Title
Using Eye Movements to Study Working Memory Rehearsal

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 31(31)

ISSN
1069-7977

Authors
Loschky, Lester
Zelinsky, Gregory

Publication Date
2009

Peer reviewed

Using Eye Movements to Study Working Memory Rehearsal for Objects in Visual Scenes

Gregory J. Zelinsky (Gregory.Zelinsky@Stonybrook.Edu)
Department of Psychology, Stony Brook University
Stony Brook, NY 11790-2500 USA

Lester C. Loschky (Loschky@KSU.Edu)
Department of Psychology, Kansas State University
Manhattan, KS 66502 USA

Abstract

Observers freely viewing a multi-object scene in preparation for a memory test often shift their gaze back to previously inspected objects. These refixations were analyzed to determine if they served a working memory rehearsal function. A gaze-contingent display paradigm was used to limit the observers' viewing of a study scene, which was followed by a spatial probe at one of the object locations and then by a 4AFC test for the probed target. The results indicated a 16% accuracy benefit linked to target refixation that disappeared if 6 or more objects were fixated after the target during study. We interpret these findings as evidence for a monitor-refixate system and a moving-anchor rehearsal strategy. According to this view, observers continuously monitor the availability of items in working memory, then refixate an item when its availability drops below a task-specific criterion. Items are shuffled in and out of this rehearsal set as the rehearsal anchor is systematically moved during scene inspection.

Keywords: memory maintenance; scene perception; gaze-contingent; rehearsal strategies; free viewing

Introduction

People freely viewing a display often return their gaze to previously fixated objects. This behavior is called an oculomotor *refixation*. Refixations are a ubiquitous property of oculomotor behavior, and have been noted in tasks as diverse as reading (Rayner, 1978, 1998), pattern copying (Ballard, Hayhoe, & Pelz, 1995; Hayhoe, Bensinger, & Ballard, 1998), portrait painting (Locher, 1996; Nodine, Locher, & Krupinski, 1993), solving arithmetic and geometry problems (Epelboim & Suppes, 1996; Hegarty, Mayer, & Green, 1992), visual search (Gilchrist & Harvey, 2000; Peterson, Kramer, Wang, Irwin, & McCarley, 2001), and undirected picture viewing (Mannan, Ruddock, & Wooding, 1996, 1997). Refixations can also comprise a significant portion of the oculomotor behavior accompanying a task, in one case up to 25% of the observed fixations (Mannan et al., 1997).

Why do people choose to refixate objects that were previously inspected? The answer to this question likely depends on the task. In the case of reading, gaze is often returned to previously read portions of a text in order to resolve specific lexical ambiguities (Frazier & Rayner, 1982; Murry & Kennedy, 1988) or to obtain elaborative details needed for narrative comprehension (Blanchard & Iran-Nejad, 1987; Just & Carpenter, 1976; Shebilske & Fisher, 1983). Refixations, however, certainly play very different roles in non-reading tasks. For example, Hayhoe, Ballard, and colleagues (Ballard et al., 1995; Hayhoe et al., 1998) had observers reconstruct a spatially complex multi-colored block pattern (the model) from a set of individual colored blocks scattered in a workspace. These observers looked first to the model pattern in order to determine the next block to select from the workspace, then refixated the model to determine the exact location to put the selected block in the pattern under construction. Refixation in this case was therefore used to acquire a specific piece of information, the spatial information needed to correctly position a block in the construction of the model replica.

Our study explores refixations in the context of an explicit working memory task. Unlike the refixations made during reading or block copying, we propose that refixations in working memory tasks may serve a straightforward rehearsal function. Many studies have shown that memory for specific stimulus properties (e.g., exact color or surface form) declines soon after gaze shifts away from an object (Carlson-Radvansky, 1999; Carlson-Radvansky & Irwin, 1995; Henderson, 1997; Henderson & Siefert, 2001; Irwin, 1991, 1996; Irwin & Gordon, 1998). Given this rapid loss of object properties following gaze movement, observers freely viewing a multi-object scene may attempt to offset this information loss by visually reacquiring an object and encoding its properties anew. Observers may therefore employ a systematic strategy of cyclically refixating or rehearsing display objects while viewing a study scene, with each refixation on an object occurring when its availability in memory drops below a critical threshold. To assess this rehearsal process, we use an explicit working memory task in which the observers' eye movements are monitored as they study an array of objects in a simple scene. We interpret refixations in this task as an attempt to actively refresh an object's availability in working memory, and refer to the specific pattern in which these refixations are deployed as the observer's memory rehearsal strategy.

