Eye movements during reading, visual search, and scene perception: An overview.

Monica S. Castelhano

Queen’s University

and

Keith Rayner

University of California, San Diego

and

University of Massachusetts, Amherst

Correspondence to:

Keith Rayner
Department of Psychology
University of Massachusetts
Amherst, MA 01003, USA
413-545-2175 (phone); 413-545-0996 (FAX)
rayner@psych.umass.edu

In this chapter, we provide an overview of research on eye movements during reading, visual search, and scene perception. Our overview of reading will be more complete than our overview of visual search or scene perception. The reason for this is quite simple. We know more about the nature of eye movements in reading than in the other two tasks. And, the reason for this is also quite apparent. In reading, there is a well-defined task for the viewer: people generally read to understand or comprehend the text. This involves a sequence of eye movements that typically move from left-to-right across the page and then down the page. Of course, the task can be varied somewhat so that, for example, readers are asked to skim the text, and this will result in different eye movements characteristics. Yet, the vast bulk of the research on eye movements during reading has utilized comprehension as the goal of the reader. On the other hand, in scene perception, the nature of the task is inherently much vaguer. Viewers may be asked to look at a scene to remember it, but the sequence in which they examine the scene may be highly idiosyncratic and variable. In visual search, there are many different types of search tasks (search for a letter, search for a colored object, search for a person in a large group picture, search for Waldo in a "Where's Waldo" children's book, and so on), and viewers can use idiosyncratic strategies in dealing with the task. Despite these differences, a fair amount of information on the nature of eye movements in each task is available. In this chapter, we will review some of the main findings concerning eye movements in these tasks as background information for the other chapters in the volume, which deal with these issues and also with the influence of cultural variables on eye movements.

Basic Characteristics of Eye Movements

While we are reading or searching a visual array for a target or simply looking at a new scene, our eyes move every 250-350 ms. These eye movements serve to move the fovea (the high resolution part of the retina encompassing 2 degrees at the center of the visual field) to an area of interest in order to process it in greater detail. Because of acuity limitations in the retina, eye movements are necessary for processing the details of the array. Our ability to discriminate fine detail drops off markedly outside of the fovea in the parafovea (extending out to about 5 degrees on either side of fixation) and in the periphery (everything beyond the parafovea). During the actual eye movement (or saccade), vision is suppressed[1] and new information is acquired only during the fixation (the period of time when the eyes remain still for about 250-350 ms). Although we have the impression that we can process the entire visual array in a single fixation and while we can rapidly obtain the gist of the scene from a single fixation, in reality we would be unable to fully process the information outside of foveal vision if we were unable to move our eyes (Rayner, 1978, 1998).

While it is true that we can move our attention independently of where the eyes are fixated, it does not seem to be the case in everyday viewing. The separation between attention and fixation is often attained in very simple tasks (Posner, 1980); however, in tasks like reading, visual search, and scene perception, covert attention and overt attention (the exact eye location) are tightly linked. Although when we look at a complicated scene, we can dissociate covert and overt attention, generally speaking it takes either a certain amount of effort to maintain it (as when we hold fixation and move our attention elsewhere) or it is a natural consequence of programming eye movements.

There is considerable evidence that before an eye movement is made, attention typically precedes it to the intended target (Deubel & Schneider, 1996; Hoffman & Subramaniam, 1995; Kowler, Anderson, Dosher, & Blaser, 1995; Rayner, McConkie, & Ehrlich 1978).

Because eye movements are essentially motor movements, it takes time to plan and execute a saccade. In simple reaction time experiments, where there is no necessity of cognitive processing of the fixated material and participants merely need to monitor when a simple fixation target moves from one location to another (and their eyes accordingly), it takes on the order of 175 ms to move the eyes under the best of circumstances (Becker & Jürgens, 1979; McPeek, Skavenski, & Nakayama, 2000; Rayner, Slowiaczek, Clifton, & Bertera, 1983). In addition, it is important to note that eye movements are more or less ballistic movements. Once initiated, it is difficult (though not impossible) to change their trajectory. While it has generally been assumed that the two eyes move in synchrony and that they fixate the same point in space, recent research summarized by Liversedge (this volume) clearly demonstrates that this is not the case and the two eyes are frequently deviated from each other.

Insert Table 1 about here

There is considerable evidence that the nature of the task influences eye movements. A summary of the average amount of time spent on each fixation and the average distance the eyes move in reading, visual search, and scene perceptionare shown in Table 1. From this table, it is immediately apparent that while the values presented in the table are quite representative of the different tasks, they show a range of average fixation durations and for each of the tasks there is considerable variability both in terms of fixation durations and saccade lengths. This is also clearly illustrated in Figure 1, which shows frequency distributions of fixation durations in the three tasks. From these graphs, it is very evident that there is a great amount of variability in fixation time measures.

Insert Figure 1 about here

At one time, researchers believed that the eyes and the mind were not tightly linked during information processing tasks like reading, visual search, and scene perception. This conclusion was based on the relatively long latencies of eye movements (or reaction time of the eyes) and the large variability in the fixation time measures. They questioned the influence of cognitive factors on fixations given that eye movement latency was so long and the fixation times were so variable. It seemed unlikely that cognitive factors could influence fixation times from fixation to fixation. Actually, an underlying assumption was that everything proceeded in a serial fashion and that cognitive processes could not influence anything very late in a fixation, if at all. However, a great deal of research since has established a tight link between the eye and the mind, and it is now clear that saccades can be programmed in parallel (Becker & Jürgens, 1979) and, furthermore, that information processing continues in parallel with saccade programming.

With this preamble (and basic information) out of the way, let’s now turn to a brief overview of eye movements in each of the three tasks. We’ll begin with reading (which will receive the most attention given that there is more research on eye movements in this task than the other two), and then move to visual search and scene perception.

Eye Movements in Reading

During reading, the average fixation duration in reading is about 225-250 ms and the average saccade size is 8-9 character spaces. In reading, unlike other tasks, character spaces are used rather than visual angle. This is because it has been demonstrated that character spaces are the more appropriate unit than visual angle. So, if the size of the print is held constant and the viewing distance varied (so that there are either more or fewer characters per degree of visual angle), how far the eyes travel is determined by character spaces, not visual angle (Morrison & Rayner, 1981). Another important characteristic of eye movements while reading is that about 10-15% of the time readers move their eyes (regress) back to previously read material in the text. These regressions, as they are called, tend to depend on the difficulty of the text. As would be expected, saccade size and fixation duration are also both modulated by text difficulty: as the text becomes more difficult, saccade size decreases, fixation durations increase, and regressions increase. From these measures alone, it is very clear that global properties of the text influence eye movements greatly. In addition, these three main global measures (saccade size, fixation duration and number of regressions) are also influenced by the type of material being read and the reader's goals in reading (Rayner & Pollatsek, 1989). For instance, reading a text for understanding produces a very different pattern of eye movement measures when compared to skimming a text while proofreading.

In addition to global effects, studies have shown clear local effects on words. Measures in these studies focus on the processing of a target word (versus looking at an average measure that is pooled from all words in a sentence, such as the average fixation duration). Local measures include: first fixation duration (the duration of the first fixation on a word), single fixation duration (those cases where only a single fixation is made on a word), and gaze duration (the sum of all fixations on a word prior to moving to another word). Global measures would be more useful if every word was fixated in a sentence and fixated only once. However, reading patterns are not so: many words are skipped during reading (i.e., never receive a direct eye fixation) and some words are fixated more than once. There is good reason to think that words that are skipped are processed on the fixation prior to the skip. There is also good reason to believe that when words are re-fixated (before moving on in the text), it is because more time is needed to process these words' meaning. The solution then is to use eye movement measures that are more local (as described above) and so, are able to provide a reasonable estimate of the processing a particular word (Rayner, 1998).

