Journal of Experimental Psychology:
Human Perception and Performance
2012, Vol. 38, No. 5, 1289–1314

0096-1523/11/$12.00 DOI: 10.1037/a0026585

Eye Movements During Scene Recollection Have a Functional Role, but They Are Not Reinstatements of Those Produced During Encoding

Roger Johansson, Jana Holsanova, Richard Dewhurst, and Kenneth Holmqvist
Lund University, Lund, Sweden

Current debate in mental imagery research revolves around the perceptual and cognitive role of eye movements to "nothing" (Ferreira, Apel, & Henderson, 2008; Richardson, Altmann, Spivey, & Hoover, 2009). While it is established that eye movements are comparable when inspecting a scene (or hearing a scene description) as when visualizing it from memory (Johansson, Holsanova, & Holmqvist, 2006), the exact purpose of these eye movements remains elusive. Are eye movements during recall purely epiphenomenal or do they have a functional purpose? Here we address this question in four experiments where eye movements were prohibited either during the encoding or recall phases. Experiments 1 and 2 showed that maintaining central fixation during visual or auditory encoding, respectively, had no effect on how eye movements were executed during recall (but it did hinder memory retrieval). Thus, oculomotor events during recall are not reinstatements of those produced during encoding. When fixation was restricted during recall, Experiments 3 and 4 revealed that scene recollection was altered and impaired, irrespective of the modality of encoding. The functional role of eye movements during mental visualization is therefore apparent in this perturbation of visuospatial capabilities.

Keywords: eye movements, visual imagery, encoding, recall, scene description, memory retrieval, spatial cognition, visual attention

We are constantly engaged in deciphering, and acting upon, the visual world around us via eye movements. However, it is not just immediate visual input that facilitates this process, but also attentional resources recruited during visual imagery. Visual imagery can be defined as "the mental invention or recreation of an experience that resembles the experience of actually perceiving a visual stimulus, but without direct sensory stimulation" (Finke, 1989, p. 2). This is something we all do when thinking back to something we have seen, or recollecting something we have heard about. We can spontaneously refer to an internal mental image which retains an approximation of the spatial relations between objects in the external environment. Based on neural findings that visual imagery draws on most of the same neural architecture as visual perception (e.g., Ganis, Thompson, & Kosslyn, 2004; Slotnick, Thompson, & Kosslyn, 2005), a prominent interpretation is to see visual imagery as a simulation of perception (e.g., Hesslow, 2011). A large body of research has also reported that the oculomotor system is recruited during these perceptual simulations (cf., Ferreira, Apel & Henderson, 2008; Richardson, Altmann, Spivey, & Hoover, 2009; Theeuwes, Belopolsky, & Oliviers, 2009). We have, for instance, previously shown that eye movements spontaneously occur with visual imagery and that they closely correspond with content and spatial relations from an original picture or scene (Johansson, Holsanova, & Holmqvist, 2006). Consequently, studies of eye movements are a powerful tool for investigating aspects of visual imagery and visuospatial memory retrieval.

Linking Eye Movements to Visual Imagery

Eye movement research during visual imagery has a rather long and winding history. Early studies reported a large amount of eye movement activity during visual imagery (Jacobson, 1932; Perky, 1910; Totten, 1935), and it was found that this effect varied for different stimuli (Clark, 1916), between individuals (Stoy, 1930), and for reported vividness (Goldthwait, 1933). But inconsistent and conflicting results in later studies, which focused on either eye movement rate (EMR) or electroculograms (EOGs), questioned a tight link between eye movements and visual imagery (e.g., Antrobus, Antrobus & Singer, 1964; Bergstrom & Hiscock, 1988; Brown, 1968; Deckert, 1964; Ehrlichman & Barrett, 1983; Hale & Simpson, 1970; Hiscock & Bergstrom, 1981; Janssen & Nodine, 1974; Marks, 1973; Weiner & Ehrlichman, 1976). However, the extent to which these experiments engaged participants in visual imagery is questionable. Furthermore, the methods of EMR and

This article was published Online First December 26, 2011.

Roger Johansson and Jana Holsanova, Cognitive Science Department, Lund University, Lund, Sweden; Richard Dewhurst and Kenneth Holmqvist, Humanities Laboratory, Lund University, Lund, Sweden.

Special thanks to Marcus Nyström for programming the MatLab scripts and to Petra Lundberg for coding eye movement correspondence, and we thank the participants for their participation in the experiments. Preliminary results from parts of this study have been presented at the Cognitive Science Conference 2010 in Portland (Johansson, Holsanova, & Holmqvist, 2010). The work was funded by the Swedish Research Council (grant no. 2007-2637), by a Fellowship at Hanse Wissenschaftskolleg (HWK) in Delmenhorst, and has been supported by the Linnaeus center for Thinking in Time: Cognition, Communication, and Learning (CCL) at Lund University, which is funded by the Swedish Research Council (grant no. 349-2007-8695).

Correspondence concerning this article should be addressed to Roger Johansson, Cognitive Science Department, Lund University, Helgonabacken 12, 223 62 Lund, Sweden. E-mail: Roger.Johansson@lucs.lu.se

EOGs are not capable of measuring the location and direction of eye movements in detail.

Since the late ’90s several studies have used more sophisticated eye-tracking techniques and have consistently reported that spontaneous eye movements occur with visual imagery and that those eye movements closely reflect the content and spatial relations from an original picture or scene (Altmann, 2004; Altman & Kamide, 2009; Brandt & Stark, 1997; Gbadamosi & Zangemeister, 2001; Holsanova, Hedberg, & Nilsson, 1999; Humphrey & Underwood, 2008; Johansson et al., 2006; Knoeferle & Crocker, 2007; Laeng & Teodorescu, 2002; Spivey & Geng, 2001; Zangemeister & Liman, 2007). Consequently, there is no doubt that directions and positions are frequently accompanied by corresponding eye movements during visual imagery. However, explanations to why these eye movements occur vary, and current studies have produced conflicting results and interpretations (cf., Ferreira et al., 2008; Laeng & Teodorescu, 2002; Richardson & Spivey, 2000; Richardson et al., 2009).

Unresolved Issues

To fully understand the role of eye movements during visual imagery one needs to grasp two puzzling issues, central in the current literature. This study aims to resolve these issues.

First, a large body of research suggests that when we are engaged in visual imagery and tasks of memory retrieval there is a cognitive system that, at least partly, reactivates processes that were involved in a preceding encoding phase (for an overview, see Kent & Lamberts, 2008). Evidence is, however, inconclusive with regard to whether the oculomotor system is a part of this reactivation (cf., Laeng & Teodorescu, 2002; Richardson & Spivey, 2000, Experiment 5). The second, related point, concerns the vibrant debate revolving around whether eye movements to blank spaces can *facilitate memory retrieval* (cf., Ferreira et al., 2008; Richardson et al., 2009). There are mixed results in the literature (cf., Laeng & Teodorescu, 2002; Richardson et al., 2009).

To make these two points clear, on the one hand, there are results suggesting that eye movements during recall are *functionally* connected with those produced during encoding (Brandt & Stark, 1997; Laeng & Teodorescu, 2002), and that they facilitate memory retrieval (Janssen & Nodine, 1974; Laeng & Teodorescu, 2002). Conversely, there are results indicating that eye movements during recall are instead *epiphenomenal* with respect to those produced during encoding (Richardson & Spivey, 2000, Experiment 5), and therefore do not assist memory retrieval processes in any useful way (Hoover & Richardson, 2008; Richardson & Kirkham, 2004; Richardson & Spivey, 2000; Spivey & Geng, 2001).

Limitations of Previous Research

To design a study which can tackle these conflicting results, it is necessary to address four crucial factors.

First, because rather simple artificial stimuli have been used in the encoding phase of previous investigations (e.g., Laeng & Teodorescu, 2002; Spivey & Geng, 2001), there is reason to question whether results can be generalized to recall of more complex and naturalistic scenes (cf., Johansson et al., 2006). For instance, Brandt and Stark (1997) describe pilot studies where they used simple geometrical figures (e.g., circles, rectangles) and failed to find consistency in eye movements during recall. We have observed similar tendencies for scenes of low complexity in previous pilot studies in our Lab.

Second, the recall task in most studies has been for participants to answer questions either related or unrelated to visual features of the encoded stimuli (e.g., Richardson & Spivey, 2000; Spivey & Geng, 2001). However, we have also observed in the Lab that eye movements during visual imagery are less likely to appear for recall tasks that are relatively easy (e.g., questions about color, shape and location for single objects in a scene). Moreover, Scholz, Melhorn, Bocklisch, and Krems (2011) have demonstrated that the degree of eye movements to blank spaces decreases when encoding and recall are repeated in a second and a third set of trials, and Micic, Ehrlichman, and Chen (2010) have shown that eye movement rate (EMR) is highly task dependent for participants who are looking at a blank screen while engaged in a variety of cognitive tasks (e.g., delayed word repetition and N-back). Taken together, these findings highlight the importance of the experimental procedure.

Third, in several studies Richardson and colleagues (Hoover & Richardson, 2008; Richardson & Kirkham, 2004; Richardson & Spivey, 2000) used multimodal stimuli during encoding, where spoken statements have been associated with visual features on the screen. It is possible therefore that their results differ from those of Laeng and Teodorescu (2002) because of recall tasks which depend more on spoken information than on visual. To disambiguate such conflicting results, it is necessary to investigate both visual and verbal modalities separately.

Fourth, eye movements to blank spaces have mostly been researched in relation to a previous *visual* encoding phase (e.g., Brandt & Stark, 1997; Ferreira et al., 2008; Richardson & Spivey, 2000). However, in Johansson et al. (2006) we also showed that participants who *listened* to a scene description while looking at a blank screen made spontaneous eye movements which closely corresponded to spatial positions and directions from the aural scene description. Here, there was no visual input in the encoding phase. Instead a mental model (Bower & Morrow, 1990) of the scene was generated and continuously updated based on *linguistic content* and knowledge about objects and spatial relationships drawn from our semantic structure of the environment. Similar results have been reported from other research groups (Demarais & Cohen, 1998; Bourlon, Oliviero, Wattiez, Pouget, & Bartolomeo, 2010; Spivey & Geng, 2001; Spivey, Tyler, Richardson, & Young, 2000). Consequently, eye movements during visual imagery do not have to rely on a preceding encoding phase, and this suggests that a connection also exists between eye movements and visual imagery that is generated whole-cloth from long term-memory. This has important implications for the encoding-recall relationship and for memory retrieval in general, but has been left relatively unnoticed in the current debate of those questions (Ferreira et al., 2008; Richardson et al., 2009).

Aim of the Current Study

The aim of the current study is to address the above-mentioned conflicting results and to overcome the limitations of previous research. We present four experiments where an eye movement restriction is introduced in either the encoding phase or the recall phase. Experiment 1 and 2 investigated whether oculomotor events during recall are reinstatements of those produced during encoding, while Experiment 3 and 4 investigated whether eye movements during recall have an effect on memory retrieval. Furthermore, because both complex pictures (Experiment 1 and 3) and a spoken scene description (Experiment 2 and 4) were used, we could clarify whether eye movement restriction has a differential effect across input modality; visual and auditory, respectively. Figure 1 shows schematics of the four experiments in the current study with regard to eye movement restriction and stimuli.

Moreover, the current investigation was designed as a follow-up to Johansson et al. (2006), to shed further light on the cognitive mechanisms underlying visual imagery and eye movements. Results and data from that study will therefore be considered in several comparisons.

The Previous Study: Free Viewing During Encoding and Recall

In our previous study (Johansson et al., 2006) participants either inspected a complex picture or looked at a blank screen while listening to a verbal scene description. Afterward, while looking at a blank screen, they recalled the picture by orally describing it and recalled the scene description by retelling it (see Figure 2). Results from this study revealed that participants to a very high degree executed eye movements to appropriate spatial locations while describing the picture from memory, while listening to the spoken scene description (that was never seen in the first place) and while retelling it from memory. The data also indicated that the effect was equally strong during recall, irrespective of whether the original elicitation was spoken or visual. Furthermore, the effect was also equally strong during encoding and recall of the verbal scene description. Figure 2 shows the encoding phase and the recall phase for two typical participants in this study.

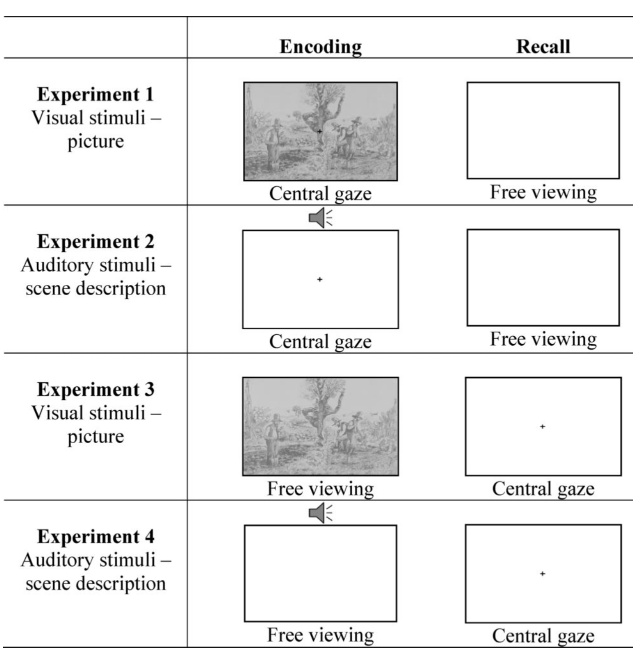

Figure 1. Schematics of all four experiments in the current study with regard to eye movement restriction and stimuli. This approach has the advantage of separating out encoding and recall processes, fixation restriction being manipulated orthogonally in each case.

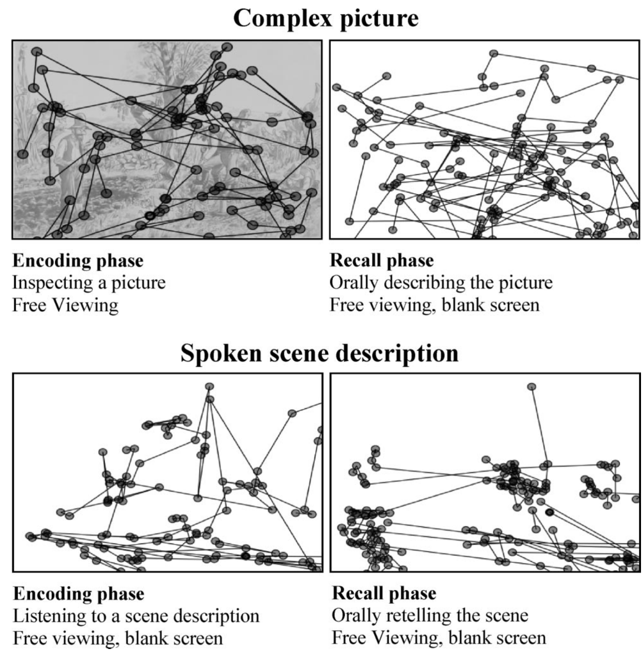

Figure 2. One participant's full scanpath during encoding and recall of a picture (above), and another participant's full scanpath during encoding and recall of a verbal scene description while looking at a blank screen (below).

Throughout the current study, results from the previous study (Johansson et al., 2006) were used for comparisons, and those will consistently be referred to as the previous study. Furthermore, in Experiments 3 and 4, verbal data from the previous study was reanalyzed, which provided crucial data for the analyses of the new experiments.

Experiment 1: Central Fixation During Encoding of a Complex Picture

The first experiment was designed to test whether oculomotor events during pictorial recall have a functional connection with those from encoding of the original picture. The method for doing this was to instruct participants to maintain central fixation when a picture was encoded. Then, in a subsequent recall phase, they were allowed to move their eyes freely while looking at a blank screen. This eye movement restriction will throughout the report be referred to as *central encoding*.

