene. *Journal of Cognitive Neuroscience*, *9*, 27–38. doi:10.1162/jocn.1997.9.1.27

Chan, J. P. K., Kamino, D., Binns, M. A., & Ryan, J. D. (2011). Can changes in eye movement scanning alter the age-related deficit in recognition memory? *Frontiers in Psychology*, *2*, 1–11. doi:10.3389/fpsyg.2011.00092

Choi, Y. S., Mosley, A. D., & Stark, L. W. (1995). String editing analysis of human visual search. *Optometry & Vision Science*, *72(7)*, 439–451. doi:10.1097/00006324-199507000-00003

Danker, J. F., & Anderson, J. R. (2010). The ghosts of brain states past: Remembering reactivates the brain regions engaged during encoding. *Psychological Bulletin*, *136*, 87–102. doi:10.1037/a0017937

Decety, J., Perani, D., Jeannerod, M., Bettinardi, V., Tadary, B., Woods, R., Mazziotta, J. C., & Fazio, F., (1994). Mapping motor representations with positron emission tomography. *Nature*, *371*, 600–602. doi:10.1038/371600a0

Decety, J., Jeannerod, M., & Prablanc, C. (1989). The timing of mentally represented actions. *Behavioral Brain Research*, *34*, 35–42.

Deutsch, G., Bourbon, W. T., Papanicolau, A. C., & Eisenberg, H. M. (1988). Visuospatial tasks compared with activation of regional cerebral blood flow. *Neuropsychologia*, *26*, 445–452. doi:10.1016/0028-3932(88)90097-8

Farah, M. J. (1995). Current issues in the neuropsychology of image generation. *Neuropsychologia*, *33(11)*, 1455-1471. doi:10.1016/0028-3932(95)00075-E

Farah, M. J., Hammond, K. M., Levine, D. N., & Calvanio, R. (1988). Visual and spatial mental imagery: Dissociable systems of representation. *Cognitive Psychology*, *20*, 439–462. doi:10.1016/0010-0285(88)90012-6

Ferreira, F., Apel, J., & Henderson, J. M. (2008). Taking a new look at looking at nothing. *Trends in Cognitive Sciences*, *11*, 405–410. doi:10.1016/j.tics.2008.07.007

Foulsham, T., Dewhurst, R., Nyström, M., Jarodzka, H., Johansson, R., Underwood, G., & Holmqvist, K. (2012). Comparing scanpaths during scene encoding and recognition: A multi-dimensional approach. *Journal of Eye Movement Research*, *5(4)*, 1-14.

Foulsham, T., & Kingstone, A. (2013). Where have the eyes been? Observers can recognize their own fixations. *Perception*, *42*, 1085-1089. doi:10.1068/p7562

Foulsham, T., & Kingstone, A. (2012). Fixation-Dependent Memory for Natural Scenes: An Experimental Test of Scanpath Theory. *Journal of Experimental Psychology: General*, 1-16. doi:10.1037/a0028227

Foulsham, T., & Underwood, G. (2008). What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. *Journal of Vision*, *8(2)*, 1–17. doi:10.1167/8.2.6

Galton, F. (1883). *Inquiries into Human Faculty and its Development*. Macmillan.

Ganis, G., Thompson, W. L., & Kosslyn, S. (2004). Brain areas underlying visual mental imagery and visual perception: an fMRI study. *Cognitive Brain Research*, *20*, 226-241. doi:10.1016/j.cogbrainres.2004.02.012

Gazzaniga, M. S., Ivry, R., & Mangun, G. R. (2008). *Cognitive Neuroscience: The Biology of the Mind*, (3rd Edition). New York: W. W. Norton.

Giudice, N. A., Klatzky, R. L., Bennett, Ch. R., & Loomis, J. M. (2013). Combining Locations from Working Memory and Long-Term Memory into a Common Spatial Image.

Spatial Cognition & Computation: An Interdisciplinary Journal, *13(2)*, 103-128. doi:10.1080/13875868.2012.678522

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and under water. *British Journal of Psychology*, *66*, 325-331. doi:10.1111/j.2044-8295.1975.tb01468.x

Goodale, M. A., Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15(1*), 20–25. doi:10.1016/0166-2236(92)90344-8

Groner, R., & Menz, C. (1985). The effects of stimulus characteristics, task requirements and individual diferences on scanning patterns. In R. Groner, G. W. McConkie, & Ch. Menz (Eds.), *Eye movements and human information processing*, (pp. 239-250). Amsterdam: Elsevier.