The Perceptual Span. A very important issue in reading is how much information is the reader able to process and use during a single fixation, which, as we've noted above, typically lasts for 200-250 ms. This measure is referred to as the perceptual span (also called the functional field of view or, to a lesser degree, the region of effective vision). Although we have the impression that we can see an entire line of text or even an entire page of text, this is an illusion. This fact has been clearly demonstrated in a number of studies over the years that use a gaze-contingent moving window paradigm (see Figure 2), introduced by McConkie and Rayner (1975; Rayner & Bertera, 1979).

The logic behind the gaze-contingent moving window paradigm is that we can determine how much of the text is being processed by varying the window size in which the text appears normal. Therefore, how much information is available to a reader is determined by how large the window (showing the normal text) has to be for readers to read normally. Or conversely, how small can the window get before there is disruption to reading. In this paradigm, the text within the window is normally presented, but the text outside the window is manipulated by either replacing the correct letters with other letters or replacing them all with X's. Studies using this paradigm have demonstrated that English readers acquire useful information from an asymmetrical region around the fixation point (extending 3-4 character spaces to the left of fixation and about 14-15 character spaces to the right)[2]. So, if the fixated word and the word to the right of fixation are normal and all other letters are replaced with visually similar letters, we find that readers are not aware that the words outside of the window have been changed, and their reading speed only decreases by about 10%. If two words to the right of fixation are normally displayed, then we see little slowdown in reading. Research has also found that readers do not utilize information from the words on the line below the currently fixated line (Rayner, 1998). Finally, the moving window paradigm can be reversed to produce a moving mask paradigm (Rayner & Bertera, 1979; Rayner, Inhoff, Morrison, Slowiaczek, & Bertera, 1981). In this case, a mask moves with the eyes on each fixation covering the letters in the center of vision (see Figure 2). From this paradigm, results demonstrate that reading is almost impossible when only letters in parafoveal vision are presented normally, while the central region is masked.

Insert Figure 2 about here

Another gaze-contingent display change paradigm (see Figure 2), called the boundary paradigm (Rayner, 1975), has been used to investigate the degree to which information outside the fixated region is processed. In this paradigm, the word to the right of fixation can be the target word (i.e., a valid preview), or it may be replaced with another word, a non-word, or random string of letters (i.e., an invalid preview). When the reader’s eye movement crosses an invisible boundary location, the preview (either the valid preview or an invalid preview) changes to the target word; since vision is suppressed during the eye movement, readers are not aware of the change. Researchers have found that when a reader has a invalid preview of the word to the right of fixation, they spend about 30-50 ms more fixating that word (following a saccade to it) than when they have a valid preview. It is also interesting to note that using this technique, studies have revealed that readers don’t combine a literal representation of the visual information across saccades, but rather combined abstract (and phonological) information across eye fixations (McConkie & Zola, 1979; Rayner, McConkie, & Zola, 1980).

Linguistic Influences on Fixation Time. As briefly mentioned above, the difficulty of the text being read has an impact on eye movement patterns (fixation duration, saccade length, and frequency of regressing to previously read text). Over the past few years, it has become very clear that how long the eyes remain in place is influenced by a host of linguistic factors. These factors include the frequency of the fixated word (Inhoff & Rayner, 1986; Rayner & Duffy, 1986), how predictable the fixated word is (Ehrlich & Rayner, 1981; Rayner & Well, 1996), how many meanings the fixated word has (Duffy, Morris, & Rayner, 1988; Sereno, O’Donnell, & Rayner, 2006), when the meaning of the word was acquired (Juhasz & Rayner, 2003, 2006), semantic relations between the word and prior words (Carroll & Slowiaczek, 1986; Morris, 1994), and how familiar the word is (Williams & Morris, 2004). For a more in-depth review of these factors and others, see Rayner (1998.)

The most compelling evidence that cognitive processing drive the eyes through the text perhaps comes from experiments in which the fixated word either disappears or is masked after 50-60 ms (Ishida & Ikeda, 1989; Liversedge, Rayner, White, Vergilino-Perez, Findlay, & Kentridge, 2004; Rayner et al., 1981; Rayner, Liversedge, & White, 2006; Rayner, Liversedge, White, & Vergilino-Perez, 2003). In these studies, readers are allowed to see the fixated word for 50-60 ms before it disappears. Under these conditions, they read quite normally; however, if the word to the right of fixation also disappears or is masked, then reading is disrupted (Rayner et al., 2006). This result indicates quite strongly that the word to the right of fixation is very important in reading. For the present purposes, it is important to note that when the fixated word disappears after 60 ms, fixation duration on that word is determined by the frequency of the word: if it is a low frequency word, the eyes remain in place longer (Rayner et al., 2003, 2006). So, even though the word is no longer visible, how long the eyes remain fixated is determined by that words' frequency. This evidence is very compelling and supports the notion that the cognitive processes associated with a fixated word are what drive eye movements through the text.

To summarize, it is clear that readers obtain information from a limited region during a fixation (extending to about 14-15 character spaces to the right of fixation). Information used for word identification is obtained from an even smaller region (extending to about 7-8 character spaces to the right of fixation). Furthermore, the word to the right of fixation is important and some information about that word is processed, as shown by preview benefit from that word. On some fixations, readers process the meaning of the fixated word and the word to the right of fixation, and will subsequently skip the word to the right of fixation. Finally, the ease or difficulty associated with processing the fixated word strongly has an effect on how long readers look at that word.

Models of Eye Movements in Reading. Based on the vast amount of information about eye movements during reading learned in the past 25-30 years, a number of models of eye movements in reading have recently appeared. The E-Z Reader model (Pollatsek, Reichle, & Rayner, 2006; Rayner, Ashby, Pollatsek, & Reichle, 2004; Rayner, Reichle, & Pollatsek, 1998; Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Pollatsek, & Rayner, 2006; Reichle, Rayner, & Pollatsek, 2003) is typically regarded as the most influential of these models. Due to space limitations, other models will not be discussed here[3]. The E-Z Reader model can account for all of the results discussed above, and also does a good job of predicting how long readers look at words, which words they skip, and which words will most likely be refixated. Importantly, it can account for global aspects of eye movements in reading, as well as more local processing characteristics; competitor models also account for similar amounts of data (see the chapter by Risse et al.). In many ways, these models share many similarities, though they differ on some precise processing details and on how certain effects are explained. As a computational model, E-Z Reader has the virtue of being highly transparent, so it makes very clear predictions and when it can't account for certain data, it is very clear why it can't (thus enabling one to change parameter values in the model). The model has also enabled us to account for data patterns that in the past may have been difficult to explain. However, this model isn't perfect and it has some limitations. For example, in this model, higher order processes due to sentence parsing and discourse variables currently have no influence. The model primarily assumes that lexical processing is driving the eyes through the text, and we believe that this isn't an unreasonable assumption.[4]

With careful experimentation and with the implementation of computational models that simulate eye movements during reading, we see from the foregoing that great advances have been made in understanding eye movements in reading (and inferring the mental processes associated with reading). In the next two sections, eye movements during visual search and scene perception will be discussed, in turn. Although there hasn't been as much research on these areas as on reading, it is still the case that some clear conclusions emerge from the work that has been done.