We hypothesized that if oculomotor events during recall are reinstatements of those produced during encoding, then participants' gaze pattern in the central encoding condition should not spread out during recall and should not correspond with directions and locations of the recalled picture.

This hypothesis has previously been supported by Laeng and Teodorescu (2002), who in two related experiments reported that participants, who were instructed to maintain their gaze centrally during visual encoding of a figure, spontaneously maintained their gaze centrally while looking at a blank screen during recall. However, Laeng and Teodorescu's visual stimuli were somewhat inadequate. For example, in their first experiment participants encoded 6 × 6 grid patterns consisting of five black grids at 10 × 10 cm. In their second, small single objects—2 × 3 cm color pictures of tropical fish—located in one of the four corners of the screen, were used. It is reasonable to question these stimuli on the grounds that one could encode and maintain them in working memory without moving the eyes (e.g., Irwin, 1996).

The question can also be raised whether the recall instruction in the study by Laeng and Teodorescu (2002) to 'build a visual image of what you have just seen' does not actually induce the participants to imitate their perceptual encoding behavior (Intons-Peterson, 1983; Pylyshyn, 1981). This is a classic objection to mental imagery experiments, and the way to avoid this critique is to never explicitly instruct the participants to "build visual images" or at all instruct them to do something that involves mental imagery.

Conflicting results to those by Laeng and Teodorescu (2002) have been reported by Richardson and Spivey (2000, Experiment 5), who also manipulated the encoding phase in one of their experiments. The visual stimuli consisted of a 2 × 2 grid in this experiment, and eye movements were manipulated by introducing a mask with a window above the grid pattern. For one group of participants, the window traveled to the four segments of the grid, which encouraged them to move their eyes to each of the four segments. For another group of participants, the window was static and located in the center of the screen, and instead the grid behind the window moved to bring each segment into view. Every time a segment was in view for either of the two groups, a spinning cross appeared within it and was synchronized with a spoken statement such that they would be perceived as one event. The spinning cross was designed to attract the participants' visual attention while they listened to the spoken statement. In the recall phase, the empty 2 × 2 grid (without the mask) was present on the screen for both groups, and they were asked to give a yes/no answer to a question that was related to one of the four statements in the encoding phase. Results revealed that both groups looked significantly more within the segment that was associated with a concerned statement than within the other three segments. No significant difference was found between the groups with regard to how much they looked within the critical segments. Therefore, contrary to Laeng and Teodorescu (2002); Richardson and Spivey (2000) concluded that eye movements during encoding are not required per se to cause systematic saccades to blank regions during recall.

However, the study by Richardson and Spivey (2002) differed in several respects from the one by Laeng and Teodorescu (2002). First, a combination of visual and spoken information was used during encoding. Second, the recall task depended exclusively on information from the spoken statements, that is, it was not dependent on visuospatial information from the encoding phase. Furthermore, visual features from the encoding phase were still present during recall (the empty 2 × 2 grid), which could have been used as important cues—or placeholders—for eye movement association.

Design

Experiment 1 consisted of an encoding phase and a recall phase. In the encoding phase, participants observed a complex picture with the instruction to fixate a cross in the center. In the recall phase, participants orally described the picture from memory while looking at a blank screen. Eye movements were recorded both during the encoding phase and during the recall phase.

As we wished to address more rigorously whether eye movements during pictorial recall are reinstatements of the oculomotor activity produced during encoding, the design of Experiment 1 differed from the studies of Laeng and Teodorescu (2002) and Richardson and Spivey (2000) in two crucial regards.

First, instead of using simple artificial visual stimuli we used a complex picture. Andersson, Ferreira, and Henderson (2011) have stressed the importance of investigating complex stimuli over simple artificial ones in studies of language-vision integration and have pointed out differences in how visual information is processed for complex natural scenes in comparison with simple artificial displays. Complex pictures are, for instance, represented both in terms of gist and associated details (Hollingworth, 2006; Bar, 2004), and the encoding procedure is different from encoding simple artificial displays of figures, grids, or small single objects (Foulsham & Underwood, 2008). Almost instantaneously the gist of the scene is encoded, and then different positions are fixated in temporal order, with more specific visuospatial information being encoded in relation to the overall scene context (Henderson, 2007).

Second, instead of an abstract task where participants are instructed to 'create a visual image' of the encoded scene, or a task where they are to answer yes/no questions associated with the encoding phase, the task here was to orally describe the picture from memory. This is a concrete recall task where several aspects of the encoded visuospatial information have to be retrieved. Thus, both the general spatial context of the scene and specific elements within it remain important for participants when the picture is verbally recalled.

Finally, by using the same stimulus picture as in the previous study (Johansson et al., 2006), we were able to compare the current experiment of central encoding with the previous study of free viewing.

Hence, the design can be described factorially as follows: There was one between-groups independent variable, experiment, with two levels (previous study, present experiment), and one within-groups independent variable, phase, with two levels (encoding, recall). Our dependent variables were the spatial dispersion of eye movements and eye movement correspondence with verbal data. These are described in full below. Figure 3 shows schematics of the design.

Participants

Twenty-two native Swedish speaking students at Lund University (11 female) participated in the experiment (mean age 21.7 years; SD = 2.2). All reported normal or corrected-to-normal vision.

Apparatus and Stimuli

Participants were seated 700 mm in front of a 480 × 300 mm monitor (running at 1680 × 1050 pixel resolution). Stimuli were presented via Experiment Center 2.4, while eye movements were measured using an SMI iView RED250, tracking binocularly at 120 Hz. Data was recorded using iView × 2.4 following five-point calibration plus validation (average measured accuracy was 0.39°;

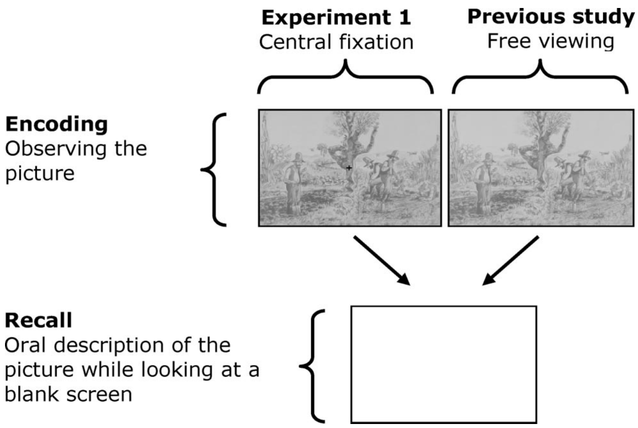

Figure 3. Schematics of Experiment 1 and how it compares to the previous study (Johansson et al., 2006).

SD = 0.31°). Fixations were detected with a dispersion threshold of 100 pixels and a duration threshold of 80 ms. BeGaze 2.4 and in-house MatLab scripts were used for analysis. The stimulus picture was a picture from Nordqvist (1990).

Procedure

To conceal the true objective of the experiment, participants were told that the experiment concerned pupil dilation in relation to mental workload. It was explained that we would be filming their eyes, but nothing was said about us recording their eye movements.

Following calibration participants received the following instruction:

> You will soon see a cross in the center of the display. This cross will be shown for five seconds. Then a picture will appear behind the cross for an additional 30 seconds. We want you to maintain fixation on this cross over the entire session. But try to perceive the picture behind the cross as well as possible. During this procedure we will measure your pupil size.

When they had understood the task instruction, they pressed the space bar and the fixation cross first appeared for five seconds against a blank background before the picture was presented.

Hereafter they received the following instruction:

> Now we want you to orally describe the picture that you tried to perceive when you were previously fixating the cross. You are free to describe it however you like and for as long as you like. Try to describe it as well as you can and tell us when you are finished. When you describe the picture you will look at a blank screen. We will again measure your pupil size. It is important that you do not close your eyes, but you may otherwise look straight ahead at the monitor, however you like.

Next, they pressed the space bar and the screen went blank. Then they orally described the picture. The experimenter always went behind a screen located behind the stimulus computer during testing.

During post experiment debriefing participants were asked what they thought the true objective of the experiment was.

Analysis

Only participants who were successful in maintaining central fixation during the encoding phase were included for analysis. We used a conservative criterion, excluding anyone who made one or more fixations to positions more than three degrees away from the center of the fixation cross (see Figure 4).

To test the central hypothesis of eye movement reinstatements between encoding and recall, we analyzed eye movements in two respects.

First, the overall spatial dispersion of eye movements was considered. Paired-samples t tests were run comparing spatial dispersion between the encoding phase and the recall phase.

Spatial dispersion alone does not capture information about how eye movements correspond to directions and positions however. Therefore, as in Johansson et al. (2006), a method for analyzing *eye movement correspondence* was used here also. This technique considers the association (i.e., correspondence) between eye movements and spatial information recounted in participants' verbal descriptions of the picture (and is outlined in full in Appendix A). There can be large individual differences in participants gaze patterns during recall (Johansson, Holsanova, & Holmqvist, 2011). For example, participants can look at a smaller part of the screen, but within this restricted area positions and directions when visualizing are frequently preserved in eye movements. Our correspondence measure accounts for such variability. Independent samples t tests were run, comparing the proportion of eye movements which correspond to directions and positions during recall.

For both spatial dispersion and eye movement correspondence, results were compared with the previous study of free viewing (Johansson et al., 2006). Cohen's (1988) d was used throughout to estimate the effect sizes.

Spatial dispersion. To analyze the overall spatial dispersion of the eye-tracking data, a modified version of the coverage measure proposed by Wooding (2002), as described in Holmqvist, Nyström, Andersson, Dewhurst, Jarodzka, and van de Weijer (2011, p. 367), was calculated for both the encoding phase and the recall phase. This measure was chosen over others because the impact of one fixation point in isolation is very small. Other coverage measures, such as the convex hull area (Goldberg & Kotval, 1999), weight all fixation points equally, and would therefore not be suited to these investigations. Wooding's measure is conservative, and infrequent distal fixations away from the main cluster do not disproportionately influence the overall spatial dispersion calculation.

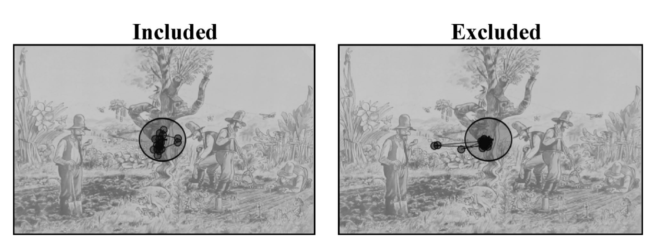

Figure 4. Gaze patterns for two participants under the central encoding restriction over the entire encoding phase (30 s). The circle represents the 3° threshold and was not present in the actual encoding phase. The participant to the left fulfilled the inclusion criterion, whereas the participant to the right did not.

The mathematics behind this measure are described in detail in Wooding (2002) and Holmqvist et al. (2011, p. 367). But the basic idea is as follows: an "attention map" is created by centering a Gaussian function at each fixation point (σ was set to span 10% of the screen width; $\sigma = 0.1 \times 1680$ pixels). Next, all the Gaussian functions are superimposed and the volume under the attention map, after being normalized to unit height, is used to estimate the spatial dispersion of the gaze pattern. Finally, the computed volume is normalized against its maximum theoretical value ($1680 \times 1050 \times 1$), which gives a proportion value between 0 and 1. Using this measure, an even distribution of fixations over the entire computer screen will yield a higher proportion value in comparison with densely packed pockets of fixations, which will return lower proportion values. The typical proportion range in scene perception is 0.2–0.4.

Eye movement correspondence. To analyze whether eye movements corresponded to directions and positions from the picture descriptions, a method combining verbal data with eye movements was used. This method was developed in Johansson et al. (2006) and is described in detail in Appendix A. The basic idea is to analyze whether eye movements correspond to the verbal focus of attention in a participants' scene recollection, according to *global* and *local* correspondence criteria. Global correspondence considers both position and direction, whereas local correspondence only considers direction. As a consequence of applying these criteria two binomial distributions are obtained: eye movements either correspond according to global or local criteria, or they do not. Interrater reliability (Cohen's kappa between two manual coders) is reported for the correspondence measures (Cohen, 1960).

The probability of a participant moving her eyes to a corresponding position or direction by chance is defined as 11% for global correspondence (four possible directions: up, down, left, right; two possible locations: full or half distance; or the eyes could be stationary, i.e., below the 1.1° amplitude threshold), and 20% for local correspondence (only four directions: up, down, left, right). Data are then analyzed for significance between total number of corresponding eye movements and the expected number corresponding by chance using Wilcoxon signed-ranks test (Wilcoxon, 1945).

Results and Discussion

None of the participants saw through the true objective of the experiment, and on this ground the data from all were retained for analysis. Five participants, however, failed to maintain central fixation during encoding of the picture and were thus excluded. Consequently, data from 17 participants contribute to the following results. See Figure 5 for an example of gaze patterns for a typical participant over both the encoding and recall phases.

Spatial dispersion. Spatial dispersion was significantly larger in the recall phase than in the encoding phase for the current experiment of central encoding, $t(16) = 5.49$, $p < .001$, $d = 1.48$. The comparison of spatial dispersion in the recall phase between the current experiment of central encoding and the previous study of free viewing (Johansson et al., 2006) revealed a significantly larger spatial dispersion in the previous study of free viewing, $t(28) = 2.57$, $p < .05$, $d = 0.90$ (see Table 1).

Eye movement correspondence. Interrater reliability was found to be very high between the two coders ($\kappa = 0.82$, $p < .001$), who discussed potentially ambiguous categorizations, coming to agreement concerning each.

Proportion of eye movement correspondence was significantly above chance for recall in the current experiment of central encoding according to both local ($Z = 3.54$, $p < .001$, $r = .89$) and global ($Z = 2.08$, $p < .05$, $r = .52$) correspondence coding. No significant difference was found for either local or global correspondence coding between the current experiment of central encoding and the previous study of free viewing (Johansson et al., 2006). See Table 1 for average values.

Discussion. The results revealed that the spatial dispersion of eye movements was significantly larger in the recall phase than in the encoding phase under the central encoding restriction, and most importantly that there was a significant degree of eye movement correspondence, both for locations (global correspondence) and directions (local correspondence). Additionally, no significant difference was found between the current experiment of central encoding and the previous study of free viewing when eye movement correspondence was compared. Consequently, despite central fixation during encoding, eye movements spread out during pictorial recall and corresponded to directions and locations from the original picture. Oculomotor events during recall can therefore not be exact recapitulations of those produced during encoding. In-

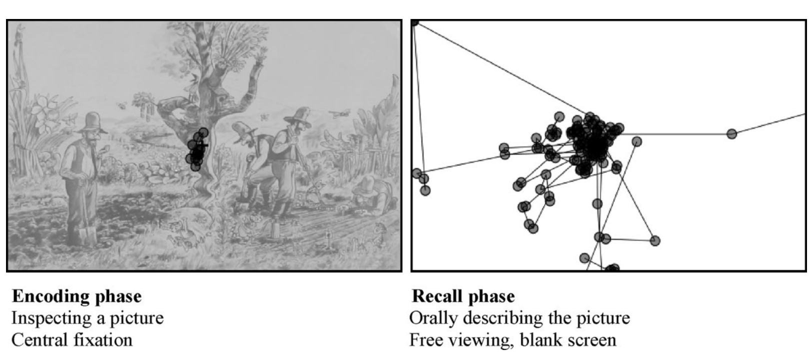

Figure 5. Gaze patterns for one typical participant under the central encoding restriction after both the encoding phase and the recall phase.

Table 1
Mean Values for Spatial Dispersion and Proportion Eye Movement (EM) Correspondence for the Current Experiment Under the Central Encoding Condition and for the Previous Study of Free Viewing, With Standard Deviations in Parentheses

	Experiment 1		Previous study	
	Encoding central fixation	Recall free viewing	Encoding free viewing	Recall free viewing
Spatial dispersion	.06 (.01)	.10 (.02)	.29 (.04)	.15 (.08)
EM correspondence				
Global	—	34 (34)%	—	54 (37)%
Local	—	76 (24)%	—	75 (28)%

stead they seem to operate independently. These results are similar to those by Richardson and Spivey (2000) but are in stark contrast to those by Laeng and Teodorescu (2002).