Gunter, R. W., & Bodner, G. E. (2008). How eye movements affect unpleasant memories: Support for a working-memory account. *Behaviour Research and Therapy*, *46*, 913–931. doi:10.1016/j.brat.2008.04.006

Harding, G., & Bloj, M. (2010). Real and predicted influence of image manipulations on eye movements during scene recognition. *Journal of Vision*, *10(2)*, Article 8. doi:10.1167/10.2.8

Hebb, D. O. (1949). *The organization of behavior*. New York: Wiley.

Hebb, D. O. (1968). Concerning imagery. *Psychological Review*, *75*, 466–477. doi:10.1037/h0026771

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, *7(11)*, 498-504. doi:10.1016/j.tics.2003.09.006

Henderson, J. M., & Ferreira, F. (2004). *The interface of language, vision, and action: Eye movements and the visual world.* New York: Psychology Press.

Holm, L., & Mäntylä, T. (2007). Memory for scenes: Refixations reflect retrieval. *Memory & Cognition*, *35(7)*, 1664 –1674. doi:10.3758/BF03193500

Hollingworth, A. (2006). Scene and position specificity in visual memory for objects. *Journal of Experimental Psychology: Learning*, *Memory, and Cognition*, *32*, 58–69. doi:10.1037/0278-7393.32.1.58

Hoover, M. A., & Richardson, D. C. (2008). When facts go down the rabbit hole: Contrasting features and objecthood as indexes to memory. *Cognition*, *108*, 533–542. doi:10.1016/j.cognition.2008.02.011

Humphrey, K., & Underwood, G. (2010). The potency of people in pictures: Evidence from sequences of eye fixations. *Journal of Vision*, *10*, 1–10. doi: 10.1167/10.10.19

Humphrey, K., & Underwood, G. (2009). Domain knowledge moderates the influence of visual saliency in scene recognition. *British Journal of Psychology*, *100(2)*, 377–398. doi:10.1348/000712608X344780

Humphrey, K., & Underwood, G. (2008). Fixation sequences in imagery and in recognition during the processing of pictures of real-world scenes. *Journal of Eye Movement Research*, *2(2)*, 1-15.

Irwin, D. E., & Zelinsky, D. J. (2002). Eye movements and scene perception: Memory for things observed. *Perception and Psychophysics*, *64*, 882-895. doi:10.3758/BF03196793

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*, 1489–1506. doi:10.1016/S0042-6989(99)00163-7

Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20, 1254–1259. doi: 10.1109/34.730558

Jacobson, E. (1932). Electrophysiology of mental activities. *American Journal of Psychology*, *44*, 677–694. doi:10.2307/1414531

Jacoby, L. L., & Craik, F. I. M. (1979). Effects of elaboration at encoding and retrieval: Trace distinctiveness and recovery of initial context. In L. S. Cermak, & F. I. M. Craik (Eds.), *Levels of processing in human memory*, (pp. 1-23). Hillsdale, NJ: Erlbaum.

James, W. (1890). *The principles of psychology*. New York: Holt.

Janssen, W. H., & Nodine, C. F. (1974). Eye movements and visual imagery in free recall. *Acta Psychologica*, *38*, 267–276. doi:10.1016/0001-6918(74)90011-0

Johansson, R., & Johansson, M. (2014). Look here, eye movements play a functional role in memory retrieval. *Psychological Science*, *25(1)*, 236–242. doi:10.1177/0956797613498260

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*, 1289–1314. doi:10.1037/a0026585

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: A Multidisciplinary Journal*, *30*, 1053–1079. doi:10.1207/s15516709cog0000_86

Just, M. A., & Carpenter, P. A. (1993). The intensity dimension of thought: Pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology*, *47*, 310–339.

Kahneman, D., & Beatty, J. (1967). Pupillary responses in a pitch-discrimination task. *Perception and Psychophysics, 2,* 101-105. doi:10.3758/BF03210302

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.