Eye Movements and Visual Search

Visual search is a research area that has received considerable effort over the past 40 years. However, the vast majority of this research has been done in the absence of considering eye movements (Findlay & Gilchrist, 1998), because it was largely assumed that eye movements are not particularly important in understanding search. With more recent research, this attitude seems to be changing, as there are now many more experiments examining aspects of visual search which report eye movement measures (see Benson, this volume; Menneer, Stroud, Cave, Donnelly, & Rayner, this volume). Though, many of these studies are often focused on using the search task to uncover properties of the saccadic eye movement system itself (see Findlay, 2004; Findlay & Gilchrist, 2003), many also use eye movements to learn about search behavior.

In this chapter, we'll focus primarily on research that has some implications for how viewers search through arrays to find specific targets. Fixation durations in search tend to be highly variable. Some studies report average fixation times as short as 180 ms while others report averages on the order of 275 ms. This is undoubtedly due to the fact that the level difficulty of a search (i.e., how dense or cluttered the array is) and the exact nature of the search task will strongly influence how long viewers pause on each item. Typically, saccade size is a bit larger than in reading (though saccades can be quite short with very dense arrays).

The Search Array Matters. Perhaps the most obvious thing about visual search is that the properties of the search array have an enormous influence on the ease with which a target is found. When an array is very cluttered (with many objects and distractors), the search becomes more demanding than when the array is simple. The eye movements on each of these types of arrays typically reflect this property of an array (Bertera & Rayner, 2000; Greene & Rayner, 2001a. 2001b). As the array becomes more complicated, we see an increase in the fixation duration and the number of fixations, as well as a decrease in the average saccade length (Vlaskamp & Hooge, 2006). Another influential property on the pattern of eye movements is the configuration of the search array. An array of objects arranged in an arc tends to have more fixations that fall in-between objects, progressively getting closer to the area where viewers think the target is located (Zelinsky, 2005; Zelinsky, Rao, Hayhoe, & Ballard, 1997). On the other hand, in randomly placed arrays, fixations tend to fall directly on an item within an array and other factors such as color of the items and shape similarity to the target object preferentially attract fixations (Williams, Henderson, & Zacks, 2005).

Does Visual Search Have a Memory? Horowitz and Wolfe (1998) initially proposed that visual search doesn't have a good memory and that the same item will be re-sampled during multiple encounters with the same search array. However, this assertion was based on reaction time functions. Ideally, eye movement data would better address this question, because one can directly examine how often the eyes return to a previously sampled part of the array. Studies using eye movement measures have made it clear that while searching, viewers generally do not return to previously searched items (Beck, Peterson, Boot, Vomela, & Kramer, 2006; Beck, Peterson, & Vomela, 2006; Peterson, Kramer, Wang, Irwin, & McCarley, 2001).

The Perceptual Span. As discussed in the reading section of this chapter, how much information is processed on each fixation is an important factor in determining eye movement patterns. Using the moving window paradigm, Rayner and Fisher (1987a, 1987b) asked viewers to search through horizontally arranged letter strings for a specified target letter. They found that the size of the perceptual span varied as a function of the difficulty of the distractor letter. That is, the perceptual span was smaller when distractor letters were visually similar to the target letter than when the distractor letters were distinctly different. This suggested that there were two qualitatively different regions within the span: a decision region (where information about the presence or absence of a target is available), and a preview region where some letter information is available, but where information on the absence of a target is not yet available.

In another study, Bertera and Rayner (2000) asked viewers to search though a randomly arranged array of letters and digits for the presence of a target letter. They used both the moving window and moving mask techniques. They varied the density of the array by changing the size of the array (so that it was 13 degrees by 10 degrees, 6 degrees by 6 degrees, or 5 degrees by 3.5 degrees), while holding the number of items constant. The size of the moving mask was directly related to the accuracy and the time it took to complete the search: increasing the mask size, increased search time and decreased accuracy. When given a moving window, viewers' search performance reached asymptote when the window was approximately 5 degrees. That is, all letters/digits falling within 2.5 degrees from the fixation point where visible, while all other items were masked.

Where and When to Move the Eyes. Many studies have tried to determine the factors involved in deciding where and when to move the eyes during search (Greene, 2006; Greene & Rayner, 2001a, 2001b; Hooge & Erkelens, 1996, 1998; Jacobs, 1986; Vaughan, 1982). However, despite this effort, no clear answers have emerged. In some cases, researchers concluded that fixation durations in visual search are the result of both preprogramming saccades and processing currently fixated information (Vaughan, 1982). In other cases, some researchers suggested that the completion of foveal analysis is not necessary to trigger an eye movement (Hooge & Erkelens, 1996, 1998), while others have suggested that it is (Greene & Rayner, 2001b). Rayner (1995) suggested that the trigger to move the eyes in a search task is determined by a check processes that determines whether the target present in the decision area of the perceptual span. If not, a new saccade is programmed to move the eyes to a location that has not been examined. As with reading, attention would first move to the region targeted for the next saccade (and as described above, would favor items that shared a similar shape or color with the target, Williams et al., 2005).

As a final point, the decision about where to fixate next and when to move the eyes will undoubtedly be influenced by characteristics of the specific search task and the density of the array. In a recent study, van Zoest, Donk, and Theeuwes (2004) investigated what type of information had more influence over the placement of fixations: goal-driven information (i.e., target knowledge) or distractor saliency. Results indicated that the timing of the fixations was linked to where the eye would land. When fixations were made quickly, viewers tended to fixate the target and distractor equally. However when the fixation latency was longer, the target was more likely fixated. So, the longer viewers took to choose a location and execute the saccade, the more likely it would be influenced by goal-driven control. Consequently, it seems that for fixations made early on, visual saliency may play a greater role in directing fixations.

Eye Movements and Scene Perception

The eye movement pattern of a viewer on a scene is shown in Figure 3a. As is very evident in this figure, not every part of the scene is fixated. This is largely because in scene perception information can be obtained over a wider region than is found in reading and possibly, visual search arrays. However, it is clear that the important aspects of the scene are typically fixated (and generally looked at for longer periods than less important parts of the scene). In Figure 3b, the fixations from multiple viewers have been compiled and show that most fixations are on the informative parts of the scene. That is, viewers do not tend fixate on the sky or the road in front of the houses. As noted at the outset, the average fixation duration in scene perception tends to be longer than that in reading, and likewise the average saccade size tends to be longer. In this section, we will briefly summarize where people tend to look in scenes, the perceptual span region for scenes, and the nature of eye movement control when looking at scenes.

Insert Figure 3 about here

Getting the Gist of a Scene. The gist of a scene has been defined as the general scene concept (Potter, 1999), and is most often referring to a scene's basic-level category when investigated in the literature (Oliva, 2005). One very important general finding is that viewers are able to acquire scene gist in a single glance. That is, the gist of the scene is understood so quickly, it is thought to be processed even before the eyes begin to move (De Graef, 2005). In a recent study, Castelhano and Henderson (2007b) showed that when viewers were shown a scene for as little as 40 ms, they were able to extract enough information to understand the scene gist. Typically, the scene gist is thought to be acquired within the first fixation in order to orient subsequent fixations to appropriate or interesting regions within the scene. We will turn to this question in the next section.

Where do Viewers Look in Scenes? Pioneering works of Buswell (1938) and Yarbus (1967) first documented how a viewer's gaze is drawn to important aspects of a visual scene and that task goal very much influences eye movements. Much of the research that followed illustrated that the eyes are drawn to informative areas in a scene quickly (Antes, 1974; Mackworth & Morandi, 1967) and an object that is out-of-place in a scene tends to attract a lot of early fixations (Friedman, 1979; Loftus & Mackworth, 1978). It is important to note however, that the out-of-place objects in these studies tended to differ from the appropriate objects on a number of dimensions (Rayner & Pollatsek, 1992). For example, an octopus in a farm scene is not only semantically out-of-place, but it also tends to have more rounded features than the objects typically in a farm scene. So, these early studies confounded visual saliency and semantic saliency. Recent experiments that control for visual saliency of the out-of-place objects raise some uncertainty about the earlier findings, and suggest that the eyes are not immediately drawn to out-of-place objects that differ in semantics alone (De Graef, Christiaens, & d'Ydewalle, 1990; Henderson, Weeks, & Hollingworth, 1999), though see Becker, Pashler, and Lubin (2007) for conflicting evidence.