The spatial dispersion of the gaze patterns during the recall phases was, however, significantly smaller in the current experiment of central encoding when compared with the previous study (Johansson et al., 2006) of free viewing. There was also a tendency (p = .09) for a lower proportion of global correspondence (see Table 1 above). Does this mean the restriction to maintain fixation during encoding affects eye movements, making them more centrally focused during recall?

We do not believe that this is the case. Instead we propose that this is an effect of not being able to look at the entire picture during encoding. When participants fixated the cross and tried to observe the picture, only a small part of it was perceived with foveal vision. Most parts were therefore encoded via parafoveal or peripheral vision, which makes it very hard to apprehend more than the general layout. Cluttered information-dense scenes like our stimulus picture are also likely to be sensitive to visual crowding effects (e.g., van den Berg, Cornelissen, & Roerdink, 2009). Therefore, by reducing the participants' possibility to inspect the picture, their picture descriptions during recall became rather sparse and short, and they primarily referred to picture elements that were in focus during encoding. For instance, Figure 5 shows a participant who mostly remembered and described aspects of the tree in the middle of the picture. In accordance, she also looked a lot in the center of the screen. However, she did mention "a man to the left of the tree," "three men to the right of the tree," "something in the top left corner," "the cat below the tree," and "the cows in the background." Each of these statements was accompanied by corresponding eye movements.

However, in order to rule out the possibility that the lesser spatial dispersion and lower proportion of global correspondence was caused by the central gaze restriction, a new experiment was conducted.

Experiment 2: Central Fixation During Encoding of a Spoken Scene Description

Experiment 2 repeated the central encoding condition from Experiment 1, but instead of using pictorial stimuli this experiment used a spoken scene description. The reason to investigate central encoding also for a spoken scene description was primarily to introduce a stimulus which does not depend on parafoveal and peripheral information. Consequently, this experiment was able to determine whether the gaze patterns of lesser spatial dispersion from Experiment 1 were driven by the inability to perceive peripheral and parafoveal information during encoding or whether they were caused by the restriction of central encoding.

Based on the results from Experiment 1, we now hypothesized that oculomotor events during recall are not reinstatements of those produced during encoding, and therefore participants' gaze patterns should spread out during recall, in correspondence with directions and locations of objects in the recalled scene description. Based on the results from the previous study of free viewing (Johansson et al., 2006), we expected that eye movements during recall of a spoken scene description should not be executed any differently than during pictorial recall (despite central encoding).

Furthermore, because Richardson and Spivey (2000) used both visual and auditory stimuli and reported different results from Laeng and Teodorescu (2002), Experiment 2 also had a secondary goal of investigating whether there is a modality-specific effect for how eye movements during encoding are related to those during recall.

Design

Experiment 2 consisted of an encoding phase and a recall phase. In the encoding phase, the participants listened to a spoken scene description with the instruction to fixate a cross in the center the screen. In the recall phase, the participants retold the scene description from memory while looking at a blank screen without eye movement restrictions.

The same scene description was used as in Johansson et al. (2006), but with a slight modification. In the previous study the scene description started with the instruction to "imagine a two-dimensional picture," and a picture was explicitly referred to. This may be challenged on the grounds that the wording may induce participants to behave according to the desired behavior (Intons-Peterson, 1983) and/or to use their tacit knowledge (Pylyshyn, 1981) of how they *think they should* visualize. Therefore, we removed the parts where the description referred to an actual picture and removed the initial "imagine" instruction. Otherwise, the task was identical.

The orthogonal design structure was preserved from Experiment 1: two levels of experiment (previous study, present experiment), and two levels of phase (encoding, recall). The dependent variables were again spatial dispersion and eye movement correspondence. Figure 6 shows the schematics of the design.

Participants

Twenty native Swedish speaking students at Lund University (10 female) participated in the experiment (mean age 21.8 years; $SD = 2.3$). All reported normal or corrected-to-normal vision.

Apparatus and Stimuli

The eye tracker and the overall set-up were the same as in Experiment 1. Average measured accuracy was 0.37° (SD = 0.28°). The auditory stimulus was a modified prerecorded scene description from Johansson et al. (2006). An English translation of the scene description follows, and Figure 7 shows spatial schematics of it.

> In the center, right in front of me, there is a large, green spruce. At the top of the spruce, a bird is sitting. To the left of the spruce—to the far left—there is a yellow house with a black tin roof and with white corners. The house has a chimney on which a bird is sitting. To the right of the large spruce—to the far right—there is a tree as high as the spruce. The leaves of the tree are colored yellow and red. Above the tree a bird is flying. Between the spruce and the tree is a man in blue overalls, raking leaves. Below the spruce, the house, the tree, and the man, that is, in front of them, there is a long red fence, which runs all the way from left to right. At the left end, a bike is leaning against the fence. Just to the right of the bike is a yellow mailbox. On top of the mailbox, a cat is sleeping. Below the fence, that is, in front of and along the fence, a road leads from the left side to the right side. On the road, to the right of the mailbox and the bike, a black-haired girl is bouncing a ball. To the right of the girl, a boy in a red cap is sitting and watching her. To the far right along the road, a lady in a big red hat is walking with some books under her arm. Just to the left of her, on the road, a bird is eating a worm.

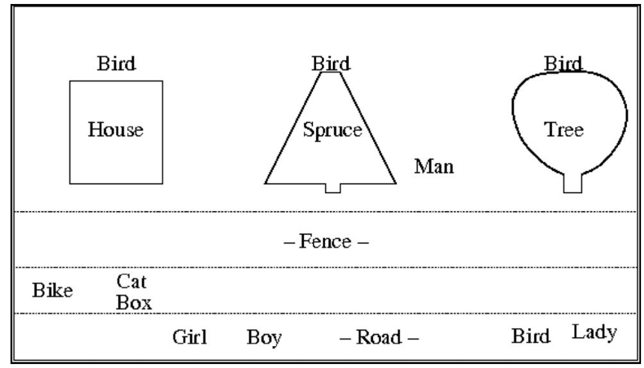

Figure 7. Spatial schematics of the spoken scene description. From "Pictures and spoken descriptions elicit similar eye movements during mental imagery, both in light and in complete darkness," by R. Johansson, J. Holsanova, and K. Holmqvist, 2006, *Cognitive Science, 30,* p. 1057. Copyright 2006 by Cognitive Science Society, Inc. Reprinted with permission.

Procedure

The procedure was repeated from Experiment 1, with minor exceptions pertaining to the auditory stimulus. For example, the instructions differed slightly:

> You will soon see a cross in the center of the display. This cross will first be shown for five seconds. Then you will listen to a prerecorded spoken scene description. We want you to maintain fixation on this cross over the entire session and listen to the scene description as carefully as possible. During this procedure we will measure your pupil size.

Also, the timing of stimulus presentation differed somewhat, the auditory scene description being 98 seconds in length. After hearing it participants received recall instructions which were also slightly adjusted:

> Now we want you to orally retell the scene description that you have just listened to. You are free to retell it with your own words, and you do not have to retell it in the same order as you heard it. Try to retell it as well as you can, and tell us when you are finished. While you retell the scene you will look at a blank screen. We will again measure your pupil size. It is important that you do not close your eyes, but you may otherwise look straight ahead at the monitor, however you like.

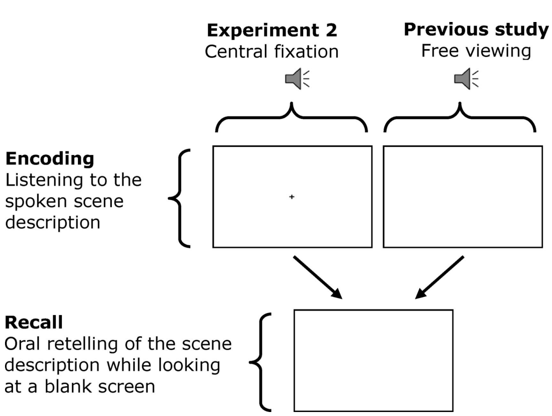

Figure 6. Schematics of Experiment 2 and how it compares with the previous study (Johansson et al., 2006)

Analysis

Eye movement data were analyzed the same way as in Experiment 1, with the same exclusion criterion with regard to maintaining central fixation.

Results and Discussion

None of the participants saw through the true objective of the experiment, and on this ground the data from all were retained for analysis. However, four participants failed to maintain central fixation during encoding of the spoken scene description and were thus excluded. Consequently, data from 16 participants contribute to the following results. See Figure 8 for an example of gaze patterns for a typical participant after both the encoding phase and the recall phase.

Spatial dispersion. Spatial dispersion was significantly larger in the recall phase than in the encoding phase for the current experiment of central encoding, $t(15) = 4.59$, $p < .001$, $d = 1.64$. No significant difference was found between the current experi-

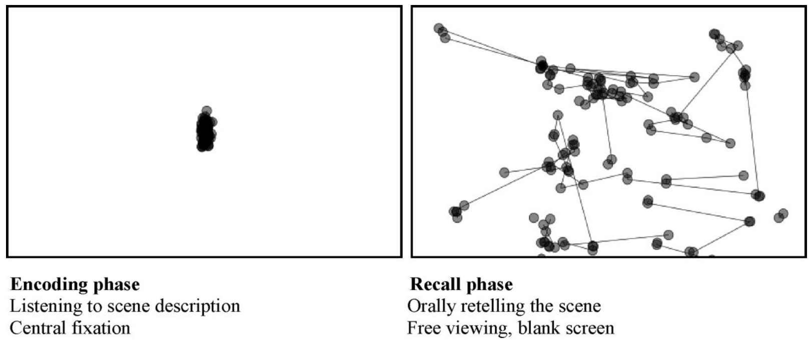

Figure 8. Gaze patterns for one typical participant under the central encoding restriction after both the encoding phase and the recall phase.

ment of central encoding and the previous study of free viewing when spatial dispersion in the recall phase was compared (see Table 2).

Eye movement correspondence. Interrater reliability was again high between the two coders ($\kappa = 0.83$, $p < .001$), who discussed categorizations as before.

Proportion of eye movement correspondence was significantly above chance for recall in the current experiment of central encoding according to both local ($Z = 3.47$, $p < .001$, $r = .87$) and global ($Z = 3.54$, $p < .001$, $r = .89$) correspondence coding. As in Experiment 1, no significant differences were observed between these eye movement correspondence data and those of the previous study. See Table 2 for average values.

Discussion. Again the results revealed that the spatial dispersion of eye movements was significantly larger in the recall phase compared with the encoding phase, where fixation was restricted. Moreover, there was again a significant degree of eye movement correspondence, indicating that participants shifted their gaze in directions and to positions according to the auditory scene description. Additionally, no significant difference was found between the current experiment and the previous study when eye movement correspondence was compared. Taken together these results clearly demonstrate that prohibiting eye movements, even during auditory perception, does not affect their execution when reactivating their mental model of the scene.

Furthermore, as encoding of the spoken scene description in Experiment 2 did not recruit parafoveal or peripheral visual processing, we can rule out the possibility that the less dispersed gaze patterns in Experiment 1 were caused by the central encoding restriction per se. Here, too, central fixation was maintained at encoding, and still a high degree of eye movement correspondence was observed at recall, making it more likely that the reduced spatial dispersion values of Experiment 1 reflect hampered peripheral processing, not reinstatements of oculomotor behavior.

Our results from Experiments 1 and 2 are similar to Richardson and Spivey (2000, Experiment 5: eye movements during encoding were not required to cause systematic saccades to blank regions during recall). In line with these findings we conclude that eye movements during scene recollection are not reactivations of the oculomotor activity produced during encoding. Instead, eye movements between encoding and recall operate independently. Consequently, the results of Laeng and Teodorescu (2002) cannot be generalized to scene recollection of more complex stimuli. Their results are, as we discussed above, most likely a consequence of simple artificial stimuli and/or task induced demands.

Do these results mean that eye movements to blank spaces are purely epiphenomenal and that they do not have an active role in how the recalled scenes are retrieved from memory?

Experiment 3 – Central Fixation During Recall of a Complex Picture

Experiment 3 was designed to test whether eye movements during pictorial recall are purely epiphenomenal or whether they have a functional connection with how the original picture is recalled and remembered. The method for doing this was to allow

Table 2
Mean Values for Spatial Dispersion and Proportion Eye Movement (EM) Correspondence for the Current Experiment Under the Central Encoding Condition and for the Previous Study of Free Viewing, With Standard Deviations in Parentheses

	Experiment 2		Previous study	
	Encoding central fixation	Recall free viewing	Encoding free viewing	Recall free viewing
Spatial dispersion	.06 (.01)	.12 (.04)	.16 (.08)	.15 (.07)
EM correspondence				
Global	—	45 (21)%	55 (37)%	55 (33)%
Local	—	70 (18)%	64 (35)%	75 (28)%

participants to look freely during encoding of a picture, and then prohibit eye movements during recall. This eye movement restriction will throughout this article be referred to as *central recall*.

We hypothesized that if eye movements during recall do not have a functional connection with how the picture is retrieved from memory then participants' picture descriptions should not be hindered by the requirement to maintain central fixation.

This hypothesis is supported by a series of experiments that have not found any memory facilitation for eye movements to nothing (Hoover & Richardson, 2008; Richardson & Kirkham, 2004; Richardson & Spivey, 2000; Spivey & Geng, 2001). However, rather simple artificial stimuli have been used in the encoding phase in all of those studies. Richardson and Spivey (2000) divided the screen into four quadrants and used movies of talking heads or animated objects synchronized with spoken statements in either of these quadrants. Spivey and Geng (2001) used four different shapes in the corners of a 3 $\times$ 3 grid. Richardson and Kirkham (2004) used animated objects to the left or right on the screen synchronized with spoken statements, while Hoover and Richardson (2008) used animated moles that emerged or descended in four locations of the screen in synchrony with a spoken fact.

Furthermore, the recall task has only to a small degree, or not at all, been dependent on visuospatial information from the encoding phase. No eye movement manipulations have been used. Moreover, it has been common for participants to either give a binary answer (yes/no or true/false) to statements that were related to the spoken information (Hoover & Richardson, 2008; Richardson & Kirkham, 2004; Richardson & Spivey, 2000), or to give a short answer to questions about the color or the shape of the absent object (Spivey & Geng, 2001).

In contrast to the studies above, Laeng and Teodorescu (2002) reported that participants who were free to look at single pictures of tropical fish in one of the four corners of the screen, maintaining their gaze centrally during a subsequent recall phase, showed a decrease in response accuracy to questions about the fishes' properties. This was in comparison with a free viewing group. But because of the reasons discussed above (Experiment 1 and 2), the generality of these findings is questionable, and we sought to shed further light on the perceptual and cognitive factors underlying them.

Design

Experiment 3 consisted of two conditions: (1) a control condition of *free viewing* where both encoding and recall of a picture was performed with free viewing; and (2) a *central recall* condition, where recall was performed with the instruction to fixate a cross in the center of the display. The two conditions were randomized for order. Participants' eye movements were recorded both during the encoding phase and during the recall phase.

To address the hypothesis in hand, the current experiment differed from previous studies (e.g., Laeng & Teodorescu, 2002; Spivey & Geng, 2001) in the same two regards as Experiment 1. First, instead of using simple artificial visual stimuli we used complex pictures. Second, a concrete recall task was used where several aspects of the visuospatial information from the encoded stimuli had to be retrieved.