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

Kessels, R. P. C., Van Zandvoort, M. J. E., Postma, A., Kappelle, L. J. & De Haan, E. H. F. (2000). The Corsi Block-Tapping Task: Standardization and Normative Data, *Applied Neuropsychology*, *7(4)*, 252-258, doi:10.1207/S15324826AN0704_8

Knoeferle, P., & Crocker, M. W. (2007). The influence of recent scene events on spoken comprehension: Evidence from eye movements. *Journal of Memory and Language*, *57*, 519-543. doi:10.1016/j.jml.2007.01.003

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology*, *4*, 219-227. doi:10.1007/978-94-009-3833-5_5

Kolers, P. A. (1973). Remembering operations. *Memory & Cognition*, *1(3)*, 347–355. doi:10.3758/BF03198119

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience*, *2*, 635–642. doi:10.1038/35090055

Kosslyn, S. M., Thompson, W. L., Kim, I. J., & Alpert, N. M. (1995). Topographical representations of mental images in primary visual cortex. *Nature*, *378(6556)*, 496-498. doi:10.1038/378496a0

Kosslyn, S. M. (1994). *Image and brain*. Cambridge, MA: MIT Press.

Laeng, B., Bloem, I. M., D'Ascenzo, S., & Tomassi, L. (2014). Scrutinizing visual images: The role of gaze in mental imagery and memory. *Cognition*, *131*, 263-283. doi:10.1016/j.cognition.2014.01.003

Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry: A window to the preconscious? *Perspectives on Psychological Science, 7,* 18-27. doi:10.1177/1745691611427305

Laeng, B., & Teodorescu, D. S. (2002). Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cognitive Science: A Multidisciplinary Journal*, *26*, 207–231. doi:10.1207/s15516709cog2602_3

Laeng, B., Waterloo, K., Johnsen, S. H., Bakke, S. J., Låg, T., Simonsen, S. S., & Høgsaet, J. (2007). The eyes remember it: oculography and pupillometry during recollection in three amnesic patients. *Journal of Cognitive Neuroscience*, *19(11)*, 1888-1904. doi:10.1162/jocn.2007.19.11.1888

Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological Assessment*, (4th Edition). New York: Oxford University Press.

Loftus, G. R. (1972). Eye fixations and recognition memory for pictures. *Cognitive Psychology*, *3*, 525-551. doi:10.1016/0010-0285(72)90021-7

Lotze, M., Montoya, P., Erb, M., Hüsmann, E., Flor, H., Klose, U., Birbaumer, N., & Grodd, W. (1999). Activation of cortical and cerebellar motor areas during executed and imagined hand movements: an fMRI study. *Journal of Cognitive Neuroscience*, *11*, 491–501. doi:10.1162/089892999563553

Mäntylä, T., & Holm, L. (2006). Gaze control and recollective experience in face recognition. *Visual Cognition*, *14*, 365–386. doi:10.1080/13506280500347992

Martarelli, C. S., & Mast, F. W. (2013). Eye movements during long-term pictorial recall. *Psychological Research*, *77*, 303–309. doi:10.1007/s00426-012-0439-7

Meister, I. G., Krings, T., Foltys, H., Müller, M., Töpper, R., & Thron, A. (2004). Playing piano in the mind — an fMRI study on music imagery and performance in pianists. *Cognitive Brain Research*, *19*, 219–228. doi:10.1016/j.cogbrainres.2003.12.005

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

Mizukawa, K. (2009). A trial of objective visual field measurement by pupillary reaction. *Kawasaki Medical Journal, 35*, 51-60.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning & Verbal Behavior*, *16*, 519-533. doi:10.1016/S0022-5371(77)80016-9

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

Noton, D., & Stark, L. W. (1971a). Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vision Research*, *11*, 929–942. doi:10.1016/0042-6989(71)90213-6

Noton, D., & Stark, L. W. (1971b). Scanpaths in eye movements during perception. *Science*, *171*, 308–311. doi:10.1126/science.171.3968.308

Nyberg, L., Habib, R., McIntosh, A. R., & Tulving, E. (2000). Reactivation of encoding-related brain activity during memory retrieval. *Proceedings of the National Academy of Sciences, USA*, *97*, 11120–11124.

Olsen, R. K., Chiew, M., Buchsbaum, B. R., & Ryan, J. D. (2014). The relationship between delay period eye movements and visuospatial memory. *Journal of Vision*, *14(1):8*, 1–11. doi:10.1167/14.1.8

Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, *42*, 107-123. doi:10.1016/S0042-6989(01)00250-4

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

Privitera, C. M., Stark, L. W., & Zangemeister, W. H. (2007). Bonnard's representation of the perception of substance. *Journal of Eye Movement Research*, *1(1)*, 1-6.