Nonetheless, it has become clear that the eye can quickly go to parts of a scene that are relevant and important. A recent study investigated the influence of scene context on the placement of eye movements while viewers searched for certain objects within pseudo-realistic scenes (Neider & Zelinksy, 2006). Viewers were asked to look for a target objects that was typically constrained to certain part of the scene (i.e., a jeep on the ground or a blimp in the sky). When a target was present, fixations were largely limited to the area one would expect to find the target object (i.e., ground or sky); whereas, when the target was absent, eventually there was less of an inclination to restrict search to these areas. In addition, when the target was in its expected area, search times were about 19% faster. From these results, they concluded that not only do viewers immediately focus fixations in the target's expected area, but also that the visual system is flexible enough to quickly switch to a "look everywhere" strategy when the first proves fruitless. Thus, it seems that although likely target context can serve as a good search strategy, it is not generally with strict adherence.

Other studies have also made it clear that saliency of different parts of the scene greatly influences where viewers tend to fixate (Parkhurst & Niebur, 2003; Mannan, Ruddock, & Wooding, 1995, 1996). As discussed with visual search, saliency is typically defined in terms of low-level components of the scene (such as contrast, color, intensity, brightness, spatial frequency, etc.). There are a number of computational models (Baddeley & Tatler, 2006; Itti & Koch, 2000, 2001; Parkhurst, Law, & Niebur, 2002) that use the concept of a saliency map to model eye fixation locations in scenes. With this approach, the bottom-up properties of a scene (i.e., saliency map) make explicit predictions about the most visually prominent regions of the scene. The models are basically used to derive predictions about the distribution of fixations on a given scene based on these prominent regions.

While these models can account for some of the variability in the distribution of fixations across a scene, they are limited by the assumption that fixations are primarily driven by bottom-up factors. Research has shown that higher-level factors also have a strong influence on where viewers direct their gaze in a scene (Castelhano & Henderson, 2007a; Henderson & Castelhano, 2005; Henderson & Ferreira, 2004). Recently, Torralba, Oliva, Castelhano, and Henderson (2006) presented a model that incorporates the influence of top-down and cognitive strategies. Although there is considerable research on localizing *where* viewers move their eyes while looking at scenes, very few studies have investigated what controls *when* the eyes move. In contrast, reading research has focused on both *where* to move the eyes and *when* to move the eyes. One recent study that attempts to investigate when the eye move looked at the effect of repeated exposure to a scene on fixation durations (Hidalgo-Sotelo, Oliva & Torralba, 2005). In this study, viewers' task was to search for a target (i.e., a person) in the scene while their eye movements were tracked. Unbeknownst to them, there were certain scene-target combinations that repeated throughout the experiment twenty times. As expected, these repeated searches showed a large decrease in response time. Interestingly, although the number of fixations did not decrease, the average fixation duration prior to fixating the target object did. Furthermore, the results showed that the proportion of target objects that were fixated before a response was made did not change with increased repetitions (85%). Furthermore, average gaze durations on the target dropped from 450 ms during the first exposure to 310 ms in the twentieth. Despite the fact that the target presence was highly correlated with the scene, it seems that observers chose to verify the target object presence before making their response. These results demonstrated that with repeated exposure, the reduced response time is primarily due to a decrease in the average duration of fixations during the search, as well as in the time taken to verify the target object. Thus, performance improvements on a repeated search stemmed from a facilitation of identifying the fixated regions as non-targets and targets, but not from a reducing the number of fixations made overall.

Another difference between scenes and reading is the question of how recently acquired information is used to guide eye movements. We know that in reading, prior context provided by the material read up until that point plays a large role in determining the directing of subsequent fixations (such as deciding between a forward-going saccade or a regressive saccade). In scenes, the role that previously viewed information plays in directing fixations is not as clear. Many of the models that use saliency as the primary driving force of the eyes do not consider how information gathered initially may influence subsequent fixations. In a recent study, Castelhano and Henderson (2007a) investigated whether the initial representation of a scene can be used to guide eye movements on a real-world scene. After a brief preview of the search scene, viewers were asked to find the target object. In the experiment, the moving window technique was used and it blocked out all scene information, with the exception of the scene at the point of fixation; thus, immediately available visual information from the surrounding scene could not be used to plan fixations. Instead, viewers had to rely on the information that was gathered during the preview. When the preview was of the search scene itself, performance was much more efficient than when the preview was a meaningless control.

Using a preview of another scene within the same semantic category (thereby providing general semantic information without the same visual details) resulted in no improvement in search efficiency; this suggests that the information used to guide the eyes was not based on general semantics. With a reduced scale preview of the search scene, performance was just as high as when the full-scale preview was shown. Taken together, this study suggests that to guide the eyes, the visual system relied on abstract visual information. Thus, the information used to guide eye movements in scenes is said to have multiple sources, such as the saliency of the regions, abstract visual information previously extracted from the scene, and top-down knowledge about the context and scene type.

The Perceptual Span. How much information is extracted from a single fixation on a scene? As noted at the beginning of this section, it is known the extent of the visual field used to extract useful information is much larger in scene viewing than it is in reading. Using a moving mask paradigm (to cover the part of the scene around the fixation point) during scene exploration, Henderson, McClure, Pierce, and Shrock (1997) found that although the foveal mask had an influence on looking time, it did not have nearly as deleterious effects for object identification as was found for word identification in reading.

In an early study, Nelson and Loftus (1980) examined object recognition as a function of the closest fixation on that object. Results showed that objects located within about 2.6 degrees from fixation were generally recognized. The results also suggested that information is acquired from the region 1.5 degrees around fixation is qualitatively different from regions further away (see also Nodine, Carmody, & Herman, 1979). A study by Parker (1978) found that the functional field of view for specific objects in a scene is quite large: with a radius of at least 10° around fixation resulting in a perceptual span of up to 20° (see Henderson & Ferreira, 2004 for discussion). However, more recent studies using more natural images suggest that the functional field of view extends about 4° away from fixation (Henderson & Hollingworth, 1999; Henderson, Williams, Castelhano, & Falk, 2003).

In an early study, Saida and Ikeda (1979) used the moving window paradigm and found that the functional field of view is quite large, and can consist of about half of the total scene regardless of the absolute size of the scene (at least for scenes that are up to 14.4 degrees by 18.8 degrees). In this and other studies using the moving window paradigm (van Diepen & d'Ydewalle, 2003; van Diepen, Wampers, & d'Ydewalle, 1998) normal scene information within the window area around a fixation point is presented normally, but the information outside of the window is degraded in some systematic way. Saida and Ikeda (1979) found a serious deterioration in recognition of a scene when the window was limited to a small area (about 3.3 degrees X 3.3 degrees) on each fixation. Performance gradually improved as the window size became larger, as noted, up to about 50% of the entire scene. Saida and Ikeda noted that there was considerable overlap of information across fixations.