To ensure that any effects in how the picture was described from memory were driven by the central recall restriction and not by individual differences in picture description style, the current experiment also introduced a control condition of free viewing. Large individual differences in how pictures are described have been demonstrated by Holsanova (2001, 2008). We also wanted to ensure that any effects in how the pictures were described from memory were not driven by specific properties of the stimulus pictures. Therefore, under the central recall condition, we used the same stimulus picture that was used in Experiment 1 and in our previous study (Johansson et al., 2006). But in the control condition of free viewing, a different picture was introduced. This picture was chosen with careful consideration with regard to comparability. Both pictures were complex scenes illustrating a depiction of a real environment. They both consisted of an immovable background with a similar amount of movable discrete objects. They both depicted both states and events and consisted of multiple agents (persons and animals). Furthermore, they both depicted a scene structure that was nameable, and they both had routes to semantic interpretation. Finally, they were both rich in color and detail.

This design allowed us to analyze the data in two ways where both individual differences and stimuli were controlled for. First, by comparing the data from the central encoding condition with the control condition of free viewing, the same participants were compared within the group over two encoded stimulus pictures. Second, by comparing the data from the central recall condition with the previous study of free viewing (Johansson et al., 2006) and with the central encoding condition of Experiment 1, three groups of participants were compared for the same encoded stimulus picture. This controls for any effects that may emerge as a result of the fixation restriction either during encoding or recall.

Hence, the design can be described factorially as follows: there was one between-groups independent variable, experiment, with three levels (Experiment 3, Experiment 1, previous study), and one within-groups independent variable, phase, with two levels (encoding, recall). For the present experiment (Experiment 3) we included a control condition of free viewing, which all participants of Experiment 3 completed. Our outcome measures were based on the verbal data during recall (total time on task, total number of idea units, verbal focus, correctly mentioned objects and correctly mentioned locations). These are described in full below. See Figure 9 for schematics of this design.

Participants

Twenty-one native Swedish speaking students at Lund University (12 female) participated in the experiment (mean age 23 years; $SD = 4.9$). All reported normal or corrected-to-normal vision.

Apparatus and Stimuli

The eye tracker and the overall set-up were the same as in Experiments 1 and 2. Average measured accuracy was 0.42° ($SD = 0.27°$).

Procedure

For both the control condition and the central recall condition participants received the following instruction:

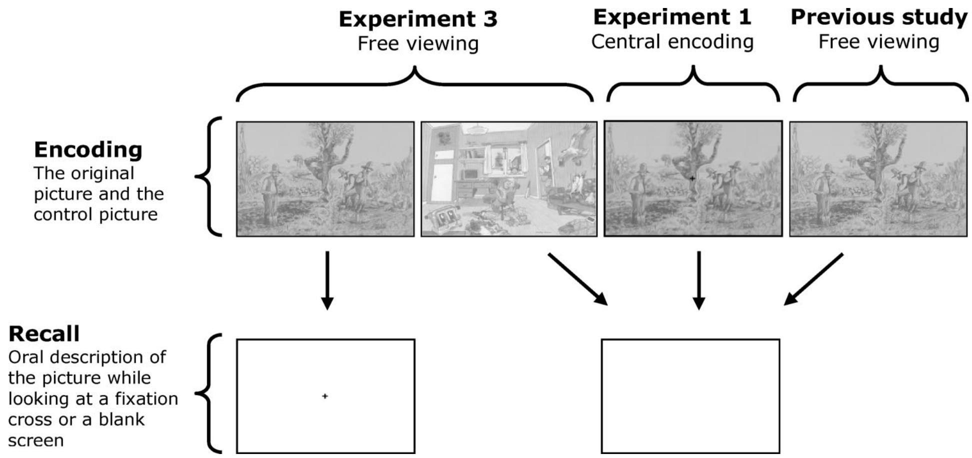

Figure 9. Schematics of Experiment 3 and how it compares with Experiment 1 and the previous study (Johansson et al., 2006).

> You will soon see a picture. We want you to study the picture as thoroughly as possible. The picture will be shown for 30 seconds. During this time we will measure your pupil size.

When they had understood the task instruction, they pressed the space bar and the picture appeared on the display. Hereafter, for the control condition, they received the following instruction:

> Now we want you to orally describe the picture that you have just studied. You are free to describe it however you like and for as long as you like. Try to describe it as well as you can and tell us when you are finished. When you describe the picture you will look at a blank screen. We will again measure your pupil size. It is important that you do not close your eyes, but you may otherwise look straight ahead at the monitor, however you like.

For the central recall condition the instruction differed slightly in the last phrases:

> [. . .] When you describe the picture we want you to maintain fixation on a cross in the center of the display. We will again measure your pupil size. It is important that you always look at the cross during this procedure.

When they had understood the task instructions, they pressed the space bar and continued to the recall phase. During the control condition of free viewing the screen went blank, and during the central recall condition a cross appeared in the center of an otherwise blank screen. Then they orally described the picture.

Other procedural details are identical to experiments 1 and 2.

Analysis

Because this experiment was conducted to investigate whether pictorial recall is affected by central fixation, the eye movement data were primarily used to exclude participants who were not successful in maintaining central fixation throughout the recall procedure. The exclusion criterion was the same as in Experiment 1 and 2. However, to make sure that the control condition of free viewing was comparable with the previous study (Johansson et al., 2006), with regard to eye movements to blank spaces, an independent samples *t* test was also conducted with spatial dispersion of gaze patterns and eye movement correspondence as dependent variables. Cohen's (1988) *d* was used to estimate the effect sizes.

If the hypothesis that eye movements during recall do not have a functional connection with how the picture is recalled and remembered, then the participants' spoken descriptions should not be affected in style and/or impaired when it comes to how much they were able to recall in the central recall condition. To test this, we chose to analyze the verbal data in two regards, which will be used as a proxy for *how* the pictures were recalled and *how much* content the participants remembered from them.

First, the spoken descriptions were transcribed and analyzed based on a method developed by Holsanova (2008, pp. 2–38) where spoken discourse is segmented into idea units with different verbal focus. An idea unit is defined as the mental unit that expresses the conscious focus of attention (Holsanova, 2008, p. 7) and consists of the amount of information to which persons can devote their central attention at a specific time. The theory is based on Chafe's (1980, 1996) studies on limitations of cognitive capacities. Because of these limitations we usually express one idea at a time. Depending on our present needs, interests, and goals we can only focus our attention on small parts, one at a time. An idea unit is usually a phrase or a short clause, delimited by prosodic, acoustic features and lexical/semantic features. It has one primary accent, a coherent intonation contour, and is usually preceded by a pause, hesitation, or a discourse marker (Holsanova, 2001, p. 15f.). Chafe (1980, 1996) has further shown that the flow of speech proceeds in brief spurts and closely reflects the flow of thoughts. Consequently, an upcoming idea unit reflects a change in the speaker's information state and expresses different aspects of information. Active information is replaced by different information at approximately two-second intervals (Chafe, 1987, p. 22). The conscious focus of attention active in these idea units will be referred to as *verbal focus* (Holsanova, 2001, 2008).

In the analyses of the spoken discourse in the present study, we considered three different verbal foci:

1. Substantive foci—refers to idea units where speakers describe referents, states and events in the picture (*there was a boy; the man was digging; there was some kind of monster*).

2. Localizing foci—refers to idea units where speakers focus on spatial relations (*in front of the tree; in the top left corner; to the far left in the room).*

3. Summarizing foci—refers to idea units where speakers focus on elements at a higher level of abstraction and introduce them as a global gestalt (*it was a living a room; there were animals everywhere; there were three birds in the tree; it depicted a family; the walls were colored in blue*).

For these three verbal foci the attention to picture content and composition differ in distinct ways. For substantive foci a particular picture element or a specific portion of the recalled picture is at the focus of attention, whereas for localizing foci spatial relations among picture elements are at the focus of attention. For summarizing foci a higher level of abstraction is used, where several elements or general aspects of the recalled picture are attended to. Apart from these three verbal foci, the focus of attention can also be dedicated to different aspects that are not per se related to picture content and compositionality (Holsanova, 2008, pp. 2–38). For example, they can be evaluative (*it felt like something from a children's book*) or introspective (*perhaps there was something else there but I don't remember*). However, for picture descriptions from memory, these kinds of foci are very few and are hard to interpret with regard to picture content and compositionality. Therefore, they will not be considered in the present study.

Finally, except for verbal foci, this analysis also considered total time on task and total number of idea units.

Second, to measure how much the participants remembered from the picture and the scene description, the number of correctly mentioned objects and the number of correctly mentioned locations were counted. The criterion for an object to be considered correctly mentioned was to in an identifiable way mention a specific object that was actually in the picture. The criterion for a location to be considered correctly mentioned was to mention a spatial relation for one object that was correct with regard to the overall composition of the picture (*in the center there was a tree; to the far left there was a daffodil*) or with regard to another object that has previously been located and identified (*to the left of the tree there was a man with a shovel; below the tree the cat jumped around*).

Paired-samples *t* tests were conducted to analyze whether the verbal data differed between the control condition of free viewing and the central recall condition. In other words, the same group of participants was compared within the group over two stimulus pictures. A one-way between subjects ANOVA was conducted to analyze whether the verbal data differed between central recall (Experiment 3), central encoding (Experiment 1), and free viewing (the previous study). In other words, participants were compared over the same stimulus picture.

It is, however, important to notice that the recall of the picture was not a memory task per se. The participants were never explicitly told to mention as many objects as possible and to mention their spatial locations. They were always instructed to try their best but were free to describe the picture as they liked with their own words. Therefore, it is possible that some participants chose to focus more on general structures in the picture than on objects and their locations, and this may be further compounded by individual differences in motivation to complete the task (cf., Holsanova, 2001, 2008).

Results and Discussion

None of the participants saw through the true objective of the experiment, and data from all participants could be included in the results. However, one participant failed to maintain central fixation during recall. Consequently, 20 participants remained to report in the results.

Spatial dispersion and eye movement correspondence—control condition. Interrater reliability was very high (κ = 0.84, $p < .001$), and coders discussed categorizations as before.

Proportion of eye movement correspondence was significantly above chance for recall under the control condition of free viewing in the current experiment, by both local ($Z = 3.74$, $p < .001$, r = .84) and global ($Z = 3.58$, $p < .001$, $r = .80$) correspondence coding. No significant differences were found between the recall phase under the control condition of free viewing in the current experiment and the previous study (Johansson et al., 2006) with regard to these aspects. See Table 3 for average values.

Verbal data, interrater reliability. Interrater reliability was found to be very high between two independent coders of verbal foci ($\kappa = 0.82$, $p < .001$), who discussed discrepancies in categorizations as before.

Verbal data, paired *t* tests. No significant differences were found for total time on task or for number of idea units between the central recall condition and the control condition of free viewing. See Table 4 for average values.

The verbal foci between the central recall condition and the control condition of free viewing was significantly different with

Table 3
Average Values for Spatial Dispersion and Proportion Eye Movement (EM) Correspondence for the Current Experiment Under the Control Condition of Free Viewing and for the Previous Study of Free Viewing, With Standard Deviations in Parentheses

	Experiment 3 – Control condition		Previous study	
	Encoding free viewing	Recall free viewing	Encoding free viewing	Recall free viewing
Spatial dispersion	.35 (.06)	.14 (.06)	.29 (.04)	.15 (.08)
EM correspondence				
Global	—	58 (31)%	—	54 (37)%
Local	—	83 (26)%	—	75 (28)%

Table 4
Mean Values for Time, Idea Units, and Proportion of Verbal Foci for Participants' Oral Picture Descriptions During Recall, With Standard Deviations in Parentheses

Pictorial recall	Time (s)	Idea units (#)
Previous study		
Free viewing	76.3 (27.4)	16.9 (5.7)
Experiment 1		
Central encoding	47.8 (19.1)	10.0 (4.0)
Experiment 3		
Free viewing	67.5 (24.3)	16.1 (5.1)
Central recall	66.8 (23.1)	16.8 (5.4)

regard to substantive foci and summarizing foci. Under the central encoding condition, 29% (SD = 15) of the idea units were on average substantive foci and under the control condition of free viewing, 55% (SD = 15) of the idea units were on average substantive foci, $t(19) = 5.628$, $p < .001$, $d = 1.20$. Under the central encoding condition, 38% (SD = 16) of the idea units were on average summarizing foci and under the control condition of free viewing, 13% (SD = 10) of the idea units were on average summarizing foci, $t(19) = 6.830$, $p < .001$, $d = 0.94$. See Figure 10 for these comparisons.

For the comparisons of correctly mentioned objects and correctly mentioned locations a significant difference was found for objects but not for locations. Under the central encoding condition, 6.5 (SD = 3.4) correct objects were on average mentioned and under the control condition of free viewing, 9.3 (SD = 3.4) correct objects were on average mentioned, $t(19) = 3.088$, $p < .01$, $d = 0.76$. See Figure 10 for these comparisons.

Verbal data, between subjects ANOVA. There was a significant main effect of experiment for both total time on task, $F(2, 46) = 5.756$, $p < .01$, partial $\eta^2 = .21$, and number of idea units, $F(2, 46) = 9.740$, $p < .001$, partial $\eta^2 = .31$. Bonferroni post hoc tests revealed that total time on task (describing the picture from memory) was significantly shorter ($p < .01$) in Experiment 1 (central encoding) than in the previous study of free viewing, and number of idea units was significantly fewer ($p < .01$) in Experiment 1 (central encoding) than in both Experiment 3 (central recall) and the previous study of free viewing. The remaining comparisons did not reveal any significant differences. See Table 4 for average values.

For verbal foci, there was a significant main effect of experiment for both substantive foci, $F(2, 46) = 11.885$, $p < .001$, partial $\eta^2 = .35$, and summarizing foci, $F(2, 46) = 11.327$, $p < .001$, partial $\eta^2 = .34$, but not for localizing foci. Bonferroni post hoc tests revealed the following: the average proportion of idea units categorized as substantive foci was significantly higher ($p < .05$) in the previous study of free viewing (Johansson et al., 2006) than in both Experiment 3 (central recall) and Experiment 1 (central encoding). The average proportion of idea units categorized as summarizing foci was significantly lower ($p < .05$) in the previous study (free viewing) than in both Experiment 3 (central recall) and Experiment 1 (central encoding). The average proportion of idea units categorized as substantive foci was significantly higher ($p < .05$) in Experiment 3 (central recall) than in Experiment 1 (central encoding). The average proportion of idea units categorized as summarizing foci was significantly lower ($p < .05$) in Experiment 3 (central recall) than in Experiment 1 (central encoding). See Figure 11 for these comparisons.

For correctly mentioned objects and locations, there was a significant main effect of experiment for objects, $F(2, 46) = 12.542$, $p < .001$, partial $\eta^2 = .36$, but not for locations. Bonferroni post hoc tests revealed that significantly ($p < .05$) more correct objects were mentioned in the previous study of free

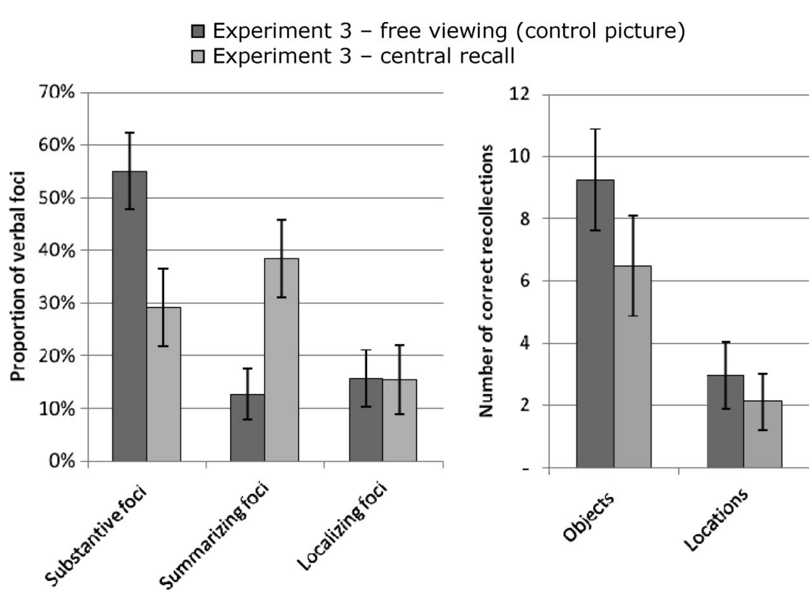

Figure 10. The same group of participants, two different pictures: Mean values for proportion of verbal foci, and correctly recalled objects and locations for participants' oral picture descriptions. Error bars represent 95 % confidence intervals.