Quinn, J. (1994). Toward a clarification of spatial processing. *Quarterly Journal of Experimental Psychology*, *47A*, 465–480. doi:10.1080/14640749408401120

Rayner, K., Li, X. S., Williams, C. C., Cave, K. R., & Well, A. D. (2007). Eye movements during information processing tasks: Individual differences and cultural effects. *Vision Research*, *47(21)*, 2714-2726. doi:10.1016/j.visres.2007.05.007

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

Richardson, D. C., & Kirkham, N. Z. (2004). Multimodal events and moving locations: Eye movements of adults and 6-month-olds reveal dynamic spatial indexing. *Journal of Experimental Psychology: General*, *133*, 46–62. doi:10.1037/0096-3445.133.1.46

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

Sanocki, T. (2003). Representation and perception of scenic layout. *Cognitive Psychology*, *47*, 63–86. doi:10.1016/S0010-0285(03)00002-1

Scholz, A., Melhorn, K., Bocklisch, F., & Krems, J. F. (2011). Looking at nothing diminishes with practice. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd annual meeting of the Cognitive Science Society*, (pp. 1070–1075). Austin, TX: Cognitive Science Society.

Schwoebel, J., Boronat, C. B., & Coslett, H. B. (2002). The man who executed "imagined" movements: evidence for dissociable components of body schema. *Brain Cognition*, *50*, 1–16. doi:10.1016/S0278-2626(02)00005-2

Sinclair, G. P., Healy, A. F., & Bourne Jr., L. E. (1997). The acquisition and long term retention of temporal, spatial and item information. *Journal of Memory and Language*, *36*, 530-549. doi:10.1006/jmla.1996.2505

Slotnick, S. D., Thompson, W. L., & Kosslyn, S. M. (2012). Visual memory and visual mental imagery recruit common control and sensory regions of the brain. *Cognitive Neuroscience*, *3(1)*, 14-20. doi:10.1080/17588928.2011.578210

Spivey, M., & Geng, J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research*, *65*, 235–241. doi:10.1007/s004260100059

Tanenhaus, M. K., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. (1995). Integration of visual and linguistic information during spoken language comprehension. *Science*, *268*, 1632-1634. doi:10.1126/science.7777863

Tatler, B. W., Baddeley, R. J., & Gilchrist, I. D. (2005). Visual correlates of fixation selection: effects of scale and time. *Vision Research*, *45(5)*, 643-659. doi:10.1016/j.visres.2004.09.017

Tatler, B. W. (2007). The central fixation bias in scene viewing: selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of Vision*, *7(14)*, 1-17. doi:10.1167/7.14.4

Thomas, N. J. T. (1999). Are theories of imagery theories of imagination? An active perception approach to conscious mental content. *Cognitive Science*, *23*, 207–245. doi:10.1016/S0364-0213(99)00004-X

Totten, E. (1935). Eye movements during visual imagery. *Comparative Psychology Monographs*, *11*, 1-46.

Tremblay, S., Saint-Aubin, J., & Jalbert, A. (2006). Rehearsal in serial memory for visual-spatial information: Evidence from eye movements. *Psychonomic Bulletin & Review*, *13(3)*, 452-457. doi:10.3758/BF03193869

Tulving, E. (1985). Memory and consciousness. *Canadian Journal of Psychology*, *26*, 1-12. doi:10.1037/h0080017

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80(5)*, 352–373. doi:10.1037/h0020071

Underwood, G., & Foulsham, T., (2006). Visual saliency and semantic incongruency influence eye movements when inspecting pictures. *Quarterly Journal of Experimental Psychology*, *59*, 1931-1949. doi:10.1080/17470210500416342

Underwood, G., Foulsham, T., & Humphrey, K. (2009). Saliency and scan patterns in the inspection of real-world scenes: Eye movements during encoding and recognition. *Visual Cognition*, *17(6-7)*, 812–834. doi:10.1080/13506280902771278

Underwood, G., Foulsham, T., van Loon, E., Humphreys, L., & Bloyce, J. (2006). Eye movements during scene inspection: A test of the saliency map hypothesis. *European Journal of Cognitive Psychology*, *18*, 321-342. doi:10.1080/09541440500236661