Based on the studies reviewed above, it should now be clear that the question of how large the perceptual span during scene viewing hasn't been answered as conclusively as it has in reading or visual search. And yet, it does appear that viewers typically gain useful information from a fairly wide region of the scene and that it probably varies as a function of the scene and the task of the viewer. For example, the ease with which an object is identified within a scene is influenced by its orientation (Boutsen, Lamberts, & Verfaille, 1998), how well camouflaged it is (De Graef, et al., 1990) and the frequency with which it is typically encountered within a scene context (Hollingworth & Henderson, 1998). As shown for the perceptual span in reading (Henderson & Ferreira, 1990), it is highly probable that identification of the fixated object has an effect on the extent to which information in the periphery is processed.

Preview Benefit. Just as with reading, before fixating an object, viewers obtain preview benefit from it (Henderson, 1992; Henderson, Pollatsek, & Rayner, 1987, 1989; Pollatsek, Rayner, & Collins, 1984; Pollatsek, Rayner, & Henderson, 1990). Further, the preview benefit for objects is about 100 ms, making it much larger than the effects found for words in reading. Interestingly, viewers are rather insensitive to large changes in a scene. In a series of experiments by McConkie and Grimes (McConkie, 1991; Grimes & McConkie, 1995; Grimes, 1996) viewers were asked to try to memorize scenes that were shown for a predetermined amount of time. In addition, they were informed that it was possible that some changes to the image could be made while they were studying it, and they were instructed to press a button when one of those changes was detected. During this study phase, changes to the scene were made when the viewer launched a saccade. As discussed earlier, vision is suppressed during saccades so that these changes would not have been visible as they were occurring. Surprisingly, observers were unaware of most changes, which included the appearance and disappearance of large objects and the changing of colors of large regions of the scene. All of these changes occurred while the scene was being viewed. Although later studies found that any disruption served to induce an inability to detect changes, such as inserting a blank screen in between two changing images (Rensink, O'Regan & Clark, 1997), movie cuts (Levin & Simons, 1997), or the simultaneous onset of patches or 'mudsplashes' covered portions of the scene (O'Regan, Rensink, & Clark, 1999), these experiments highlighted the relation between what is viewed during the initial exploration of a scene and what is remembered about that scene. Further studies have shown that this lack of awareness does not mean that there is no recollection of any visual details, but rather that the likelihood of remembering visual information is highly dependent on the processing of that information (Hollingworth & Henderson, 2002; Hollingworth, 2003). Thus, knowing something about the processes that go on during a fixation on a scene is extremely important if one would want to predict how well visual information will be remembered later on, or even while the scene is being explored further.

When do viewers move their eyes when looking at scenes? Past studies have shown that attention precedes an eye movement to a new location within a scene (Henderson, 1992; van Diepen & D'Ydewalle, 2003). So, it would follow that the eyes will move once the visual information at the center of vision has been processed and a new fixation location has been selected and programmed. In a recent study, van Diepen and D'Ydewalle (2003) investigated when this shift in attention (from the center of fixation to the periphery) took place in the course of a fixation. They had observers view scenes whose information at the center of fixation was masked during the initial part of fixations (from 20- 90 ms). In another case, the periphery was masked at the beginning of each fixation (for 10-85 ms). As expected, this study demonstrated that when the center of fixation was masked initially, fixation durations increased with longer mask durations (61% increase). When the periphery was masked, they found a slight increase in fixation durations, but not as much as with a central mask (15% increase). Interestingly, their results also showed that as mask durations in the periphery were increased, the average distance of saccades decreased and the number of fixations increased. They surmised that increasing the duration of the peripheral mask does not prevent the visual system from making a saccade, but instead information that is immediately available is often chosen for the next fixation. Taken together, these results suggest that at the fovea information is extracted very rapidly, and attention is directed to the periphery almost immediately following the extraction of information (70-120 ms) to choose a viable saccade target. The general timing of the switch between central and peripheral information processing is currently being investigated; however, the inherent variability across scenes makes it difficult to find as specific a time frame as in reading.

Finally, a very interesting issue that has received attention lately is the extent to which emotional aspects of a scene influence eye movements. Do emotional aspects of a scene lead viewers to quickly look at that part of the scene conveying strong emotional content? Nummenmaa, Hyönä, and Calvo (this volume) review this issue. It seems quite clear that emotional aspects of scenes do have an immediate impact on eye movements.

General Comments on Eye Movements

In the preceding sections, we have reviewed research on eye movements in three tasks: reading, visual search and scene perception. Although there are obviously many differences between these tasks, there are some general principles that are likely to hold across them. First, how much information is processed on any fixation (the perceptual span or functional field of view) varies as a function of the task. The perceptual span is clearly smaller in reading than in either scene perception or visual search. Hence, for example, fixations in scene perception tend to be longer and saccades are longer because more information is being processed in a single fixation. Second, the difficulty of the stimulus influences eye movements: in reading, when the text becomes more difficult, eye fixations get longer and saccades get shorter; likewise in scene perception and visual search, when the stimulus is more crowded, cluttered, or dense, fixations get longer and saccades get shorter. Fourth, the difficulty of the specific task (reading for comprehension versus reading for gist, searching for a person in a scene versus looking at the scene for a memory test, and so on) obviously influences how the eyes move. Finally, in all three tasks there is some evidence (Najemnik & Geisler, 2005; Rayner, 1998) that viewers integrate information poorly across fixations and it is more critical that information is processed efficiently on each fixation.

Cultural Influences on Eye Movements

In the prior sections of this chapter, we have not commented on cultural influences on eye movements. However, there is currently a great deal of interest in this issue and many chapters in the present volume deal specifically with this issue. The chapter by Rayner, Li, and Pollatsek (this volume) provides a discussion of the extension of the E-Z Reader model to Chinese and a number of the chapters in the volume deal specifically with eye movements when reading Chinese text. The other recent development receiving attention is the extent to which cultural differences can influence how people look at a scene. The chapter by Boland, Chua, and Nisbett (this volume) reviews their interesting finding that culture can influence what people look at (and how early they look at) certain aspects of a scene. The chapter by Li, Williams, Cave, Well, and Rayner (this volume) raises some qualifications concerning the cultural effect. Clearly, more research is needed on this issue.

Summary

In this chapter, we have reviewed the basic findings concerning eye movements when (1) reading, (2) looking at a scene, and (3) searching through a visual array. Although there is no question that these tasks differ considerably, and that eye movements also differ considerably as a function of the task, it is the case that eye movements can be very informative about what exactly viewers do in each type of task. Each of these points has been discussed in the preceding sections. It will also be interesting to see how well the findings we have described above will hold up when viewers look at dynamically changing scenes (virtually all of the work that we described has dealt with static scenes).

We also suspect, as is clearly evident in the present volume, that there will be more and more research examining cultural influences on eye movements. It is also worth noting that eye movement data are now widely used to examine cognitive processing in a number of real world tasks such as driving (see Pollatsek, Fisher, and Pradhan, this volume), mental problem solving and visual comparisons (Cave et al., this volume), making tea and sandwiches (Land & Hayhoe, 2001), and so on. And, eye movement data are now also widely used in the so-called “visual world” paradigm (see Ferreira & Henderson, 2004) in which people listen to speech while simultaneously looking at a visual array. This research shows quite clearly that people tend to look at that part of the array that is being discussed in the narrative. Finally, as should be apparent from the foregoing discussion, our expectation is that eye movements will continue to play a valuable role for those interested in visual cognitive processing.

References

Antes, J.R. (1974). The time course of picture viewing. *Journal of Experimental Psychology, 103,* 62-70.

Baddeley, R. J., & Tatler, B. W. (2006). High frequency edges (but not contrast) predict where we fixate: a Bayesian system identification analysis. *Vision Research, 46,* 2824-2833.

Beck, M.R., Peterson, M.S., Boot, W.R., Vomela, M., & Kramer, A.F. (2006). Explicit memory for rejected distractors during visual search. Visual Cognition, 14, 150-174.