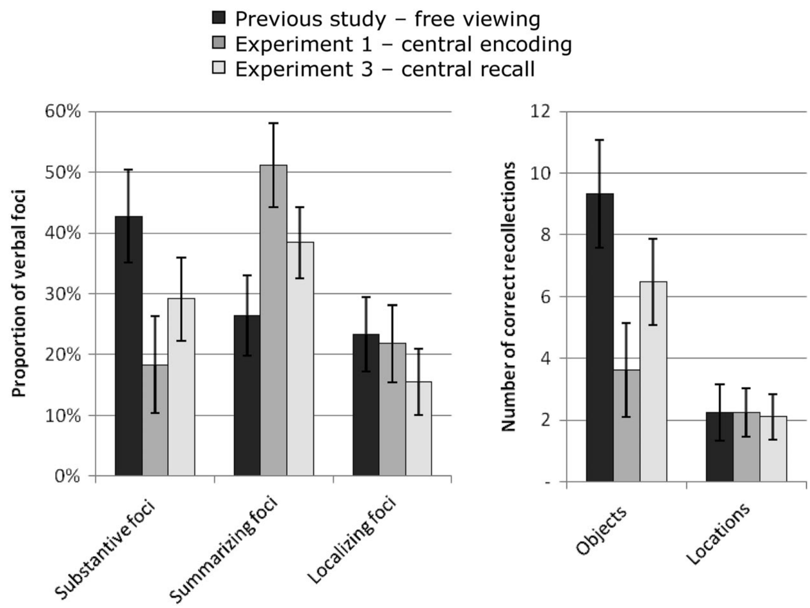

Figure 11. Three groups of participants, the same stimulus picture: Mean values for proportion of verbal foci, and correctly recalled objects and locations for participants' oral picture descriptions. Error bars represent 95% confidence intervals.

viewing (Johansson et al., 2006) than in both Experiment 3 (central recall) and Experiment 1 (central encoding), and significantly more correct objects were mentioned in Experiment 3 (central recall) than in Experiment 1 (central encoding). See Figure 11 for these comparisons.

Discussion. The results revealed significant differences in *how* the participants described the picture under the central recall condition when compared with recall under free viewing. This effect was reliable both when the results were compared over two comparable pictures within the group (see Figure 10), and over the same picture between groups (see Figure 11); that is, when compared with the results from the previous study of free viewing (Johansson et al., 2006). A larger proportion of idea units where the participants focused on elements at a higher level of abstraction and introduced them as a global gestalt (summarizing foci) was found under the central recall condition and a larger proportion of idea units focusing on objects, events and details (substantive foci) was found under the free viewing condition. Moreover, it was found that the focus of attention under the central encoding condition in Experiment 1 to an even larger extent was dedicated to summarizing foci than to substantive foci, when compared with both the current experiment of central recall and the previous study of free viewing (Johansson et al., 2006).

The results also revealed significant differences in *how many* correct objects the participants mentioned under the central recall condition when compared with recall under free viewing. This effect was reliable both when the results were compared over two comparable pictures within the group (see Figure 10), and over the same picture between groups (see Figure 11); that is, when compared with the results from the previous study of free viewing (Johansson et al., 2006). Moreover, it was found that under the central encoding condition in Experiment 1 objects were correctly mentioned even more infrequently when compared with both the current experiment of central recall, and the previous study of free viewing (Johansson et al., 2006).

Consequently, the central fixation restriction changed participants' verbal focus of attention during pictorial recall. The focus of attention was dedicated to more global aspects of the picture than to separate picture elements, and a more holistic stance was favored over a more analytical part-by-part one. Describing the picture more globally in this way was even more tangible under the central encoding condition of Experiment 1. Those descriptions were on average also much shorter and consisted of fewer idea units. Nonetheless, that is exactly what could be expected when the participants could only foveate a very limited part of the picture during encoding. That is, they were retrieving a very sparse memory representation of the picture. More remarkable is the finding that, despite a rich and detailed memory representation of the picture (free viewing during encoding), the central recall restriction in the current experiment affected and impaired the participants' ability to recall it.

However, no significant differences were found for localizing foci or number of correctly mentioned locations, and only around two correct locations were mentioned on average. This might seem a bit odd and rather scarce. Nonetheless, spontaneously describing scenes from memory is not a task where frequent and detailed descriptions about specific spatial relations are encouraged. Instead, as Holsanova (2008, p. 9) has pointed out, spatial expressions in the discourse are mostly guided by general conventions of picture composition. For example, concepts like "in the center," "in the background," "on the left" or "to the right" were usually sufficient when the describer referred to spatial aspects of the picture.

To sum up, this experiment revealed that pictorial recall was significantly affected by a central gaze restriction. We interpret this result as strong evidence that eye movements during pictorial recall cannot be purely epiphenomenal. They do have a functional connection to how the picture is recalled and remembered. These results are comparable with those by Laeng and Teodorescu (2002), who also used a central gaze condition during recall. But they are in contrast to those by Richardson and colleagues (Hoover & Richardson, 2008; Richardson & Kirkham, 2004; Richardson & Spivey, 2000; Spivey & Geng, 2001), who have not found any effects of memory facilitation. However, those studies did not manipulate eye movements as an independent variable during recall, and the recall tasks did only to a small degree, or not at all, depend on visuospatial information from the encoding phase. Furthermore, except for Spivey and Geng (2001), those studies have used multimodal stimuli of both visual and spoken information, and it is possible that they have failed to find a memory effect because their recall tasks have depended on the auditory and not the visual information.

Therefore, to rule out the possibility that the role for eye movements to blank spaces is different for scenes that have recently been encoded through visual perception and for abstract scenes that have been encoded on the basis of linguistic content through auditory perception, a final experiment was conducted.

Experiment 4 – Central Fixation During Recall of a Spoken Scene Description

Experiment 4 repeated the central recall condition from Experiment 3, but instead of using pictorial stimuli this experiment used a spoken scene description.

Based on the results from Experiment 3 we now hypothesized that eye movements during recall should have a functional connection also to how a verbal scene description is recalled and remembered. Consequently, the participants' scene retellings should be impaired by the central recall restriction when compared with the previous study of free viewing (Johansson et al., 2006). This hypothesis has partly been supported by Janssen and Nodine (1974), who reported that a group of participants, who maintained central fixation, was impaired in visual recall of spoken nouns when compared with another group, who was instructed to look as if they were actually looking at the noun. However, in contrast to those results, Hale and Simpson (1970) did not find any significant differences in response latencies for image generation of nouns under the experimental manipulations of making eye movements and not making eye movements.

Design

The experiment consisted of an encoding phase and a recall phase. In the encoding phase, the participants listened to a spoken scene description while looking at a blank screen. In the recall phase, the participants orally retold the scene description from memory with the instruction to fixate a cross in the center the display. Eye movements were recorded both during the encoding phase and during the recall phase.

By using the same scene description as in Experiment 2—which was a slightly modified version of the one used in Johansson et al., 2006)—we were able to compare the data from the current experiment of central recall with the previous study of free viewing (Johansson et al., 2006).

Additionally, by comparing the verbal data from Experiment 4 with the verbal data from Experiment 2, we were also able to investigate how a restriction of not being allowed to move the eyes to blank spaces during recall (central recall) compares to a restriction of not being allowed to move the eyes to blank spaces during encoding (central encoding). This comparison has important theoretical implications for understanding the mechanism that mediates eye movements to nothing. For instance, as we have shown in our previous study (Johansson et al., 2006), eye movements were to a high degree executed to corresponding positions on a blank screen also when the spoken scene description was encoded (for an example of this, see Figure 2 above).

Hence, the design can be described factorially as follows: There was one between-groups independent variable, experiment, with three levels (Experiment 4, Experiment 2, previous study), and one within-groups independent variable, phase, with two levels (encoding, recall). Our outcome dependent variables were based on the verbal data during recall (correctly mentioned objects and correctly mentioned locations). Figure 12 shows schematics of the design.

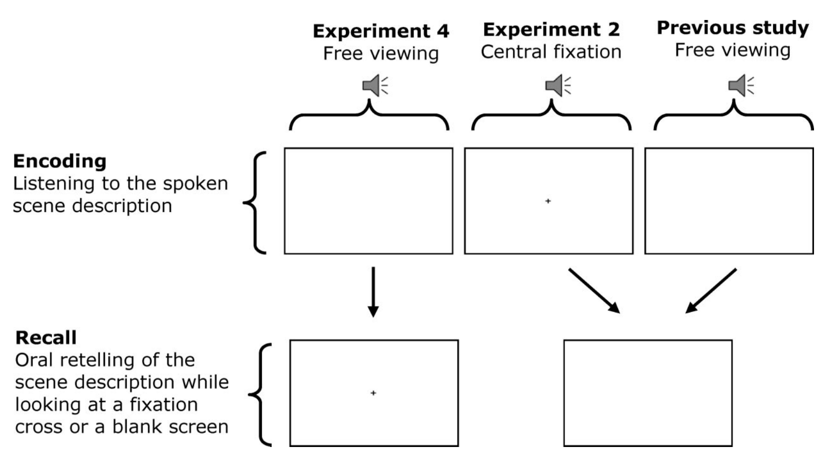

Figure 12. Schematics of Experiment 4 and how it compares with Experiment 2 and the previous study (Johansson et al., 2006).

Participants

Twenty students at Lund University—10 females and 10 males—participated in the experiment. All subjects reported either normal vision or vision corrected to normal (i.e., with contact lenses or glasses). All participants were native Swedish speakers. The mean age of the participants was 22.3 years ($SD = 3.7$).

Apparatus and Stimuli

The eye tracker and the overall set-up were the same as in Experiment 1, 2, and 3, and the auditory stimulus was the same as in Experiment 2. Average measured accuracy was 0.42° ($SD =$ 0.29°).

Procedure

For the encoding phase, participants received the following instruction:

> You will soon hear a prerecorded spoken scene description. We want you to listen to the description as carefully as possible. While you listen to the scene description you will look at a blank screen. During this procedure we will measure your pupil size. It is important that you do not close your eyes, but you may otherwise look straight ahead at the monitor, however you like.

When they had understood the task instruction, they pressed the space bar and the display went blank. Then the prerecorded scene description was played from speakers in front of the participants. Hereafter the participants received the following instruction:

> Now we want you to orally retell the scene description that you have just listened to. You are free to retell it with your own words and you do not have to retell it in the same order as you heard it. Try to retell it as well as you can and tell us when you are finished. While you retell the scene we want you to maintain fixation on a cross in the center of the display. We will again measure your pupil size. It is important that you always look at the cross during this procedure.

When they had understood the task instruction, they pressed the space bar and a cross appeared in the center of an otherwise blank screen. Then they retold the scene description.

Other procedural details are identical to Experiments 1, 2, and 3.

Analysis

If eye movements during recall have a functional connection with how the scene description is recalled and remembered, then the participants' retellings should be impaired when it comes to how much they were able to recall under the central recall restriction.

The recall of the scene description in the current experiment differed from recalling the pictures of Experiment 3 in two important regards. First, the scene description consists of a fixed amount of objects and a fixed amount of spatial references. There were 22 possible objects to remember and 16 possible locations to remember.

Second, the instruction during recall was to *retell* the scene description. The participants were—as for the picture description—instructed that they were free to retell it as they liked with their own words, and to try their best. But even if they were free in how they verbalized their retellings, they were still bound to the structure and contents of the previously heard scene description. Therefore, recall was to a large extent performed as a memory task where the participants struggled to remember as much from the spoken description as possible. Thus, the verbal data from their retellings were not, as in Experiment 3, dependent on different picture description styles and verbal foci. Consequently, the verbal data was analyzed with regard to the amount of correctly mentioned objects and the amount of correctly mentioned locations, but not with regard to verbal foci.

A one-way between subjects ANOVA was conducted to compare the effect of eye movement restrictions on correctly mentioned objects and correctly mentioned locations for central recall (Experiment 4), central encoding (Experiment 2), and free viewing (the previous study; Johansson et al., 2006).

Finally, because this experiment was conducted to investigate whether recall of a scene description is affected by central fixation the eye movement data were—as in Experiment 3—primarily used to exclude participants who were not successful in maintaining central fixation throughout the recall procedure. The exclusion criterion was the same as in Experiments 1, 2, and 3. However, to make sure that the encoding phase of free viewing in the current experiment was comparable with the encoding phase of the previous study (Johansson et al., 2006), with regard to eye movements to blank spaces, an independent samples t test was also conducted with spatial dispersion of gaze patterns and eye movement correspondence as dependent variables.

Results and Discussion

None of the participants saw through the true objective of the experiment, and data from all participants could be included in the results. However, four participants failed to maintain central fixation during recall. Consequently, 16 participants remained to report in the results.

Eye movement correspondence—encoding phase. The interrater reliability was found to be very high between the two coders ($\kappa = 0.89$, $p < .001$), who followed the same procedure as before.

Consistent with the previous study (Johansson et al., 2006), encoding of the verbal scene description generated gaze patterns that spread out with a high proportion of corresponding eye movements, which was significantly above chance according to both local ($Z = 3.39$, $p < .001$, $r = .82$) and global ($Z = 2.68$, $p < .01$, $r = .65$) correspondence (see Table 5). No significant differences

Table 5
Mean Values for Eye Movement Correspondence in the Encoding Phase Under the Central Recall Condition, as Well as for the Previous Study of Free viewing, With Standard Deviations in Parentheses

Encoding the scene description	Experiment 4 free viewing	Previous study free viewing
Spatial dispersion	.18 (.12)	.16 (.08)
EM correspondence		
Global	47 (40)%	55 (37)%
Local	62 (32)%	64 (35)%

were found between the current experiment and the previous study (Johansson et al., 2006).

Verbal data. There was a significant main effect of eye movement restriction for both correctly mentioned objects, $F(2, 41) = 4.9$, $p < .05$, partial $\eta^2 = .18$, and correctly mentioned locations, $F(2, 41) = 4.0$, $p < .05$, partial $\eta^2 = .15$. Bonferroni post hoc tests revealed that significantly ($p < .05$) more objects and locations were correctly mentioned (objects: $M = 18.2$, $SD = 1.7$; locations: $M = 10.5$, $SD = 2.0$) in the previous study of free viewing (Johansson et al., 2006) than in both the current experiment of central recall (objects: $M = 14.4$, $SD = 4.7$; locations: $M = 7.5$, $SD = 4.1$) and Experiment 2 of central encoding (objects: $M = 14.5$, $SD = 3.2$; locations: $M = 7.7$, $SD = 2.6$). However, no significant difference was found for either objects or locations between Experiment 4 (central recall) and Experiment 2 (central encoding). See Figure 13 for average values.

Discussion. Consistent with Experiment 3, the results revealed that participants performance during recall were affected by the central recall restriction. They mentioned significantly fewer objects and significantly fewer of their locations under the central recall condition, when compared with the previous study of free viewing (Johansson et al., 2006). Therefore, we conclude that a central gaze restriction during recall *does* impair memory retrieval, also for scenes that are encoded on the basis of linguistic content through auditory perception.