Valuch, C., Becker, S. I., & Ansorge, U. (2013). Priming of fixations during recognition of natural scenes. *Journal of Vision*, *13(3)*, 1-22. doi: 10.1167/13.3.3

Wexler, M., Kosslyn, S. M., & Berthoz, A. (1998). Motor processes in mental rotation. *Cognition*, *68*, 77-94. doi:10.1016/S0010-0277(98)00032-8

Wheeler, M. E., Petersen, S. E., & Buckner, R. L. (2000). Memory's echo: Vivid remembering reactivates sensoryspecific cortex. *Proceedings of the National Academy of Sciences*, *USA*, *97*, 11125-11129. doi:10.1073/pnas.97.20.11125

Whiteside, J. A. (1978). Visual display recognition and duplication of inspection sequences. *Perception & Psychophysics*, *23(4)*, 350–355. doi:10.3758/BF03199721

Winograd, E., & Church, V. (1988). Role of spatial location in learning face-name associations. *Memory and Cognition*, *16(1)*, 1-7. doi:10.3758/BF03197739

Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum Press.

Appendices

Appendix 1: Complete set of stimuli for the pilot study with the images of paintings.

Appendix 2: Consent Form.

University of Oslo

Psychology Department

Consent Form

Purpose:
The purpose of this study is to examine how people memorize geometric patterns resembling checkerboards and retrieve this visual information from their long-term memory.

The study is a project for master thesis in Cognitive Neuroscience under the supervision of Professor Bruno Laeng.

Procedure:
If you agree to be in this study, you will be asked to do the following:

1. You will need to take part in two experimental sessions with 48 hours delay from each other.
2. In the first session you will be looking at several images resembling checkerboards and try to memorize them as well as you can.
3. In the second session you will be asked to decide which images you have seen in the first session and which ones you have not seen before. You will be answering by pressing the keys on the keyboard.
4. During both first and second session, eye-tracker will be measuring your eye movements and the size of your pupils.
5. You will be additionally asked to take part in short and simple task that will measure your spatial memory. It will be so-called Corsi Block Tapping Task.

The total time required to complete the study should be approximately 30 - 40 minutes (around 15 minutes for each session).

Voluntary Nature of the Study/Confidentiality:
Your participation in this study is entirely voluntary and you may refuse to complete the study at any point during the experiment, or refuse to answer any questions with which you are uncomfortable. You may also stop at any time and ask the researcher any questions you may have. Your name will never be connected to your results or to your responses on the questionnaires; instead, a number will be used for identification purposes. Information that would make it possible to identify you or any other participant will never be included in any sort of report. The data will be accessible only to those working on the project.

Questions:
At this time you may ask any questions you may have regarding this study.

Statement of Consent:
I have read the above information. I have asked any questions I had regarding the experimental procedure and they have been answered to my satisfaction. I consent to participate in this study, and I agree that the results can be published later on.