Beck, M.R., Peterson, M.S., & Vomela, M. (2006). Memory for where, but not what, is used during visual search. Journal of Experimental Psychology: Human Perception and Performance, 32, 235-250.

Becker, M.W., Pashler, H., & Lubin, J. (2007). Object-intrinsic oddities draw early saccades. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 20-30

Becker, W., & Jürgens, R. (1979). Analysis of the saccadic system by means of double step stimuli. *Vision Research, 19,* 967-983.

Bertera, J. H., & Rayner, K. (2000). Eye movements and the span of effective stimulus in visual search. *Perception & Psychophysics, 62*, 576-585.

Boutsen, L., Lamberts, K. & Verfaillie, K. (1998) Recognition times of different views of 56 depth-rotated objects: A note concerning Verfaillie and Boutsen (1995). *Perception & Psychophysics*, 60, 900-907.

Buswell, G. T. (1935). *How people look at pictures.* Chicago: University of Chicago Press.

Carroll, P.J., & Slowiaczek, M.L. (1986). Constraints on semantic priming in reading: A fixation time analysis. *Memory & Cognition, 14,* 509-522.

Castelhano, M. S., & Henderson, J. M. (2007a). Initial scene representations facilitate eye movement guidance in visual search. *Journal of Experimental Psychology: Human Perception and Performance*, in press.

Castelhano, M. S., & Henderson, J. M. (2007b). The influence of color on perception of scene gist. *Manuscript under review.*

De Graef, P. (2005). Semantic effects on object selection in real-world scene perception. In G. Underwood (ed), Cognitive processes in eye guidance. Oxford: Oxford University Press.

De Graef, P., Christiaens, D., & d'Ydewalle, G. (1990). Perceptual effects of scene context on object identification. *Psychological Research, 52*, 317-329.

Deubel, H., & Schneider, W.X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research, 36*, 1827-1837.

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language, 27,* 429-446.

Ehrlich, S. E, & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior, 20,* 641-655.

Findlay, J. M. (2004). Eye scanning and visual search. In J. M. Henderson, and F.Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp 135-160). New York: Psychology Press.

Findlay, J. M., & Gilchrist, I. D. (1998). Eye guidance and visual search. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 295-312). Oxford, England: Elsevier.

Findlay, J. M., & Gilchrist, I. D. (2003). *Active vision. The Psychology of looking and seeing*. Oxford: Oxford University Press.

Friedman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General, 108,* 316-355.

Greene, H. (2006). The control of fixation duration in visual search. *Perception, 35*, 303-315.

Greene, H., & Rayner, K. (2001a). Eye movements and familiarity effects in visual search. *Vision Research, 41*, 3763-3773.

Greene, H., & Rayner, K. (2001b). Eye-movement control in direction-coded visual search. *Perception, 29*, 363-372.

Grimes, J. (1996). On the failure to detect changes in scenes across saccades. In K. Akins (Ed.), *Vancouver studies in cognitive science: Vol. 5. Perception* (pp. 89–110). New York: Oxford University Press.

Grimes J., & McConkie, G. (1995). On the insensitivity of the human visual system to image changes made during saccades. In K. Akins, (Ed.), *Problems in Perception.* Oxford, UK: Oxford University Press.

Henderson, J. M. (1992). Identifying objects across saccades: Effects of extrafoveal preview and flanker object context. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 521-530.

Henderson, J.M, & Castelhano, M.S. (2005). Eye Movements and Visual Memory for Scenes. In G. Underwood (Ed.), *Cognitive Processes in Eye Guidance* (pp. 213-235). Oxford University Press.

Henderson. J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16(3),* 417-429.

Henderson. J. M., & Ferreira, F. (2004). Scene perception for psycholinguists. In J. M. Henderson, and F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp 1-58). New York: Psychology Press.

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology, 50*, 243–271.

Henderson, J. M., McClure, K. K., Pierce, S., & Schrock, G, (1997). Object identification without foveal vision: Evidence from an artificial scotoma paradigm. *Perception & Psychophysics, 59*, 323-346.

Henderson, J. M., Pollatsek, A., & Rayner, K. (1987). The effects of foveal priming and extrafoveal preview on object identification. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 449-463.

Henderson, J.M., Pollatsek, A., & Rayner, K. (1989). Covert visual attention and extrafoveal information use during object identification. *Perception & Psychophysics, 45,* 196-208.

Henderson, J. M., Weeks, P. A. Jr., & Hollingworth, A. (1999). Effects of semantic consistency on eye movements during scene viewing. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 210-228.

Henderson, J. M., Williams, C. C., Castelhano, M. S., & Falk, R. J. (2003). Eye movements and picture processing during recognition. *Perception & Psychophysics, 65,* 725-734.

Hidalgo-Sotelo B, Oliva A, Torralba A (2005). Human Learning of Contextual Priors for Object Search: Where does the time go? *Proceedings of 3rd Workshop on*

Attention and Performance at the International Conference in Computer Vision and Pattern Recognition (CVPR), San Diego CA.

Hoffman, J.E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics, 57*, 787-795.

Hollingworth, A. (2003). Failures of retrieval and comparison constrain change detection in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 388-403.

Hollingworth, A., & Henderson, J. M. (1998). Does consistent scene context facilitate object perception? *Journal of Experimental Psychology: General, 127,* 398-415.

Hollingworth, A., & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 113-136.

Hooge, I. T. C., & Erkelens, C. J. (1996). Control of fixation duration during a simple search task. *Perception & Psychophysics, 58,* 969-976.

Hooge, I. T. C., & Erkelens, C. J. (1998). Adjustment of fixation duration during visual search. *Vision Research, 38,* 1295-1302.

Horowitz, T. S. & Wolfe, J. M. (1998). Visual search has no memory. *Nature, 94,* 575-577.

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics*, *40*, 431-439.

Irwin, D.E. (2004). Fixation location and fixation duration as indices of cognitive processing. In J.M. Henderson & F. Ferreira (eds) *The interface of language, vision, and action: Eye movements and the visual world* (pp 105-134). New York: Psychology Press.

Ishida, T., & Ikeda, M. (1989). Temporal properties of information extraction in reading studied by a text-mask replacement technique. *Journal of the Optical Society A: Optics and Image Science, 6,* 1624-1632.

Itti., L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*, 1489-1506.

Itti, L., & Koch, C. (2001). Computational modeling of visual attention. *Nature Reviews: Neuroscience, 2*, 194-203.

Jacobs, A.M. (1986). Eye movement control in visual search: How direct is visual span control? *Perception & Psychophysics, 39*, 47-58.

Juhasz, B.J., & Rayner, K. (2003). Investigating the effects of a set of intercorrelated variables on eye fixation durations in reading. *Journal of Experimental Psychology: Learning, Memory & Cognition, 29*, 1312-1318.

Juhasz, B.J., & Rayner, K (2006). The role of age-of-acquisition and word frequency in reading: Evidence from eye fixation durations. *Visual Cognition, 13,* 846-863.

Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in programming saccades. *Vision Research, 35*, 1897-1916.

Land, M.F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research, 41*, 3559-3565.

Levin, D. T. & Simons, D. J. (1997). Failure to detect changes to attended objects in morion pictures. *Psychonomic Bulletin & Review, 4 (4),* 501-506.

Liversedge, S.P., Rayner, K., White, S.J., Vergilino-Perez, D., Findlay, J.M., & Kentridge, R.W. (2004). Eye movements while reading disappearing text: Is there a gap effect in reading? *Vision Research, 44*, 1013-1024.

Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 565-572.