However, when compared with the central encoding condition of Experiment 2, no significant differences were found. Consequently, applying central fixation either during recall or encoding impaired the participants' performance of retelling the scene from memory. We interpret this result as further evidence that eye movements to blank spaces are not epiphenomenal and that they do have a functional role in all situations where visual imagery is engaged. By restricting these eye movements the perceptual simulation which constitutes visual imagery experiences is affected and impaired. Consequently, it did not matter whether the central gaze restriction was introduced when the scene was constructed and stored in memory (as in Experiment 2) or when the scene was retrieved and reconstructed from memory (as in Experiment 4). Both encoding and recall depend on visual imagery in these cases, and an eye movement restriction will therefore affect and impair the perceptual simulation in either case.

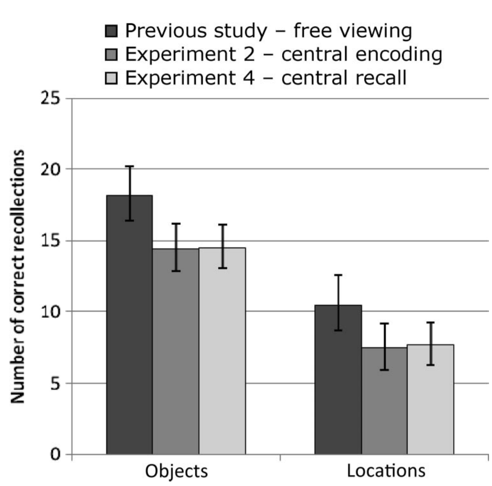

Figure 13. Mean values for correctly recalled objects and correctly recalled locations for participants' oral retellings of the scene description. Error bars represent 95% confidence intervals.

To sum up, we conclude that eye movements to blank spaces do have a functional role and that they do facilitate memory retrieval during scene recollection. However, we suggest that the functional role is not one of memory retrieval per se, but is related to the processes involved in mental simulations of perceiving visuospatial scenes. Moreover, it does not matter whether these scenes are generated as mental models directly from long-term memory or whether they are recalled in an encoding-recall procedure. Finally, we conclude that the reason why so few studies have been able to find a link between memory retrieval and eye movements during recall (cf., Ferreira et al., 2008; Richardson et al., 2009) is a result of recall tasks that only to a small degree, or not at all, have been dependent on visuospatial information from the encoding phase.

General Discussion

The four experiments presented in the current study produced the following main results:

(1) Experiment 1 showed that despite maintaining central fixation during encoding of a complex picture, eye movements spread out and correspond to positions and directions during recall of the original picture;

(2) Experiment 2 showed that maintaining central fixation during encoding of a spoken scene description did not affect how eye movements spread out and correspond to positions and directions during recall of the original scene;

(3) Experiment 3 showed that maintaining central fixation during pictorial recall affected how the original picture was recalled (the picture descriptions focused significantly more on general aspects of the picture than on objects, details, and events, and significantly fewer objects were reported than when recall was performed without eye movement restrictions);

(4) Experiment 4 showed that maintaining central fixation during recall of a scene description impaired how the scene was remembered during recall (significantly fewer objects and significantly fewer of their locations were remembered in the retellings of the scene description).

Our results have the following implications. First, the results from Experiments 1 and 2 showed that oculomotor events during scene recollection cannot be reinstatements of those produced during encoding. These results are similar to those by Richardson and Spivey (2000, Experiment 5) but are in stark contrast to those by Laeng and Teodorescu (2002). Consequently, Laeng and Teodorescu's (2002) result that eye movements during recall reenact those of encoding do not generalize to complex scene recollection. Their results are, as discussed previously, more likely to be a consequence of using simple artificial stimuli in combination with a recall task that is sensitive to task induced demands.

Second, even if there was no functional link between the oculomotor events during encoding and those from recall, the results from Experiment 3 and 4 showed that eye movements to blank spaces are not purely epiphenomenal and that they do have a functional connection to how a scene is recalled with respect to memory retrieval processes. Experiment 3 showed that central fixation during pictorial recall affected participants' focus of attention in a way where global and holistic reports of the original picture were favored over a detailed report of separate picture elements. Experiment 4 showed that central fixation during recall of the scene description reduced the number of objects and spatial referents that the participants mentioned in the retellings. Additionally, a similar impairment was also found for Experiment 2, where central fixation was applied during encoding of the scene description. These results clearly show that eye movements to blank spaces can facilitate memory retrieval and suggests that they may also facilitate memory encoding. Moreover, this aspect of the results is in line with those by Laeng and Teodorescu (2002) and those by Janssen and Nodine (1974). We propose that the reason why so few studies (Richardson et al., 2009) have found an effect of memory performance for eye movements to nothing depends on stimuli and recall tasks.

Theoretical implications of our results are elaborated below.

Why Would One Look at nothing?

Neisser (1967) argued that eye movements, or the processes that drives them, are actively associated with the construction of a visual image, and Hebb (1968) suggested that eye movements are necessary to assemble and organize "part images" into a whole visualized image. In a similar fashion, Laeng and Teodorescu (2002) suggested that eye movements are, during scene encoding, stored along with a visual representation of the scene, and that they are used to properly arrange the parts during scene recollection. They concluded that eye movements during recall reenact those of encoding and argued that eye movements have a necessary and functional role in the "constructing" of mental images. Brandt and Stark (1997) argued for a similar interpretation within the framework of "scanpath theory" (Noton & Stark, 1971a, 1971b). Scanpath theory proposes that eye movements during scene recollection are necessary and that they should play out in the same sequential order as during encoding.

Experiments 1 and 2 in the current study showed that eye movements were executed to blank spaces that corresponded with a mental image of the recalled scene even if the gaze had been maintained on a fixation cross during encoding. Therefore, contrary to previous research and theories, our results provide strong evidence that the oculomotor activity of overt eye movements to blank spaces during visual imagery cannot be a reactivation of the oculomotor activity produced during the perceptual encoding process. Such a strong connection would predict *similar* oculomotor activity in each phase, which is clearly not the case.

Our results are instead in line with those by Richardson and Spivey (2000), who in a situated and embodied account of cognition showed that eye movements are not required per se to cause systematic saccades to blank regions during recall. Such eye movements were instead explained in terms of *spatial indexing* (Richardson & Kirkham, 2004), where spatiotemporal information is used to link internal representations to objects in the world to reduce working memory demands (Ballard, Hayhoe, Pook, & Roe, 1997). A similar theoretical approach has been taken by Altmann (2004), who in a blank screen version of the so called 'visual world paradigm' (e.g., Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995), reported that participants' eyes returned to locations where objects had previously been located (see also, Altmann & Kamide, 2009; Knoerferle & Crocker, 2007). Altmann (2004) argued that spatial pointers are a part of the episodic trace associated with the objects from the encoding phase, and when this trace is subsequently activated the eyes are automatically driven toward the encoded locations.

Nonetheless, even if participants maintained central fixation during encoding in Experiments 1 and 2, they most likely engaged in covert attention shifts when trying to inspect the picture and possibly even when they listened to the scene description. For instance, Thomas and Lleras (2007) have shown that shifts in covert attention can produce identical results in a problem-solving task to overt eye movements, and Theeuwes, Belopolsky and Oliviers (2009) have shown that attention precedes eye movements and proposed that attention is the vehicle by which information is stored in visuospatial working memory. Moreover, it has been suggested that shifts of spatial attention are by-products of preparing saccadic eye movements, or nothing but such programming of saccades (e.g., Deubel & Schneider, 1996; Irwin & Gordon, 1998; Theeuwes, Oliviers & Chizk, 2005). In favor of this view, Geng, Ruff and Driver (2009) have reported similar activity in visual cortex for saccades to remembered locations and for programmed saccades that were never executed to a remembered location. Therefore, it is possible that the mechanisms that drive eye movements were, despite central fixation, encoded through covert attention shifts. Then in the subsequent recall phase of free viewing these covert attention shifts, or programming of saccades, were reenacted and manifested as overt eye movements.

Consequently, even if our results show that eye movements are executed to blank spaces during recall, despite central fixation during encoding, they do not necessarily contradict the idea that memory retrieval reactivates the processes of attention shifts involved in encoding (Kent & Lamberts, 2008).

The Consequence of Looking at Nothing

Laeng and Teodorescu (2002) reported that participants who were free to look at single pictures in one of the four corners of the screen and who maintained their gaze centrally during a subsequent recall phase, showed a decreased ability to give a correct response to questions about the pictures properties. Based on these results, Ferreira et al. (2008) have argued that looking at nothing facilitates memory retrieval of previously encoded information.

But a large body of research by Richardson and colleagues has failed to find a memory effect for eye movements to nothing (Hoover & Richardson, 2008; Richardson & Kirkham, 2004; Richardson & Spivey, 2000; Spivey & Geng, 2001). Those studies have, however, not focused on scene recollection and visual imagery per se and did not manipulate eye movements as an independent variable during recall. Additionally, stimuli have been relatively simple in all of these studies and the recall tasks have only to a small degree, or not at all, been dependent on visuospatial information from the encoding phase. The participants were either to give a binary answer (yes/no or true/false) to statements that were related to spoken information that were presented in synchrony with the visual stimuli in the encoding phase (Hoover & Richardson, 2008; Richardson & Kirkham, 2004; Richardson & Spivey, 2000) or to give a short answer to questions about the color or the shape of the visual stimulus from the encoding phase (Spivey & Geng, 2001).

In the current study, eye movements were manipulated as an independent variable, and the results from Experiment 3 and 4 showed that maintaining central fixation during recall affected recall of a complex picture and impaired memory performance for the recall of a scene description. These results provide strong evidence that eye movements to nothing have a functional role in memory retrieval.

We propose that the reason why we found an effect of memory performance and the studies by Richardson and colleagues did not are related to differences in task and stimuli. For example, Richardson and Spivey (2000) used single visual cues as stimuli, and the task was to give binary answers to statements, whereas we used stimuli of high complexity and a recall task that depended on complex visuospatial information from the encoding phase. Furthermore, the studies by Richardson and colleagues were not investigating eye movements during visual imagery of scenes per se, and they only analyzed what quadrant of the screen the eyes were located in during a response. Additionally, the information that was to be retrieved during recall was either unrelated to visuospatial information from the original stimulus (Richardson & Kirkham, 2004; Richardson & Spivey, 2000) or related to its color, shape, or texture (Spivey & Geng, 2001). Location or spatial relations were never part of the task during recall.

It is possible that the gain of looking at nothing is not primarily to process information related to appearance but to assist in visualizing locations and spatial relations. For instance, Postle, Idzikowski, Della Sala, Logie, and Baddeley (2006) have shown that a saccadic distraction task impairs memory for location, but not for shape. There is a large body of research which suggests that location memory and object memory are driven by the two relatively independent systems of the dorsal (“where”) pathway and the ventral (“what”) pathway (e.g., Gazzaniga, Ivry, & Mangun, 2008; Pollatsek, Rayner, & Henderson, 2002), and Farah, Hammond, Levine, and Calvanio (1988) have demonstrated that these two cortical pathways also drive spatial imagery (dorsal) and object imagery (ventral). Object imagery refers to information processing of scenes with regard to their appearance, color, shape, and texture (e.g., Farah et al., 1988; Paivio, 1991) and primarily activates neural architecture in the ventral stream (e.g., Gazzaniga et al., 2008), whereas spatial imagery refers to information processing of objects location, spatial relationships, movement, image maintenance, and spatial transformations (e.g., Farah et al., 1988; Paivio, 1991) and primarily activates neural architecture in the dorsal stream (e.g., Gazzaniga et al., 2008).

This explanation is further supported by the result that the participants under the central recall condition in Experiment 3 focused less on individual objects when compared with a condition of free viewing. Instead, the focus of attention was dedicated to more global and holistic aspects of the picture. This focus of attention is characteristic of imagery strategies that are more related to object imagery than to spatial imagery. For instance, Kozhevnikov, Kosslyn, and Shephard (2005), in their studies of individual styles, demonstrated that object visualizers encode and process images holistically, as a single perceptual unit, whereas spatial visualizers process images analytically, part-by-part.

Consequently, eye movements to blank spaces would play a bigger role when scenes of high complexity are to be arranged and recalled in a context, and where the scene recollection depends to a high degree on spatial relationships, spatial transformations, and image maintenance. However, Hollingworth (2006) has also shown that memory of specific details for an object was improved when it was presented at its original location and within the same scene context as it was originally encoded.

Moreover, applying central fixation either during recall (Experiment 4) or encoding (Experiment 2) impaired the participants’ performance of retelling the scene description from memory. Consequently it did not matter whether a mental model of the scene was constructed directly from long-term memory and encoded as a coherent scene or whether it was retrieved from memory in an encoding-recall procedure. Both encoding and recall depended on visuospatial imagery in these cases. The consequence of looking at blank spaces would therefore not be one of memory retrieval per se. For instance, Huijbers, Pennartz, Rubin, and Daselaar (2011) have recently identified separate neural correlates for imagery and memory retrieval. The gain of these eye movements could instead primarily be related to the processes that are involved in the perceptual simulations which constitute visuospatial imagery experiences, and specifically to processes that weight more on spatial imagery than on object imagery.

Cognitive Load and Central Fixation

One might, however, claim that the effects of memory performance in Experiments 3 and 4 are not a functional consequence of not being allowed to execute appropriate eye movements to blank spaces but rather are attributable to the cognitive cost of performing the additional task of maintaining central fixation. If so, the task of concentrating upon the fixation cross would tap into general cognitive resources to such an extent that the participants were impaired in their ability to describe the picture. This is an alternative explanation of the results which the current study cannot rule out. We do, however, propose that an explanation of this kind is not very likely for several reasons.

First of all, in Experiment 3, the central recall condition did not impair participants’ overall performance in the verbal reports. That is, there were no significant differences for total time on task or for number of produced idea units when the central recall condition and the control condition of free viewing were compared within subjects over two comparable pictures, or when the central recall condition in Experiment 3 and the previous study of free viewing (Johansson et al., 2006) were compared between-subjects over the same picture (see Table 4 above). Nonetheless, despite there being no apparent differences in the performance of the verbal reports, participants *changed* the focus of attention to a more holistic one under the central recall condition.

Moreover, if it is the case that the additional task of concentrating upon the fixation cross still drains cognitive resources in general, then this task should affect performance on any demanding cognitive task. Micic et al. (2010) have investigated performance in verbal fluency tasks and n-back tasks in a condition of central fixation and a condition of free viewing (looking at a blank screen), and reported no significant differences between these two conditions. Although verbal fluency tasks and n-back tasks are very different from describing a picture from memory, they are demanding cognitive tasks that require a great deal of executive attention, especially the n-back task. Furthermore, Postle et al. (2006) have investigated the relationship between visuospatial working memory and eye movements in a series of experiments. In an encoding phase in those experiments, the task was to mentally "put" eight numbers at eight different positions in an imagined 4 × 4 matrix according to an auditory instruction (e.g., *in the starting square put a 1* . . . in the next square to the right put a 2 . . . In the next square up put a 3 . . ., etc.). Subsequently, in a recall phase, the task was to report the position of those numbers. In Experiment 2 of those experiments (Postle et al., 2006) the recall task was performed both in a condition of free viewing and in a condition of central fixation. However, no significant difference in memory performance was found between the conditions.

Consequently, as the cognitive cost of keeping the eyes still did not impair cognitive performance in the study by Micic et al. (2010) or impair memory performance in the study by Postle et al. (2006), we do not find it likely that it would explain how the focus of attention changed in Experiment 3 or that memory performance was reduced in Experiment 2 and 4.