The current study addresses two questions regarding the relationship between refixations and rehearsal in an explicit memory task. First, can the refixation behavior observed during the viewing of a study scene be serving a rehearsal function? A rehearsal strategy cannot be inferred from the mere existence of refixations, and in fact answers to the following questions would need to be obtained before this relationship can be seriously entertained. First, how common were refixations in the task? If this behavior was relatively rare, then even if it was used as a rehearsal tool it may not meaningfully affect memory performance. Second, are these refixations patterned or simply the product of random oculomotor inspection? We will address this question by comparing the actual refixation rate to a chance baseline, and by looking for evidence of systematic patterns in the refixation data. Third, does the refixation of an object improve recognition accuracy for that object later during test? If this relationship is not observed, then refixations obviously cannot be serving a rehearsal function.

Methods

The events comprising a typical trial are shown in Figure 1. A trial began with the presentation of a 9-object study scene (Figure 1a). The stimuli were real-world objects (toy, tool, or food items) arranged on an appropriate background surface (a crib, workbench, or dining table). Each scene was presented in color and subtended 18° × 11.6° of visual angle. Individual objects were scaled to fit within a 2.4° bounding box and their locations in the scene were constrained to 18 positions, creating the appearance of a random arrangement of items on a surface. Multiple trials for a given scene type were created by randomly pairing objects to locations.

Observers studied scenes in preparation for a recognition test. At test a spatial probe would be presented and they would have to indicate the object from the study scene that appeared at a probed location. Observers were therefore required to encode from the study scene information about both the identity and location of each object.

Distinguishing our paradigm from many other working memory paradigms is the fact that testing was contingent upon the observer's oculomotor behavior while viewing the study scene. Although the experiment instructions made no reference to oculomotor behavior, observers invariably chose to make eye movements during this challenging memory task. Eye position was sampled at 1000 Hz using a Generation VI dual Purkinje-image eye tracker and analyzed in real time to determine the object being fixated. Unknown to the observer, one of the objects was pre-designated as the memory target for that particular trial (e.g., the butter dish in Figure 1). As the observer freely viewed the study scene, gaze would eventually be directed to this target object (Figure 1a). This fixation event was detected by the program controlling the stimulus presentation, which then started to count the number of different objects fixated after gaze left the target. We refer to these items fixated after the target as intervening objects (Waugh and Norman, 1965), in recognition of their potential for interfering with memory for the target. Because these objects would be processed after the target but before the memory test, memory for the target may vary as a function of the number of fixated intervening objects (Zelinsky & Loschky, 2005).