Mackworth, N. H., & Morandi, A. J. (1967). The gaze selects informative details within pictures. *Perception & Psychophysics, 2,* 547-552.

Mannan, S. K., Ruddock, K. H., & Wooding, D. S. (1995). Automatic control of saccadic eye movements made in visual inspection of briefly presented 2-D images. *Spatial Vision, 9,* 363-386.

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

McConkie, G.W. (1991). Perceiving a stable visual world. In *Proceedings of the Sixth European Conference on Eye Movements*, pages 5–7. Leuven, Belgium: Laboratory of Experimental Psychology.

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics, 17,* 578-586.

McConkie, G. W., & Zola, D. (1979). Is visual information integrated across successive fixations in reading? *Perception & Psychophysics, 25,* 221-224.

McPeek, R. M., Skavenski, A. A., & Nakayama, K. (2000). Concurrent processing of saccades in visual search. *Vision Research, 40*, 2499–2516.

Morris, R. K. (1994). Lexical and message-level sentence context effects on fixation times in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 92-103.

Morrison, R. E., & Rayner, K. (1981). Saccade size in reading depends upon character spaces and not visual angle. *Perception & Psychophysics, 30*, 395-396.

Najemnik, J. & Geisler, W.S. (2005) Optimal eye movement strategies in visual search. *Nature*, *434*, 387-391.

Neider, M. B., & Zelinsky, G. J. (2006). Scene context guides eye movements during search.*Vision Research*, *46(5),* 614-621.

Nelson, W.W., & Loftus, G.R. (1980). The functional visual field during picture viewing. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 391-399.

Nodine, C. E, Carmody, D. P., & Herman, E. (1979). Eye movements during visual search for artistically embedded targets. *Bulletin of the Psychonomic Society*, *13,* 371-374.

Oliva, A. (2005). Gist of the scene. In L. Itti, G. Rees, and J.K. Tsotsos (Eds.),*The Encyclopedia of Neurobiology of Attention*. (pp. 251-256) San Diego: Elsevier.

O'Regan JK, Rensink RA, & Clark JJ (1999). Change blindness as a result of 'mudsplashes'. *Nature*, *398*, 34.

Parker, R. E. (1978). Picture processing during recognition. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 284-293.

Parkhurst, D. J., & Niebur, E. (2003). Scene content selected by active vision. *Spatial Vision*, *16*, 125–154.

Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, *42*, 107-123.

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.

Pollatsek, A., Rayner, K.,& Collins,W. E. (1984). Integrating pictorial information across eye movements. *Journal of Experimental Psychology: General, 113,* 426-442.

Pollatsek, A., Rayner, K., & Henderson, J. M. (1990). Role of spatial location in integration of pictorial information across saccades. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 199-210.

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology, 52*, 1-52.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32,* 3-25.

Potter, M. (1999). Understanding Sentences and Scenes: The role of conceptual short-term memory. In V. Coltheart (Ed.), *Fleeting Memories,* (pp. 13-46). Boston: MIT Press.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7,* 65-81.

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

Rayner, K. (1995). Eye movements and cognitive processes in reading, visual search, and scene perception. In J. M. Findlay, R. Walker, & R.W. Kentridge (Eds.), *Eye movement research: Mechanisms, processes and applications* (pp. 3-22). Amsterdam: North Holland.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 85*, 618-660.

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E.D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z Reader model. *Journal of Experimental Psychology: Human Perception and Performance, 30*, 720-732.

Rayner, K., & Bertera, J. H. (1979). Reading without a fovea. *Science*, *206*, 468-469.

Rayner, K., & Duffy, S.A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, *14*, 191-201.

Rayner, K., McConkie, G.W., & Ehrlich, S.F. (1978). Eye movements and integrating information across saccades. *Journal of Experimental Psychology: Human Perception and Performance*, *4,* 529-544.

Rayner, K., & Fisher, D. L. (1987a). Eye movements and the perceptual span during visual search. In J. K. O'Regan &A. Ltvy-Schoen (Eds.), *Eye movements: From physiology to cognition* (pp. 293-302). Amsterdam: North Holland.

Rayner, K., & Fisher, D. L. (1987b). Letter processing during eye fixations in visual search. *Perception & Psychophysics, 42,* 87-100.

Rayner, K., Inhoff, A. W., Morrison, R. E., Slowiaczek, M. L., & Bertera, J. H. (1981). Masking of foveal and parafoveal vision during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 167--179.

Rayner, K., Liversedge, S.P. & White, S.J. (2006). Eye movements when reading disappearing text: The importance of the word to the right of fixation. *Vision Research* 46: 310-323.

Rayner, K., Liversedge, S.P., White, S.J., & Vergilino-Perez, D. (2003). Reading disappearing text: Cognitive control of eye movements. *Psychological Science,* 14, 385-389.

Rayner, K., McConkie, G. W., & Zola, D. (1980). Integrating information across eye movements. *Cognitive Psychology, 12,* 206-226.

Rayner, K., & Pollatsek, A. (1989). The *psychology of reading.* Englewood Cliffs, NJ: Prentice Hall.

Rayner, K., & Pollatsek, A. (1992). Eye movements and scene perception. *Canadian Journal of Psychology, 46,* 342-376.

Rayner, K., Reichle, E. D., & Pollatsek, A. (1998). Eye movement control in reading: An overview and model. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 243-268). Oxford, England: Elsevier.

Rayner, K., Slowiaczek, M.L., Clifton, C., & Bertera, J.H. (1983). Latency of sequential eye movements: Implications for reading. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 912-922.

Rayner, K., Warren, T., Juhasz, B.J., & Liversedge, S.P. (2004). The effect of plausibility on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 1290-1301.

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review, 3,* 504-509.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105,* 125-157.

Reichle, E. D., Pollatsek, A. & Rayner, K. (2006). E-Z Reader: A cognitive-control, serial-attention model of eye-movement behavior during reading. *Cognitive Systems Research, 7*, 4-22.

Reichle, E.D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye movement control in reading: Comparison to other models. *Behavioral and Brain Sciences. 26*, 507-526.

Rensink RA, O'Regan JK, & Clark JJ (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, **8**:368-373.

Saida, S., & Ikeda, M. (1979). Useful field size for pattern perception. *Perception & Psychophysics, 25,* 119-125.

Sereno, S.C., O'Donnell, P.J., & Rayner, K. (2006). Eye movements and lexical ambiguity resolution: Investigating the subordinate bias effect *Journal of Experimental Psychology: Human Perception and Performance, 32*, 335-350.

Torralba, A., Oliva, A., Castelhano, M.S., & Henderson, J.M. (2006). Contextual guidance of attention in natural scenes. *Psychological Review, 113, 766-786.*

van Diepen, E M. J., & d'Ydewalle, G. (2003). Early peripheral and foveal processing in fixations during scene perception. *Visual Cognition, 10*, 79-100.

van Diepen, E M. J., Wampers, M., & d'Ydewalle, G. (1998). Functional division of the visual field: Moving masks and moving windows. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 337-356). Oxford, England: Elsevier.

van Zoest, L.J.F.M., Donk, M., & Theeuwes, J. (2004). The role of bottom-up control in saccadic eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 746-759.

Vaughan, J. (1982). Control of fixation duration in visual search and memory search: Another look. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 709-723.

Vlaskamp, B.N.S., & Hooge, I.T.C. (2006). Crowding degrades saccadic search performance. *Vision Research*, *46*, 417-425.

Williams, C. C., Henderson, J.M., & Zacks, R. T. (2005). Incidental visual memory for targets and distractors in visual search. *Perception & Psychophysics*, *67*, 816-827.