An "Inner Space" in Working Memory

Despite a large body of research, it is still an issue of debate whether visuospatial imagery involves an "inner space" of analog image representations (cf., Kosslyn, Thompson, & Ganis, 2006; Pylyshyn, 2002). Nonetheless, assuming that it does, the most common claim is that the internal image representations are constructed in a "visual buffer" of working memory (e.g., Kosslyn et al., 2006). Distance, location, and orientation of the internal image can be represented in this visual buffer, and it is possible to shift attention to certain parts or aspects of it (Kosslyn et al., 2006). Eye movements during visual imagery would thus be connected either with the construction of an image or when different parts of it are inspected through internal attention shifts. Scanpath theory (e.g., Brandt & Stark, 1997) and similar accounts, which assume a strong functional link between oculomotor events during encoding and recall (e.g., Laeng & Teodorescu, 2002), have argued that eye movements are used to arrange "part images" into a whole image when it is constructed in the visual buffer. However, if the visual buffer model (e.g., Kosslyn et al., 2006) is applied to the current study, the results from Experiments 1 and 2 would instead indicate that eye movements during recall are connected with attention shifts of *inspection*. Otherwise, those eye movements should not spread out and correspond to directions and positions from the original picture or scene description, that is, the construction of the mental image should depend on the central gaze restriction from encoding. Kosslyn has actually suggested that eye movements during mental imagery might reflect the process of "sliding" an image in discrete jumps "across" the visual buffer, which would be similar to how eyes, head, and body movements are ordinarily used in visual perception (Kosslyn, 1994, p. 367). This suggestion has, however, not been further elaborated (or investigated).

Nonetheless, as discussed above, it is still possible that the processes that drive eye movements are sufficient to encode visual information through covert attention shifts and that these processes become manifested in overt eye movements during recall.

Embodied Cognition, Situated Vision, and External Memory

Another approach where the role for eye movements is important during visual imagery, but not necessary for internal image construction per se, is to see them as a support that can relieve visuospatial working memory load. This is an embodied and situated view of cognition where motor processes and/or the visual world are used as an aid to memory. For instance, Ballard et al. (1997) have argued that eye movements are used as spatial indices to coordinate elements of an internal model with the external world, and where those looks are used to minimize working memory demands. Consequently, it may often be advantageous to minimize working memory load and direct the eyes toward perceptual cues in the external world when more information is needed. However, to what degree visuospatial information in spatial indexing relies on internal representation versus the external world is debatable, and it might seem peculiar how looks to "nothing" could offload working memory demands. Pylyshyn (e.g., 2002) has argued that there is no internal visuospatial information at all. Instead, all internal representations have a descriptive format that can be bound to visual features in the external world. But in Johansson et al. (2006) we showed that eye movements reflect content and spatial relations even when visualizations are performed in complete darkness, which eliminated the possibility of features in the external world as the exclusive explanation to those eye movements.

However, as Richardson et al. (2009) have pointed out, spatial indexing does not necessarily mean that the external memory store (e.g., O'Regan, 1992) is the only memory source at hand. Internal visuospatial representations can still have an important part of the memory traces that are activated during visual imagery tasks. Consequently, eye movements to blank spaces would not be necessary to achieve a simulation of perception but could instead serve an important and supportive role during demanding tasks that involve visuospatial imagery. In this view, eye movements would become more likely to appear when difficult recall tasks are performed.

This interpretation is also supported by our finding that those with lesser spatial imagery ability show an increased degree of eye movements that spread out and correspond with locations and directions during recall (Johansson et al., 2011). We interpreted this finding in an embodied view of cognition, where those with low spatial imagery ability "needed" an eye movement support to a larger degree than those with high spatial imagery ability. For instance, Keehner, Hegarty, Cohen, Khooshabeh, and Montello (2008) have shown a strong relationship between success in spatial reasoning tasks and the ability to offload cognition in the external world. Consequently, this could also explain why we found an effect of memory performance and why Richardson and colleagues did not (cf., Richardson et al., 2009). It is a much harder task to visualize and describe a complex picture or retell a scene description than to answer questions either related to, or unrelated to, the much simpler stimuli used in their studies. Furthermore, we used eye movement manipulations and tasks that were dependent on visuospatial information from the encoding phase. In support of this view, Scholz et al. (2011) have demonstrated that the degree of eye movements to blank spaces diminishes with practice. In their study the overall design from Richardson and Spivey (2000) was replicated, but with two important differences. First, the spoken information, which during encoding was associated with one of four areas on the screen, explicitly referred to visual scenes (e.g., *there is a place with a purple lighthouse, a sickle bay, and a wooden church*). Second, encoding and recall were repeated in a second and a third set of trials. Results revealed that the degree of eye movements to blank spaces significantly decreased with practice over trials.

Analogous results to our finding that an eye movement restriction affects and impairs recall have also been reported by Goldin-Meadow and colleagues in gesture research (e.g., Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001). In their studies, speakers were either permitted to use gestures or instructed not to use gestures while they explained how they had solved a previous math problem. It was found that speakers remembered more when they gestured than when they did not. Goldin-Meadow et al. (2001) interpreted their results in an embodied and situated framework of cognition where gestures can be used to offload memory resources related to spatial information.

Mental Simulations, Grounded Cognition and Enactive Theory

An interpretation of spatial indexing and where working memory demands can be offloaded through eye movements to either visual features or to blank spaces is also much in line with theories of grounded cognition (e.g., Barsalou, 2008) and simulation theory (e.g., Hesslow, 2011), where modal simulations, motor processes, and situated action underlie cognition. For instance, Spivey and Geng (2001) have proposed that eye movements should respond to an active visual representation regardless of whether that visual representation was generated by visual input, linguistic input, or from memory. The oculomotor system would in this sense not know the difference between scenes that are generated whole-cloth from long-term memory and those that are recalled in an encoding-recall procedure. Therefore, it would be no different if the perceptual simulation is driven by a preceding encoding phase (as in Experiments 1 and 3) or based on linguistic content and our semantic knowledge of the scene structure (as in Experiments 2 and 4). If this simulation is altered, by an eye movement restriction (as in the current study), it is plausible that either memory retrieval or encoding would be affected and impaired.

Thus, in an embodied and situated view, the cognitive system would in an opportunistic way use both external and internal memories to minimize memory demands and to use all information at hand to achieve a goal as effectively as possible. For example, Spivey and Geng (2001) reported a larger amount of eye movements to corresponding locations when the recall phase included the grid from the original encoding phase than when a completely blank screen was used. Moreover, in Johansson et al. (2006) we reported a significant degree of appropriate eye movements also when recall was performed in complete darkness. However, the proportion of eye movement correspondence for locations (global correspondence coding) was significantly lower when compared with a group of participants who looked at a blank screen (Johansson et al., 2006). Consequently, eye movements seem to increase with respect to how much support the external world can provide. On the other hand, if the external information provides no relevance to the task, then as Glenberg, Shroeder, and Robertson (1998) showed, the eyes are more likely to be averted. In a case like this, the most effective strategy for the cognitive system is instead to exclude external visual information to devote more resources to the memory task. Furthermore, in the experiments by Postle et al. (2006), memory performance was impaired when participants were instructed to look at a moving object, when compared with conditions of looking at a stationary object or for free viewing. In those experiments, both eye movements and external information were disruptive with regard to the recall task and the participants were not able to devote sufficient memory resources to it. This is exactly what would be expected in a situated and embodied account of cognition, where eye movements depend on both internal and external representations.

Finally, another prominent approach to explain eye movements during visual imagery has been favored by Thomas (2009). In his enactive theory of imagery (Thomas, 2009), perceptual experiences are not considered to consist of representations or states in one's brain. Instead imagery experiences arise when we actively seek for information, even if there is nothing to be found. Recalling a scene is in this sense similar to actually perceiving the scene, but without perceptual feedback from the scene itself. Consequently, a similar information-seeking behavior as in visual perception will be performed and, for instance, eye movements will be reenacted as if one were actually looking at the imagined scene. However, according to enactive theory, eye movements during recall are not—as in scanpath theory—a reactivation of the oculomotor activity produced during encoding. The term "reenactment" in these accounts refers to the reenactment of visual perception in general, not of behavior during a specific encoding phase.

Comparable to grounded cognition (e.g., Barsalou, 2008) and simulation theory (e.g., Hesslow, 2011), enactive theory (Thomas, 2009) would also predict that an eye movement restriction of central fixation during recall (as in Experiment 3 and 4) detrimentally affects memory performance.

Conclusion

In the current study we have shown that eye movements during visual imagery are not a reactivation of those produced during encoding. Despite this, we have demonstrated that eye movements to "nothing" still have an active and functional role, affecting memory recall (and encoding) if they are prohibited. This finding challenges the longstanding view that we direct our eyes solely to acquire new sensory visual input.

References

Altmann, G. T. M., & Kamide, Y. (2009). Discourse-mediation of the mapping between language and the visual world: Eye-movements and mental representation. *Cognition, 111,* 55–71. doi:10.1016/j.cognition.2008.12.005

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

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

Antrobus, J. S., Antrobus, J. S., & Singer, J. L. (1964). Eye movements accompanying daydreaming, visual imagery, and thought suppression. *Journal of Abnormal and Social Psychology, 69,* 244–252. doi:10.1037/h0041846

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

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience, 5,* 617–629. doi:10.1038/nrn1476

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

Bergstrom, K. J., & Hiscock, M. (1988). Factors influencing ocular motility during the performance of cognitive tasks. *Canadian Journal of Psychology/Revue Canadienne de psychologie, 42,* 1–23. doi:10.1037/h0084174

Bourlon, C., Oliviero, N., Wattiez, P., Pouget, P., & Bartolomeo, P. (2011). Visual mental imagery: What the head's eye tells the mind's eye. *Brain Research, 1367,* 287–297. doi:10.1016/j.brainres.2010.10.039

Bower, G. H., & Morrow, D. G. (1990). Mental models in narrative comprehension. *Science, 247,* 44–48. doi:10.1126/science.2403694

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

Brown, B. B. (1968). Visual recall ability and eye movements. *Psychophysiology, 4,* 300–306. doi:10.1111/j.1469-8986.1968.tb02771.x

Chafe, W. (1980). The deployment of consciousness in the production of a narrative. In W. Chafe (Ed.), *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production* (pp. 9–55). Norwood, NJ: Ablex.

Chafe, W. (1987). Cognitive constraints on information flow. In Russell Tomlin (ed.), *Coherence and grounding in discourse* (pp. 21–51). Amsterdam, The Netherlands: John Benjamins.

Chafe, W. (1996). How consciousness shapes language. *Pragmatics and Cognition, 4,* 35–54. doi:10.1075/pc.4.1.04cha

Clark, H. (1916). Visual imagery and attention: An analytical study. *The American Journal of Psychology, 27,* 461–492. doi:10.2307/1412990

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20,* 37–46. doi:10.1177/001316446002000104

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: L. Erlbaum Associates.

Deckert, G. H. (1964). Pursuit eye movements in the absence of a moving visual stimulus. *Science, 143,* 1192–1193. doi:10.1126/science.143.3611.1192

Demarais, A. M., & Cohen, B. H. (1998). Evidence for image-scanning eye movements during transitive inference. *Biological Psychology, 49,* 229–247. doi:10.1016/S0301-0511(98)00042-8

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research, 36,* 1827–1837. doi:10.1016/0042-6989(95)00294-4

Ehrlichman, H., & Barrett, J. (1983). "Random" saccadic eye movements during verbal-linguistic and visual–imaginal tasks. *Acta Psychologica, 53,* 9–26. doi:10.1016/0001-6918(83)90013-6

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, A., & Henderson, J. M. (2008). Taking a new look at looking at nothing. *Trends in Cognitive Science, 12,* 405–410. doi:10.1016/j.tics.2008.07.007

Finke, R. A. (1989). *Principles of mental imagery*. Cambridge, MA: MIT Press.

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*:,* 6, 1–17.

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

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

Gbadamosi, J., & Zangemeister, W. H. (2001). Visual imagery in hemianopic patients, *Journal of Cognitive Neuroscience. 13,* 855–866. doi:10.1162/089892901753165782

Geng, J. J., Ruff, C. C., & Driver, J. (2009). Saccades to a remembered location elicit spatially specific activation in the human retinotopic visual cortex. *Journal of Cognitive Neuroscience, 21,* 230–245. doi:10.1162/jocn.2008.21025

Glenberg, A. M., Schroeder, J. L., & Robertson, D. A. (1998). Averting the gaze disengages the environment and facilitates remembering. *Memory & Cognition, 26,* 651–658. doi:10.3758/BF03211385

Goldberg, J. H., & Kotval, X. P. (1999). Computer interface evaluation using eye movements: Methods and construct. *International Journal of Industrial Ergonomics, 24,* 631–645. doi:10.1016/S0169-8141(98)00068-7

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science, 12,* 516–522. doi:10.1111/1467-9280.00395

Goldthwait, C. (1933). Relation of eye movements to visual imagery. *The American Journal of Psychology, 45,* 106–110. doi:10.2307/1414189

Hale, S. M., & Simpson, H. M. (1971). Effects of eye movements on the rate of discovery and the vividness of visual images. *Perception & Psychophysics, 9,* 242–246.

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

Henderson, J. M. (2007). Regarding scenes. *Current Directions in Psychological Science, 16,* 219–222. doi:10.1111/j.1467-8721.2007.00507.x

Hesslow, G. (2011). The current status of the simulation theory of cognition. *Brain Research*. (Epub ahead of print.). doi:10.1016/j.brainres.2011.06.026

Hiscock, M., & Bergstrom, K. J. (1981). Ocular motility as an indicator of verbal and visuospatial processing. *Memory & Cognition, 9,* 332–338. doi:10.3758/BF03196967

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

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures.* New York, NY: Oxford University Press.

Holsanova, J., Hedberg, B., & Nilsson, N. (1999). Visual and verbal focus patterns when describing pictures. In Becker, Deubel, & Mergner (Eds.), *Current oculomotor research: Physiological and psychological aspects.* New York, NY: Plenum.

Holsanova, J. (2001). *Picture viewing and picture description: Two windows on the mind.* Doctoral dissertation. Lund University Cognitive Studies 83. ISSN 1101–8453. ISBN 91–856843-76.

Holsanova, J. (2008). Discourse, vision, and cognition. *Human Cognitive Processes 23.* Amsterdam, The Netherlands: John Benjamins Publishing Company.

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

Huijbers, W., Pennartz, C. M., Rubin, D. C., & Daselaar, S. M. (2011). Imagery and retrieval of auditory and visual information: Neural correlates of successful and unsuccessful performance. *Neuropsychologia, 49,* 1730–1740. doi:10.1016/j.neuropsychologia.2011.02.051

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*(3), 1–15.