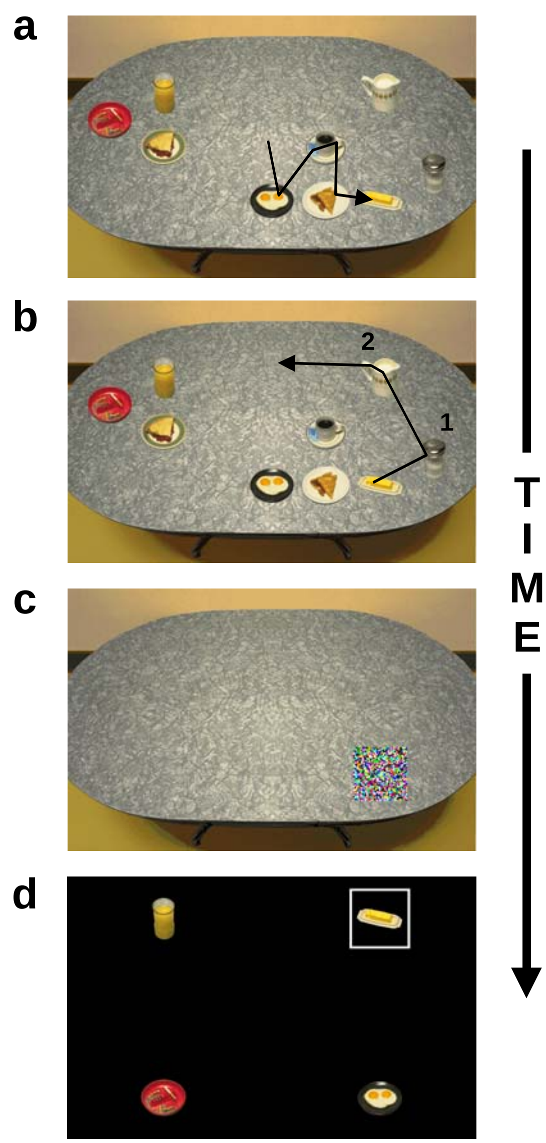

Figure 1: Procedure and representative stimuli.

A criterion placed on the number of intervening objects was used to terminate the study display. If this intervening object criterion was preset to 2, the observer would be allowed to fixate exactly 2 objects following the target (Figure 1b). As gaze shifted away from the second post-target object, the study scene would be replaced by a colored noise mask appearing at the target's location on an "emptied" background surface (Figure 1c). This mask, which was visible for 1 second, served as the spatial probe for the memory test. There were 7 intervening object conditions (1-7). In the 1 intervening object condition, the study display was terminated during the saccade away from the first post-target object (i.e., the eye was not allowed to land on a second non-target item after leaving the target); in the 7 intervening object condition the observer was allowed to fixate exactly 7 non-target objects following target fixation. The intervening object conditions were randomly interleaved throughout the experiment. Post-experiment questioning revealed that none of the six participants realized that their opportunity for study depended on their own pattern of eye movements, in conjunction with an intervening object criterion, instead attributing the variable study scene duration to some random presentation schedule.

Following probe offset, a 4-object array was presented and observers had to indicate which one of these objects appeared at the probed location (Figure 1d). One of these objects was always the target, with the other three randomly selected from the study scene. The observer registered a response by looking at the desired object, which caused a white box to be drawn around the item, then pressing a button when satisfied with their selection. Observers were instructed to respond as accurately as possible without regard for time, and each participated in 378 trials, 54 per each of the 7 intervening object conditions. See Zelinsky and Loschky (2005) for additional details.

Results and Discussion

How frequent are target refixations?

One consequence of a free viewing paradigm is that observers could shift their gaze back to the target object before the scene-terminating intervening object criterion was achieved. To determine if this refixation behavior was common enough to serve a rehearsal function, we isolated and analyzed those trials in which observers made at least one target refixation and found an overall refixation rate of 35%. However, this refixation rate varied with intervening object condition. Figure 2 shows that the frequency of target refixation trials increased with the number of fixated intervening objects required by the gaze contingent paradigm. When observers were allowed to fixate only one object following the target, a target refixation could not occur because the study scene would terminate upon gaze leaving the non-target object. This contingency is indicated in the data by a refixation frequency of zero in the 1 intervening object condition. However, as soon as refixations were allowed by the paradigm, the refixation rate jumped to 20% and continued to increase with the number of intervening objects. This monotonic increase in refixation rate is not surprising. Each additional fixation forced by our intervening object manipulation creates another opportunity for observers to direct gaze back to the target. What is surprising is the fact that refixations occurred on up to 53% of the trials, making this behavior frequent enough to be used as a rehearsal strategy for maintaining object information in working memory.

Is this refixation rate greater than chance?

The analysis from the previous section only demonstrated the feasibility of using refixations as a rehearsal tool; it offered no direct evidence that refixations were actually serving this function. A first step in drawing a link between refixations and rehearsal is to show that refixation behavior is patterned and not simply random. The refixation function illustrated in Figure 2 by itself cannot demonstrate this; with sufficient time and eye movements, observers would inevitably refixate the target even if they were shifting their gaze randomly from one object to the next.