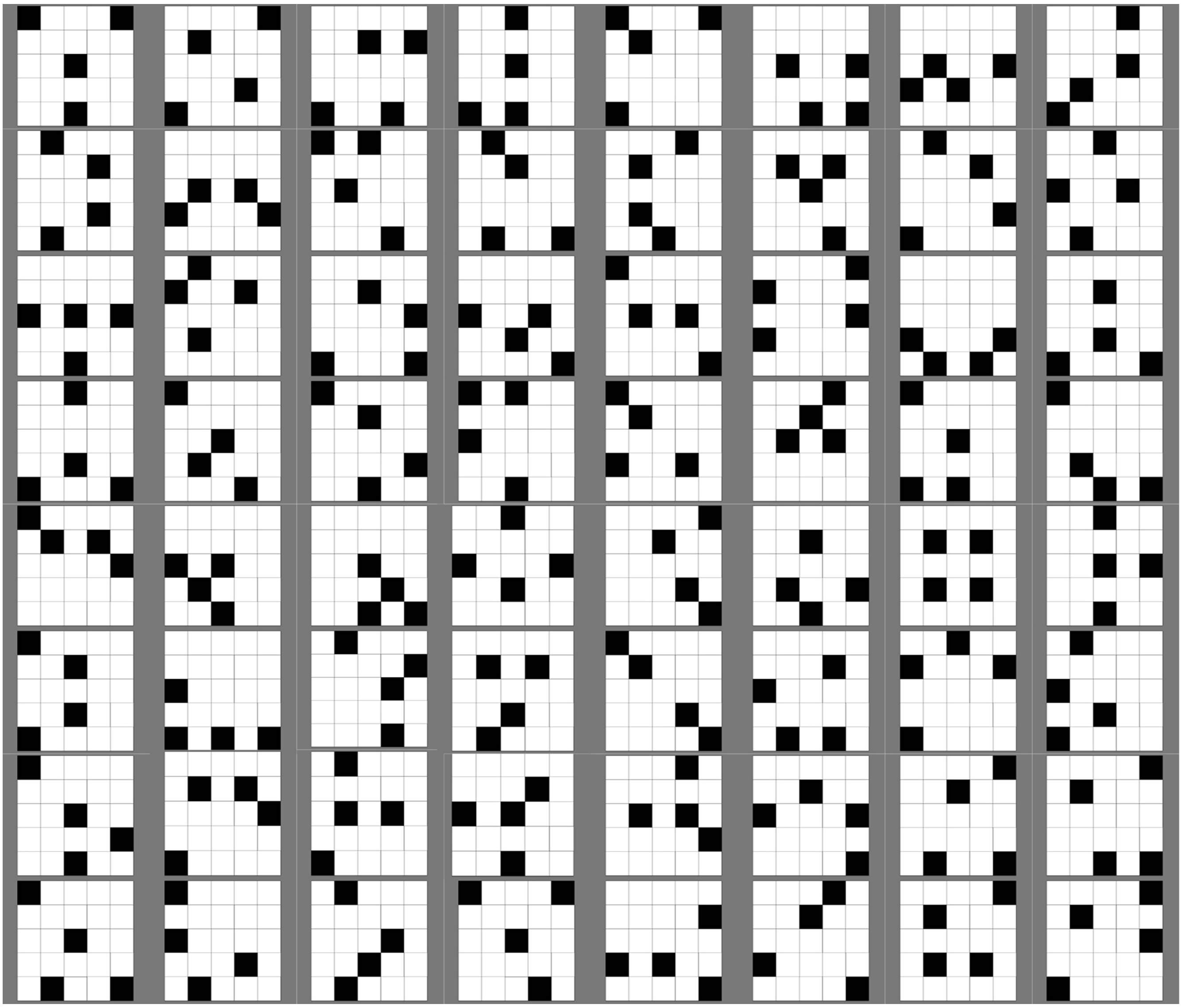
Name of Participant____Date: __________
(in capital letters)