Intons-Peterson, M. J. (1983). Imagery paradigms: How vulnerable are they to experimenters' expectations? *Journal of Experimental Psychology: Human Perception and Performance, 9,* 394–412. doi:10.1037/0096-1523.9.3.394

Irwin, D. E., & Gordon, R. D. (1998). Eye movements, attention and

trans-saccadic memory. *Visual Cognition, 5,*127–155. doi:10.1080/713756783
Irwin, D. E. (1996). Integrating information across saccadic eye movements. *Current Directions in Psychological Science, 5,* 94–100. doi:10.1111/1467-8721.ep10772833
Jacobson, E. (1932). Electrophysiology of mental activities. *The American Journal of Psychology, 44,* 677–694. doi:10.2307/1414531
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., 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
Johansson, R., Holsanova, J., & Holmqvist, K. (2010). Eye movements during mental imagery are not reenactments of perception. In: Ohlsson, Stellan & Catrambone, Richard (Eds.), *Proceedings of the 32nd annual meeting of the Cognitive Science Society* (pp. 1968–1973). Austin, TX: Cognitive Science Society.
Johansson, R., Holsanova, J., & Holmqvist, K. (2011). The dispersion of eye movements during visual imagery is related to individual differences in spatial imagery ability. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd annual meeting of the Cognitive Science Society* (pp. 1200–1205). Austin, TX: Cognitive Science Society.
Keehner, M., Hegarty, M., Cohen, C. A., Khooshabeh, P., & Montello, D. R. (2008). Spatial reasoning with external visualizations: What matters is what you see, not whether you interact. *Cognitive Science: A Multidisciplinary Journal, 32,* 1099–1132. doi:10.1080/03640210801898177
Kent, C., & Lamberts, K. (2008). The encoding-retrieval relationship: Retrieval as mental simulation. *Trends in Cognitive Sciences, 12,* 92–98. doi:10.1016/j.tics.2007.12.004
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
Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). *The case for mental imagery.* New York, NY: Oxford University Press. doi:10.1093/acprof:oso/9780195179088.001.0001
Kosslyn, S. M. (1994). *Image and brain.* Cambridge, MA: MIT Press.
Kozhevnikov, M., Kosslyn, S., & Shephard, J. (2005). Spatial versus object visualizers: A new characterization of visual cognitive style. *Memory & Cognition, 33,* 710–726. doi:10.3758/BF03195337
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
Marks, D. F. (1973). Visual imagery differences and eye movements in the recall of pictures. *Perception & Psychophysics, 14,* 407–412. doi:10.3758/BF03211175
Micic, D., Ehrlichman, H., & Chen, R. (2010). Why do we move our eyes while trying to remember? The relationship between non-visual gaze patterns and memory. *Brain and Cognition, 74,* 210–224. doi:10.1016/j.bandc.2010.07.014
Neisser, U. (1967). *Cognitive psychology.* New York: Appleton-Century-Crofts.
Nordqvist, S. (1990*). Kackel i trädgårdslandet* [Festus and mercury: Ruckus in the Garden]. Bromma, Sweden: Opal.
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
O'Regan, J. K. (1992). Solving the 'real' mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology/Revue Canadienne de psychologie, 46,* 461–488. doi:10.1037/h0084327
Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology/Revue Canadienne de psychologie, 45,* 255–287. doi:10.1037/h0084295
Perky, C. W. (1910). An experimental study of imagination. *The American Journal of Psychology, 21,* 422–452. doi:10.2307/1413350
Pollatsek, A., Rayner, K., & Henderson, J. M. (1990). Role of spatial location in integration of pictorial information. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 199–210. doi:10.1037/0096-1523.16.1.199
Postle, B. R., Idzikowski, C., Della Salla, S., Logie, R. H., & Baddeley, A. D. (2006). The selective disruption of spatial working memory by eye movements. *Quarterly Journal of Experimental Psychology. 59,* 100–120. doi:10.1080/17470210500151410
Pylyshyn, Z. (1981). The imagery debate: Analogue media versus tacit knowledge. *Psychological Review, 88,* 16–45. doi:10.1037/0033-295X.88.1.16
Pylyshyn, Z. W. (2002). Mental imagery: In search of a theory. *Behavioral and Brain Sciences, 25,* 157–182. doi:10.1017/S0140525X02000043
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
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.
Slotnick, S. D., Thompson, W. L., & Kosslyn, S. M. (2005). Visual mental imagery induces retinotopically organized activation of early visual areas. *Cerebral Cortex, 15,* 1570–1583. doi:10.1093/cercor/bhi035
Spivey, M. J., & Geng, J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research/Psychologische Forschung, 65,* 235–241. doi:10.1007/s004260100059
Spivey, M., Tyler, M., Richardson, D., & Young, E. (2000). Eye movements during comprehension of spoken scene descriptions. *Proceedings of the 22nd annual conference of the Cognitive Science Society* (pp. 487–492). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
Stoy, E. G. (1930). A preliminary study of ocular attitudes in thinking of spatial relations. *Journal of General Psychology, 4,* 379–385. doi:10.1080/00221309.1930.9918326
Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*(5217), 1632–1634. doi:10.1126/science.7777863
Theeuwes, J., Belopolsky, A., & Olivers, C. N. L. (2009). Interactions between working memory, attention and eye movements. *Acta Psychologica, 132,* 106–114. doi:10.1016/j.actpsy.2009.01.005
Theeuwes, J., Olivers, C. N. L., & Chizk, C. L. (2005). Remembering a location makes the eyes curve away. *Psychological Science, 16,* 196–199. doi:10.1111/j.0956-7976.2005.00803.x
Thomas, L. E., & Lleras, A. (2009). Covert shifts of attention function as an implicit aid to insight. *Cognition, 111,* 168–174. doi:10.1016/j.cognition.2009.01.005
Thomas, N. J. T. (2009). Visual imagery and consciousness. In W. P. Banks (Ed.), *Encyclopedia of consciousness* (volume 2, pp. 445–457).

Oxford, UK: Elsevier/Academic Press. doi:10.1016/B978-012373873-8.00083-9

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

van den Berg, R., Cornelissen, F. W., & Roerdink, J. B. (2009). A crowding model of visual clutter, *Journal of Vision, 9,* 24, 1–11.

Weiner, S. L., & Ehrlichman, H. (1976). Ocular motility and cognitive process. *Cognition, 4,* 31–43. doi:10.1016/0010-0277(76)90009-3

Wilcoxon, F. (1945). Individual comparison by ranking methods. *Biometrics, 1,* 80–83. doi:10.2307/3001968

Wooding, D. S. (2002). Fixation maps: Quantifying eye-movement traces. *Proceedings of the 2002 symposium on Eye tracking research & applications,* New Orleans, Louisiana.

Zangemeister, W. H., & Liman, T. (2007). Foveal versus parafoveal scanpaths of visual imagery in virtual hemianopic subjects. *Computers in Biology and Medicine, 37,* 975–982. doi:10.1016/j.compbiomed.2007.01.015

Appendix A

Correspondence Between Eye-Movement and Verbal Data

Most studies of eye movements during visual imagery have investigated how well positional and sequential eye movement data during scene recollection (looking at a blank screen) corresponds with positional and sequential eye movement data from a preceding encoding phase of visual scene perception (e.g., Brandt & Stark, 1997; Gbadamosi & Zangemeister, 2001; Laeng & Teodorescu, 2002; Spivey & Geng, 2001). However, as it has been found that eye movement patterns during visual imagery are very idiosyncratic and frequently compressed to a smaller area (e.g., Johansson et al., 2011), we believe it is unsound to ascribe actual physical coordinates of objects from the encoded stimulus onto the blank screen used during recall and to use these coordinates to analyze how well eye movements during a scene recollection resembles the original scene. Therefore, we instead developed a method where eye movement data is recorded concurrent with verbal data and where the analysis of how well eye movements resemble the recalled scene is dependent on how the eye movement data corresponds to the verbal data, and the overall structure of each participants individual gaze pattern (Johansson et al., 2006).

Figure A1 shows the entire gaze pattern for one participant after she has orally described a scene from memory while looking at a blank screen. Schematics of this scene are illustrated to the right of the gaze pattern. Figure A2 shows eye movement data and verbal data over four successive points in time for this particular participant.

Criteria for Eye Movement Correspondence

To evaluate whether an eye movement corresponds with the direction or location of an object mentioned in the verbal data, this method considers specific temporal and spatial criteria.

Temporal Criteria

When describing a scene, eye movements frequently *precede* the mentioning of an object. But eye movements may also occur *after* an object is mentioned. That is, sometimes participants first move their eyes to a new position and then start talking about that object, while others start talking about an object and then move their eyes to the new location. In Johansson et al. (2006), both the maximum voice-eye latency (eye movements *after* an object was mentioned) and the maximum eye-voice latency (eye movements *before* an object was

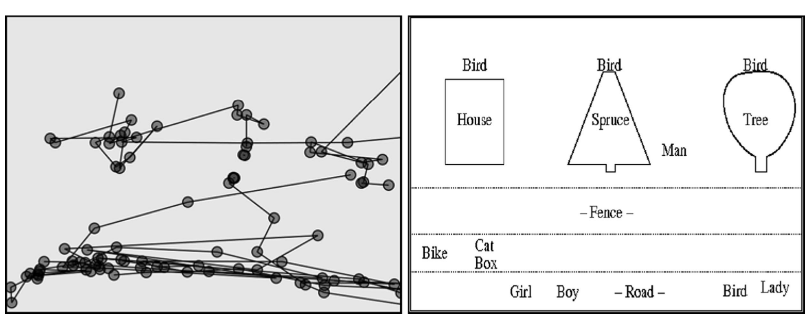

Figure A1. The entire gaze pattern for one participant (left) after she has looked at a blank screen and orally described a scene from memory. A schematics of the scene is illustrated to the right. Circles represents fixations, and lines represents saccades. From “Pictures and spoken descriptions elicit similar eye movements during mental imagery, both in light and in complete darkness,” by R. Johansson, J. Holsanova, and K. Holmqvist, 2006, *Cognitive Science, 30,* p. 1057. Copyright 2006 by Cognitive Science Society, Inc. Adapted with permission.

(*Appendix continues*)

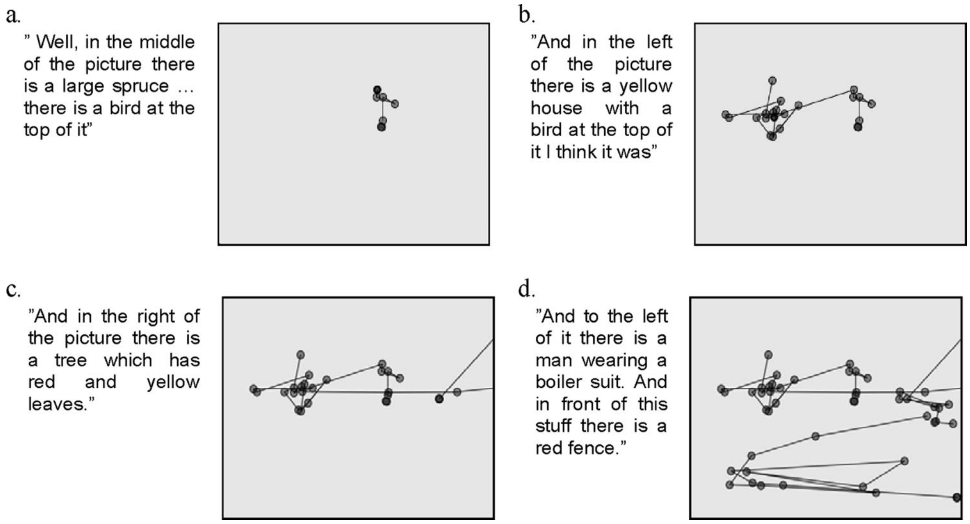

Figure A2. Gaze patterns and verbal data (here translated from Swedish into English) of the first 57 sec. over four succesive points in time for the participant in Figure A1. (a) 0 to 10 sec. (b) 11 to 26 sec. (c) 26 to 40 sec. (d) 40 to 57 sec.

mentioned) were found to be 5 seconds when all participants were considered (with average values between 0 to 2 seconds). Therefore, the temporal criterion is defined as follows:

> The eye movement, from one position to another, must appear within 5 seconds before or after the participant mentions the object.

Spatial Criteria

Apart from the temporal criteria in relation to the verbal it is also necessary to define a minimum threshold for the saccadic amplitude (the distance of the eye movement between two consecutive fixations) to be considered an actual movement from one imagined object to another. Because eye movements during visual imagery are very idiosyncratic, this threshold is not obvious to set. However, in line with Zangemeister and Liman (2007), the threshold was set to 1.1° (this represents about 10 mm on the blank computer screen). All saccades below this threshold were not considered to be an actual movement to a new area in the "mental image" of the recalled scene (see Figure A3).

Moreover, as it is possible to imagine a scene either using the whole blank screen or only a certain part of it, we did not ascribe imagined objects to specific coordinates on the blank screen. Instead we analyze the relative position of an eye movement compared with the overall structure of each participants individual gaze pattern (see Figure A4).

Eye movements of the participants were then scored as corresponding with the verbal data or not corresponding with the verbal data according to two forms of spatial coding: **global correspondence** and **local correspondence**. Global correspondence was considered when fulfilling the following spatial criterion (together with the temporal criterion above):

Figure A3. Example of one fixation that was not considered an actual movement (left) and one that was (right). This is the first three fixations from the gaze pattern illustrated in Figure A2 where this particular participant describes a spruce in the center and with a bird in the top of it.

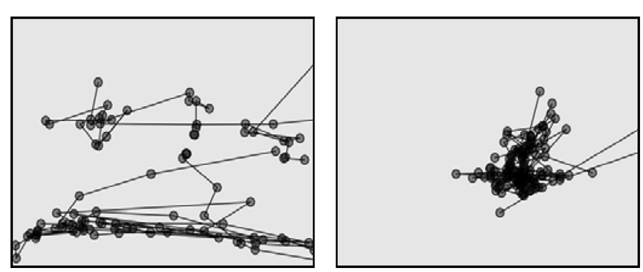

Figure A4. Examples of two participants after they had looked at the blank screen and described the scene illustrated in Figure A1. To the left, one participant who looked at the entire screen during the scene recollection (the same participant as in Figures A1–A3), and to the right, another participant who looked at a much smaller area of the screen.

(*Appendix continues*)

> When an eye movement shifts from one object to another it must finish in a *position* that is spatially correct relative to all fixation points over the participants' entire verbal report.

Local correspondence was considered when fulfilling the following spatial criterion (together with the temporal criterion above):

> When an eye movement shifts from one object to another it must move in the correct *direction***.**

> The possible directions were up, down, left, or right, which gives four possible quadrants of 90° for the eyes to move in (Figure A5).

The key difference between global and local correspondence is that global correspondence requires fixations to take place at the categorically correct spatial position relative to the whole individual eye-tracking pattern. Local correspondence only requires that the eyes move in the correct direction between two consecutive objects in the verbal data. See Figure A6 for schematics and eye movement examples of how local and global correspondence differs. If the temporal or spatial criteria for neither local correspondence nor global correspondence are fulfilled, eye movements are not considered to correspond with the verbal data (typically, when the eyes move with an amplitude below the threshold of 1.1° or move in the wrong direction).

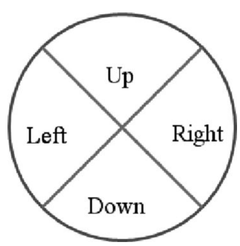

Figure A5. The four directional quadrants.

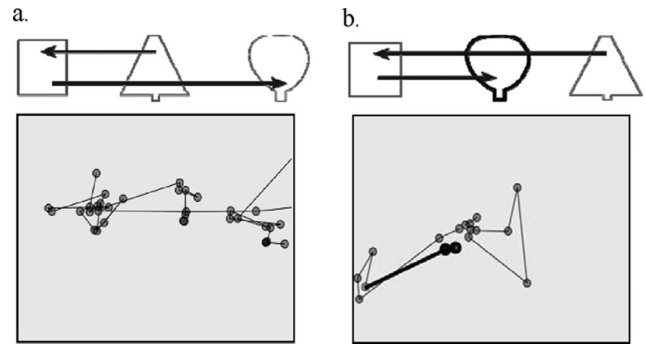

Figure A6. To the left (a), one participant (the same as in Figures A1 – A3) who first describes the spruce in the center, then the house to left and the tree to the right of both the house and the spruce. This participant executes eye movements that correspond to this spatial layout according to global correspondence criteria, that is, first she looks in the center and describes the spruce, then she looks at an area to the left and describes the house, and finally when she describes the tree she looks at an area that is to the right of both the area she was looking at when she described the spruce and the area she was looking at when she described the house. To the right (b), another participant who describes the exact same three objects. She also describes the spruce in the center, then moves her eyes to an area to the left and describes the house. However, when this participant describes the tree, her eyes (black and bold) do not end up in a position that is correct in relation to both the spruce and the house. Thus, this participant was not scored as having global correspondence when she described the tree. Nevertheless, she did move her eyes in the correct direction (to the right) when she shifted her attention from the house and started to describe the tree, and thus this participant was scored as having local correspondence when she described the tree. From "Pictures and Spoken Descriptions Elicit Similar Eye Movements During Mental Imagery, Both in Light and in Complete Darkness," by R. Johansson, J. Holsanova, and K. Holmqvist, 2006, *Cognitive Science, 30,* p. 1059. Copyright 2006 by Cognitive Science Society, Inc.

Received June 2, 2011
Revision received October 24, 2011
Accepted October 28, 2011 ■