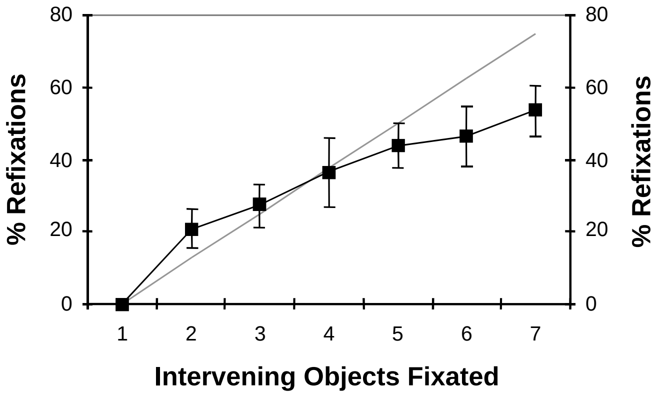

Figure 2: Frequency of target refixations as a function of intervening objects fixated after the target during study. Bars extending from each marker show the standard error of the mean. The line indicates the random target refixation frequency as estimated by Monte Carlo simulation.

To exclude the possibility that these refixations were due to random oculomotor inspection, we conducted Monte Carlo simulations of the refixation rate generated by random eye movements between 9 objects. Separate simulations were performed for all seven of the intervening object conditions, and each simulation implemented the identical display termination criteria used by the gaze-contingent methodology in the behavioral experiment. The gray line in Figure 2 shows the results from these 105,000 simulations; 15,000 per each of the intervening object conditions. If the observers' were refixating the target as a result of random oculomotor inspection, the behavioral refixation function should not differ from this random probability baseline. However, this was clearly not the case. Observers in the 2 intervening object condition refixated the target more frequently than would be expected by chance. This refixation rate dropped down to the random probability level in the 3 and 4 intervening object conditions, then continued to drop far below chance when there were 5 or more intervening objects. With regard to the question motivating this analysis, these deviations in the refixation rate from chance clearly indicate a patterned behavior. Observers preferentially revisited the target when the paradigm allowed fixations on only 2 intervening objects, but avoided returning to the target when the paradigm forced a larger number of intervening object fixations.

When are targets refixated during study?

We now know that refixations were common in this explicit memory task and that they were not just the product of random eye movements between objects, but we do not yet have a clear picture of when these refixations actually occurred during viewing of the study scene. Figure 3 replots the refixation data from Figure 2, segregated by four intervening object conditions. Each panel is a relative frequency histogram of refixations showing when in the intervening object sequence observers elected to shift gaze back to the target. This analysis shows that target refixations within a given condition did not occur equally often after every allowable intervening object fixation. Although observers had multiple opportunities to return gaze to the target in the 3-6 intervening object conditions, we see that the modal refixation behavior occurred after gaze left the first post-target object. Observers were in fact almost twice as likely to refixate the target after 1 intervening object than at any other time during the study scene presentation. If an observer was going to refixate the target, that refixation was most likely to occur soon after initially fixating the target (see also Dickinson & Zelinsky, 2007).

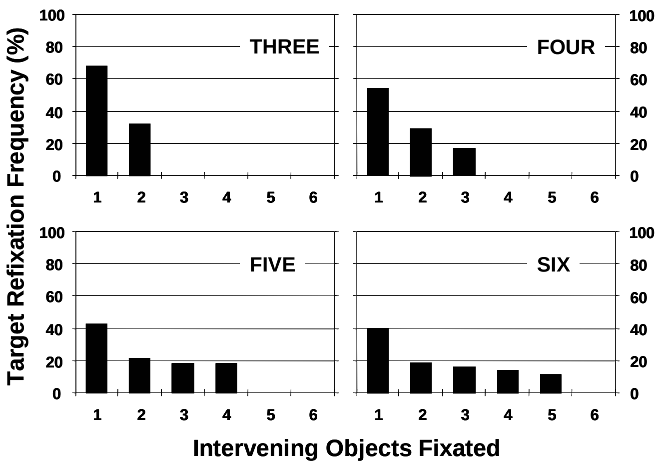

Figure 3: Refixation histograms for the three, four, five, and six intervening object conditions. Each shows when gaze shifted back to the target as a function of fixated intervening objects during scene viewing. A similar pattern characterized the seven intervening object data (not shown).

Do refixations improve memory?