Signature of Participant ____

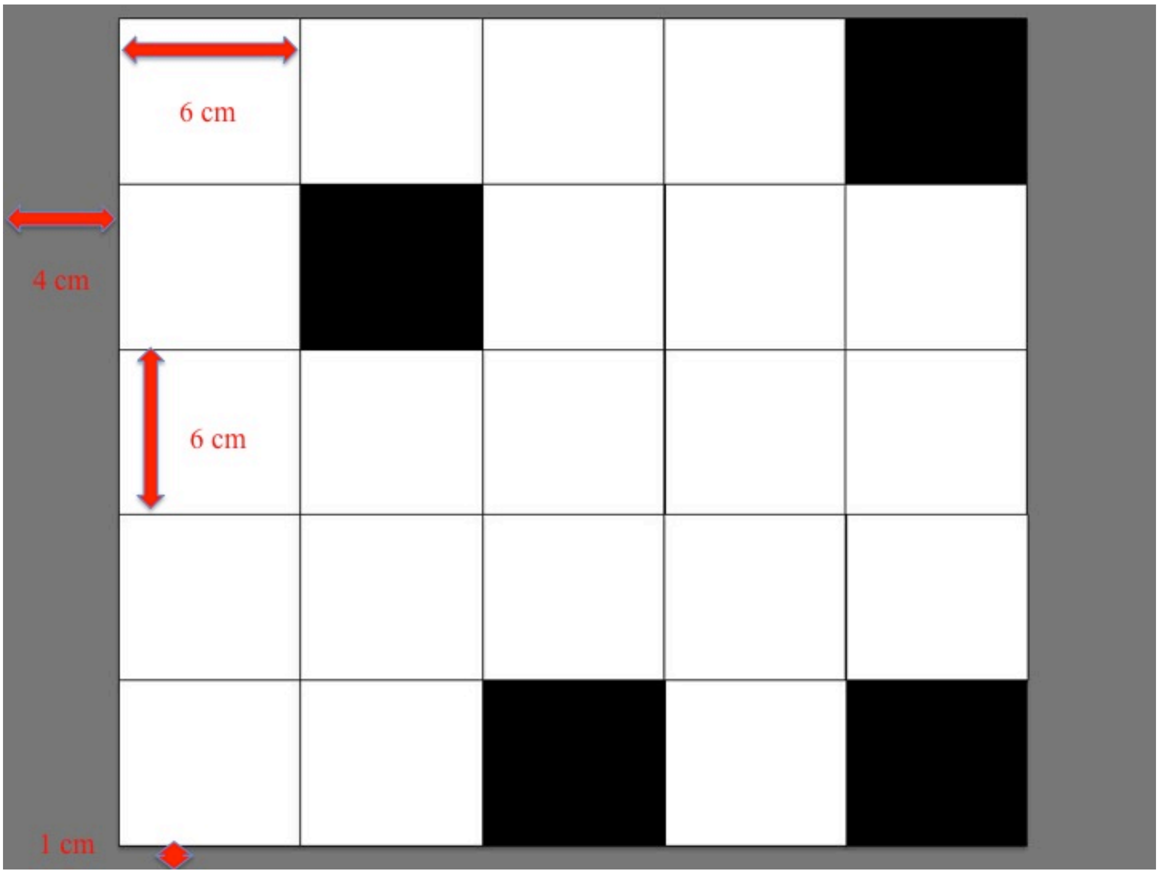
Age: _________

Thanks for your participation!

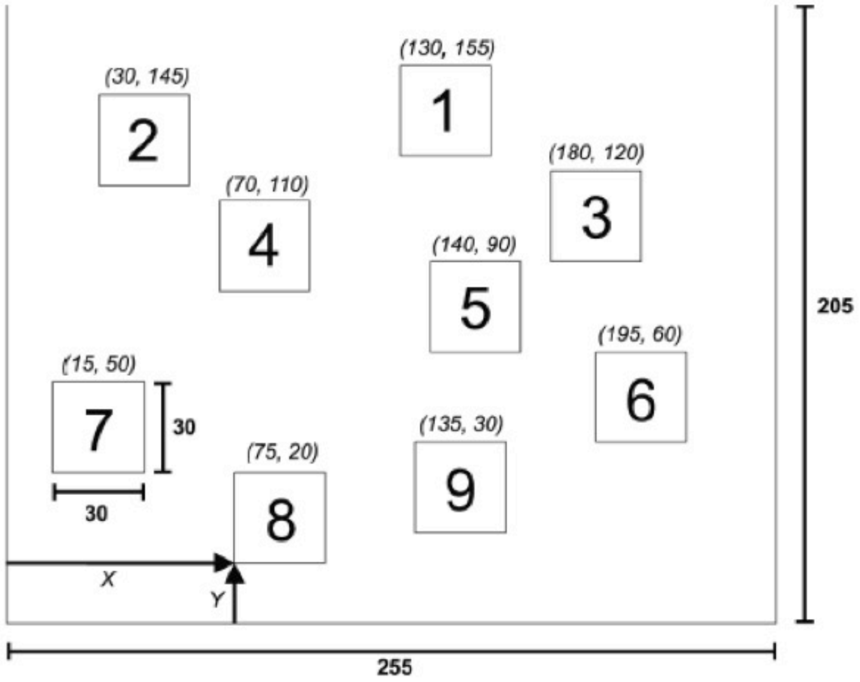
Appendix 3: Complete set of stimuli for the current study with the images resembling checkerboards.



Appendix 4: The size of the images of grids on the computer screen (resolution of 1680 x 1050).



Appendix 5A: Detailed illustration of the test materials for the Corsi Block-Tapping Task (Kessels et al., 2000). Coordinates in millimeters.



Appendix 5B: The block sequences in the Corsi Block-Tapping Task (Kessels et al., 2000):

8–5
6–4
4–7–2
8–1–5
3–4–1–7
6–1–5–8
5–2–1–8–6
4–2–7–3–1
3–9–2–4–8–7
3–7–8–2–9–4
5–9–1–7–4–2–8
5–7–9–2–8–4–6
5–8–1–9–2–6–4–7
5–9–3–6–7–2–4–3
5–3–8–7–1–2–4–6–9
4–2–6–8–1–7–9–3–5