Williams, R.S., & Morris, R.K. (2004). Eye movements, word familiarity, and vocabulary acquisition. *European Journal of Cognitive Psychology, 16,* 312-339.

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

Zelinsky, G. (2005). Specifying the components of attention in a visual search task. In L. Itti, G. Rees, & J. Tsotsos (Eds.), *Neurobiology of attention* (pp. 395-400). Elsevier, Inc.

Zelinsky, G. J., Rao, R. P. N., Hayhoe, M. M., & Ballard, D. H. (1997). Eye movements reveal the spatiotemporal dynamics of visual search. *Psychological Science, 8(6),* 448-453.

Acknowledgments

Preparation of this chapter was supported by a Grant from the Microsoft Corporation and by Grant HD26765 from the National Institute of Health.
Correspondence should be addressed to Keith Rayner, Department of Psychology, University of Massachusetts, Amherst, MA 01003, USA.

Footnotes

1. It is important to note that although vision is suppressed, mental processes continue during the saccade for many cognitive tasks, (see Irwin, 2004 for a review of when cognition is also suppressed during saccades).

2. The size of the perceptual span is also influenced by the nature of the writing system, but a description of this influence is beyond the scope of this chapter (see Rayner, 1998 for a review).

3. For a comprehensive overview of these models, see the special issue of *Cognitive Systems Research*, 2006 (vol. 7). In addition, Risse, Engbert, and Kliegl (this volume) discuss the other major model, SWIFT.

4. Our primary argument is that lexical processing drives the eyes through the text and higher order processes primarily serve to intervene when something doesn't compute (see Rayner, Warren, Juhasz, & Liversedge, 2004).

Table 1. Eye movement characteristics in reading, scene perception, and visual search.

Task	Mean Fixation Duration (ms)	Mean Saccade Size (degrees)
Silent reading	225-250	2 (8-9 letter spaces)
Oral reading	275-325	1.5 (6-7 letter spaces)
Scene perception	260-330	4
Visual search	180-275	3

Table 2. Mean viewing time (in seconds) and number of fixations for the text and picture parts of ads as a function of task. Values in parentheses equal the percent of time looked at the text or picture (for the Viewing Time) and the percent of fixations in the text or picture (for the Number of fixations).

	Viewing Time		Number of Fixations	
	Text	Picture	Text	Picture
Rayner et al. (2006)	3.64 (39%)	5.72 (61%)	14.7 (39%)	22.7 (61%)
Rayner et al. (2001)				
Intended	5.61 (73%)	2.12 (27%)	25.2 (72%)	9.8 (28%)
Non-intended	3.60 (71%)	1.50 (29%)	16.4 (70%)	6.9 (30%)

Note: In the Rayner et al. (2001) study, intended refers to ads that viewers were instructed to look at to purchase whereas non-intended refers to the other ads they viewed.

Figure Captions

Figure 1. Fixation duration frequency distributions for reading, visual search, and scene perception. The data are from the same 24 observers engaged in the three different tasks. No lower cutoffs of fixation duration were used in these distributions while an upper cutoff of 1000 ms was used.

Figure 2. Examples of a moving window (with a thirteen character window), a moving mask (with a 7 character mask), and the boundary paradigm. When the reader's eye movement crosses an invisible boundary location (the letter *n*), the preview word *feeding* changes to the target word *reading*. The asterisk represents the location of the eyes in each example.

Figure 3. Examples of where viewers look in scenes. The top portion of the figure shows where one viewer fixates in the scene (the dots represent fixation points and the lines represent the sequence). The bottom portion shows where a number of different viewers fixate (with the dots representing fixation locations across a number of viewers).

Figure 1

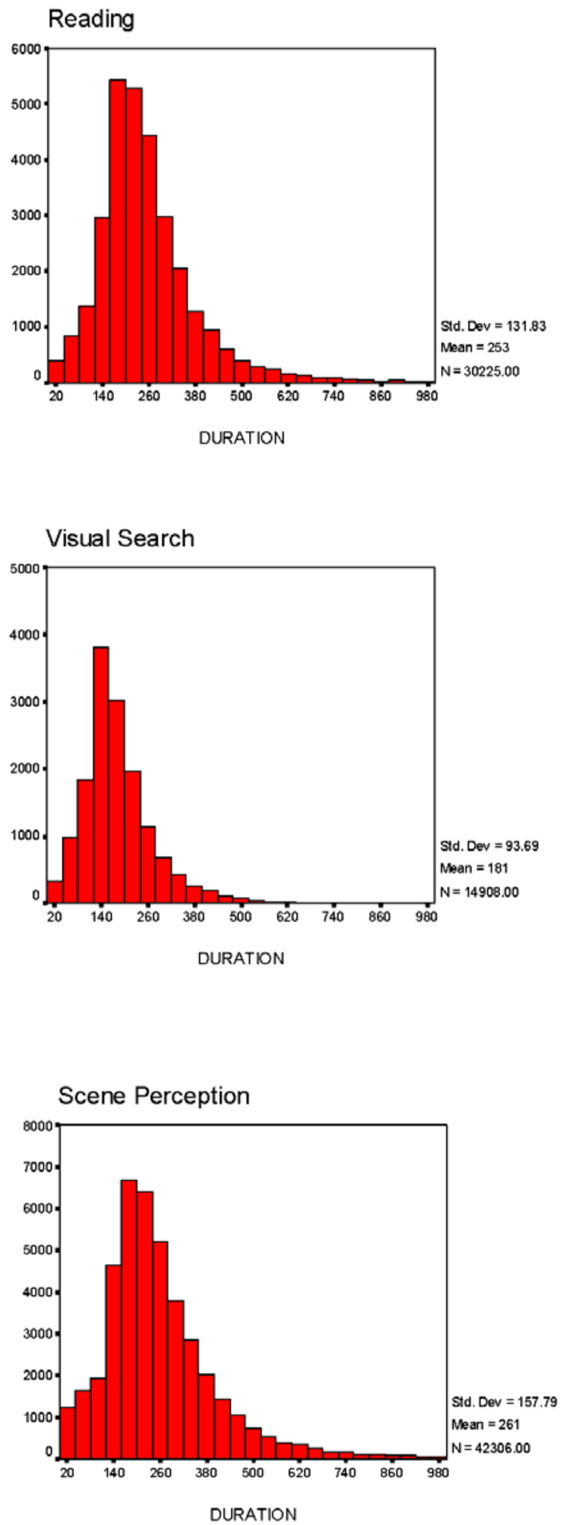
Reading
6000
5000
4000
3000
2000
1000
0
20 140 260 380 500 620 740 860 980
Std. Dev = 131.83
Mean = 253
N = 30225.00
DURATION
Visual Search
5000
4000
3000
2000
1000
0
20 140 260 380 500 620 740 860 980
Std. Dev = 93.69
Mean = 181
N = 14908.00
DURATION
Scene Perception
8000
7000
6000
5000
4000
3000
2000
1000
0
20 140 260 380 500 620 740 860 980
Std. Dev = 157.79
Mean = 261
N = 42306.00
DURATION

Figure 2

Normal Line:

```
Where do people while reading versus looking at
```

Moving Window Paradigm (13 character window):

```
Xxxxx xx xeople while rxxxxxx xxxxxx xxxxxxx xx
                *
Xxxxx xx xxxxxx xxile reading vxxxxx xxxxxxx xx
                        *
```

Moving Mask Paradigm (7 character mask):

```
Where do people wxxxxxxxading versus looking at
                    *
Where do people while rexxxxxxxersus looking at
                           *
```

Boundary Paradigm:

```
Where do people while feeding versus looking at
                  *
Where do people while reading versus looking at
                         *
```

Figure 3