A final property required by any behavior believed to serve a memory rehearsal function is, of course, a demonstrated relationship to an actual memory benefit. Figure 4 compares recognition accuracy for trials in which there was a target refixation to trials in which there were no target refixations.[1] Refixating the target in the 2-5 intervening object conditions improved our observers' ability to later pick this object out of a 4AFC display, one-tailed t (5) = -2.91, p = .017, effect size (Eta^2) = .628. This improvement amounted to a roughly constant 16% benefit over the 2-5 intervening object range, followed by a decline into the no-refixation level of performance with additional intervening objects. An essential component of the asserted relationship between refixation and a memory rehearsal function is therefore supported. To the extent that refixations serve a rehearsal function, they would be expected to improve memory for the refixated object. That appears to be the case. Returning gaze to an object during the study phase of an explicit memory task increased the probability of that object being correctly identified in a subsequent recognition test.

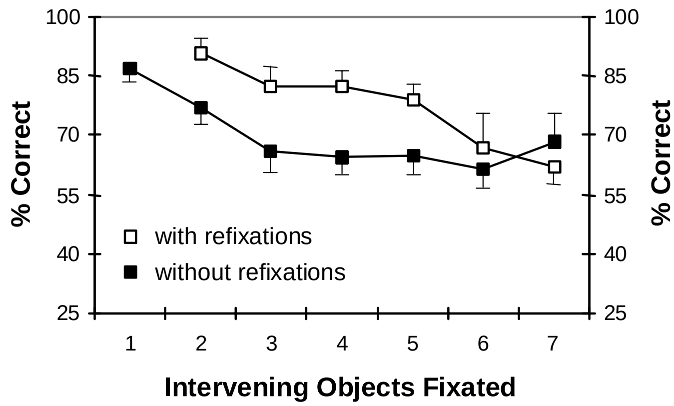

Figure 4: Recognition accuracy as a function of intervening objects fixated after the target. Filled markers show data from trials in which the target was not refixated during study; the unfilled markers show data from trials in which there was a target refixation. Bars indicate one SEM.

General Discussion

Memory rehearsal strategies have long been implicated in tasks involving the sequential presentation of list items (Murdock & Metcalfe, 1978; Rundus & Atkinson, 1970; see Baddeley, 1986, and Neath, 1998, for reviews). The current study builds on this work by proposing the use of an analogous rehearsal strategy for objects presented simultaneously as part of a visual scene.

The need for such a rehearsal strategy arises from the observation of serial order effects introduced by the sequence of fixations made to objects during scene viewing (Zelinsky & Loschky, 2005). The availability of an object in memory depends on when it was fixated during study, with more recently fixated objects enjoying a higher probability of memory retrieval. This is clear in the case of the Figure 4 no-refixation data; accuracy was 87% after fixating only 1 intervening object, 76% after 2, and 65% after 3 intervening object fixations. Such a pattern suggests that the target was highly available for retrieval immediately after fixation, but that interference introduced by the following object fixations caused this initially high level of availability to rapidly decline. Eventually, after fixating three intervening objects, the recency effect disappeared entirely, asymptoting into an above-chance pre-recency level of performance.

To attenuate this rapid decline in immediate memory for objects, we propose that gaze refixations are used in an attempt to actively maintain or rehearse object information, which we refer to as a *monitor-refixate rehearsal strategy*.

[1] A detailed discussion of the no-refixation data, with emphasis on the shape of the intervening object function and its theoretical implications, can be found in Zelinsky and Loschky (2005).

According to this strategy, observers viewing a study scene continuously monitor the availability of each object encoded into memory, then use refixations to maintain this object at the level of availability needed to achieve a desired level of recognition performance. Refixations are triggered by a threshold that the observer sets on this availability dimension. When the availability of a previously fixated object drops below this criterion, the oculomotor system is signaled that this item is in danger of being forgotten and an eye movement is programmed back to the item in an attempt to maintain its place in working memory. The goal of an object refixation, whether explicit or otherwise, is therefore to offset the interference introduced by fixations on other objects (Zelinsky & Loschky, 2005) by reinstating its previously above-criterion level of memory availability.

This simple monitor-refixate model can account for many of the patterns observed in the current data. First, it explains the improved accuracy resulting from refixations reported in Figure 4, and the disappearance of this advantage with a large number of intervening objects. Such benefits derive from the fact that each target refixation essentially resets the intervening object function. If a target is refixated during study, the non-target object inspected immediately following this refixation would represent only one intervening object, regardless of when the target was initially fixated during viewing. Accuracy on 2 intervening object refixation trials (90%) therefore resembles the level found on 1 intervening object no-refixation trials (86%). A similar logic can be applied to the larger intervening object conditions. Because refixations in the 3 intervening object condition would occur after fixations on either the first or second post-target objects, accuracy in this condition should fall between the 87% and 76% levels delineated by the no-refixation function, and indeed this was the case (82%). Note, however, that this relationship would be expected to break down at very large intervening object separations. Consider a 6 intervening object trial. Although a target refixation could have occurred following 2-5 intervening objects in this condition, as Figure 3 reminds us these refixations were not uniformly distributed over this range. On the majority of these trials the target would have been refixated after the first two intervening objects in the viewing sequence. These early refixations would mean that any recency benefit would have disappeared into the pre-recency asymptote by the time the 6 intervening object criterion was satisfied. This relationship between recency and intervening objects therefore nicely explains the gradual convergence of the refixation and no-refixation functions after 6-7 intervening objects were fixated.

A monitor-refixate model also explains why the refixation rate was above-chance in the 2 intervening object condition (Figure 2). As speculated by Hooge and Erkelens (1999), quite often gaze leaves an object before the properties of that object have been fully coded. This incomplete encoding, when coupled with the interference created by shifting gaze to an intervening object, would cause the target's availability in memory to drop below criterion. This event would in turn trigger an eye movement bringing gaze immediately back to the target, which over trials would elevate the target refixation rate above chance.

As for why the target refixation rate dipped below chance over the 5-7 intervening object range (Figure 2), we believe that this reflects the abandonment of objects fixated early in the viewing sequence from the rehearsal process. Assuming a capacity limit on the size of the rehearsal set, strict adherence to the above-described maintenance rehearsal strategy would quickly lock the observer in a perpetual loop, with the same subset of objects continuously refixated and forgotten. In order to carry out the instructed task of fully inspecting the study scene, observers would therefore have to periodically break this monitor-refixate loop long enough to include new items in the rehearsal set. However, a fixed capacity rehearsal set means that the addition of each new object to the set must be accompanied by the abandonment of a previously maintained member. In the sequence of object fixations A, then B, then C, an observer might therefore elect to override the signal from A asking for a refixation, and make an eye movement instead from C to D. Following this fixation on D, a refixation would then be made to B rather than A, despite the fact that the availability of both objects would now be below threshold. This refixation amounts to a decision on the part of the observer to abandon maintenance of object A and to shift their starting point or "anchor" in the rehearsal loop to object B.

By systematically moving their anchor as they study a scene, observers would dynamically change the composition of the rehearsal set throughout a trial. Note however that refixations are not increasing working memory capacity under this scheme, the membership of the rehearsal set is simply being changed as items are shuffled in and out of memory. Each time an observer wishes to add a new item to the rehearsal set, an old item must be abandoned. In this sense, working memory in this task can be characterized as a sort of moving window of highly available scene objects, with the size of this window determining the number of objects that can be maintained above criterion, and the location of this window in the scene determining the contents of the rehearsal set.

Conclusion

We propose that the refixations occurring during the study of a scene reflect the use of a monitor-refixate memory rehearsal strategy. This strategy assumes the moment-by-moment monitoring of working memory representations for retrieval availability. When one of these object representations drops below an availability threshold, the eye movement system is signaled to return gaze to that object so as to refresh its representation. Just as inner speech is used to maintain a list of digits long enough to dial them into a telephone, refixations may therefore serve to maintain the visuo-spatial information from simultaneously presented objects long enough to perform whatever task is at hand.

Acknowledgments

G.J.Z. was supported by NIMH grant 2-R01-MH63748 and NSF grant IIS-0527585. L.C.L. was supported by funds from the Kansas State University Office of Research and Sponsored Programs, and by the NASA Kansas Space Grant Consortium.

References

Baddeley, A. (1986). *Working memory*. (Vol. 11). Oxford: Clarendon Press.

Ballard, D., Hayhoe, M., & Pelz, J. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience*, *7*, 66-80.

Blanchard, H. E., & Iran-Nejad, A. (1987). Comprehension processes and eye movement patterns in the reading of surprise-ending stories. *Discourse Processes*, *10*, 127-138.

Carlson-Radvansky, L. A. (1999). Memory for relational information across eye movements. *Perception & Psychophysics*, *61*, 919-934.

Carlson-Radvansky, L. A., & Irwin, D. E. (1995). Memory for structural information across eye movements. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*, 1441-1458.

Dickinson, C. A., & Zelinsky, G. J. (2007). Memory for the search path: Evidence for a high-capacity representation of search history. *Vision Research*, *47*, 1745-1755.

Epelboim, J., & Suppes, P. (1996). Window on the mind? What eye movements reveal about geometric reasoning. In G. Cottrell (Ed.). *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society* (pp. 59-60). Mahwah, New Jersey: Erlbaum Associates.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, *14*, 178-210.

Gilchrist, I. D., & Harvey, M. (2000). Refixation frequency and memory mechanisms in visual search. *Current Biology*, *10*, 1209-1212.

Hayhoe, M., Bensinger, D., & Ballard, D. (1998). Task constraints in visual working memory. *Vision Research, 38*, 125-137.

Hegarty, M., Mayer, R. E., & Green, C. E. (1992). Comprehension of arithmetic word problems: Evidence from students' eye fixations. *Journal of Educational Psychology*, *84*, 76-84.

Henderson, J. M. (1997). Transsaccadic memory and integration during real-world object perception. *Psychological Science*, *8*, 51-55.

Henderson, J. M., & Siefert, A. B. (2001). Types and tokens in transsaccadic object identification: Effects of spatial position and left-right orientation. *Psychonomic Bulletin & Review*, *8*, 753-760.

Hooge, I., & Erkelens, C. (1999). Peripheral vision and oculomotor control during visual search. *Vision Research*, *39*, 1567-1575.

Irwin, D. E. (1991). Information integration across saccadic eye movements. *Cognitive Psychology*, *23*, 420-456.

Irwin, D. E. (1996). Integrating information across saccadic eye movements. *Current Directions in Psychological Science*, *5*, 94-99.

Irwin, D. E., & Gordon, R. D. (1998). Eye movements, attention and trans-saccadic memory. *Visual Cognition*, *5*, 127-155.

Just, M., Carpenter, P. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, *8*, 441-480.

Locher, P. J. (1996). The contribution of eye-movement research to an understanding of the nature of pictorial balance perception: A review of the literature. *Journal of the International Association of Empirical Studies*, *14*, 143-163.

Mannan, S., Ruddock, K., & Wooding, D. (1996). The relationship between the location of spatial features and those of fixations made during visual examination of briefly presented images. *Spatial Vision*, *10*, 165-188.

Mannan, S., Ruddock, K., & Wooding, D. (1997). Fixation sequences made during visual examination of briefly presented 2D images. *Spatial Vision*, *11*, 157-178.

Murdock, B. B., & Metcalfe, J. (1978). Controlled rehearsal in single-trial free recall. *Journal of Verbal Learning and Verbal Behavior*, *17*, 309-324.

Murry, W. S., & Kennedy, A. (1988). Spatial coding in the processing of anaphor by good and poor readers: Evidence from eye movement analyses. *Quarterly Journal of Experimental Psychology*, *40A*, 693-718.

Neath, I. (1998). *Human memory: An introduction to research, data, and theory*. New York: Brooks/Cole.

Nodine, C. F., Locher, P. J., & Krupinski, E. A. (1993). The role of formal art training on perception and aesthetic judgment of art compositions. *Leonardo*, *26*, 219-227.

Peterson, M. S., Kramer, A. F., Wang, R. F., Irwin, D. E., & McCarley, J. S. (2001). Visual search has memory. *Psychological Science*, *12*, 287–292.

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin*, *85*, 618-660.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372-422.

Rundus, D., & Atkinson, R. C. (1970). Rehearsal processes in free recall: A procedure for direct observation. *Journal of Verbal Learning and Verbal Behavior*, *9*, 99-105.

Shebilske, W. L., & Fisher, D. F. (1983). Eye movements and context effects during reading of extended discourse. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 153-179). New York: Academic Press.

Waugh, N. C., & Norman, D. A. (1965). Primary memory. *Psychological Review*, *72*, 89-104.

Zelinsky, G. J., & Loschky, L. C. (2005). Eye movements serialize memory for objects in scenes. *Perception & Psychophysics*, *67*, 676–690.