CHAPTER THREE

Visual Attention and Eye Movements

James E. Hoffman
University of Delaware, USA

Visual information regarding the spatial layout and identity of objects in our environment requires several kinds of orienting mechanisms. Our eyes have parallel visual axes, providing overlap of the views in each eye, a prerequisite for binocular vision (Julesz, 1971). This arrangement results in an inability to see the approximately 180 degrees of space that lie behind our head. Thus body and head movements are required to access the entire 360 degrees available to us. Similarly, the retina of the eye is nonhomogeneous, containing a foveal area at the center which provides a small area of high-acuity form analysis (one can see a string of about eight letters in this area). Eye movements are required to bring this foveal area to bear on peripheral objects to clearly discern their shapes. These overt orienting mechanisms are, in turn, supplemented by a covert or “hidden” attention system that provides enhanced visual processing of selected areas through internal neural adjustments that can be made much faster than overt orienting (Hoffman, 1975; Posner, 1980).

This chapter reviews evidence supporting the claim that the covert attention system plays an important role in guiding overt orienting based on eye movements. In particular, it appears that eye movements directed to a location in space are preceded by a shift of visual attention to the same location and, furthermore, this coupling of eye movements and attention is mandatory. This relationship holds regardless of whether the eye movement is triggered by an external event such as a sudden movement or onset (exogenous control, see Yantis, 1996) or is “internally”
directed by the subject (endogenous control) on the basis of instructions or expectations.

The claim that eye movements are guided by attention does not entail the proposition that these two systems are completely interdependent. It has been known at least since Helmholtz’s time that one can attend to peripheral objects without making an eye movement. Helmholtz (1909) noted that he could direct his attention “at will” to different characters printed on a card that was illuminated by a brief electrical spark. The duration of the spark was too short to allow an eye movement to occur so the impression that he could “look” at a given letter must have been due to an internal attentional system being directed at a stored memory representation of the array. These impressions have been confirmed by modern work that has shown that there is indeed a visual copy of briefly presented stimuli that persists for a short time after termination of a stimulus (Sperling, 1960). In addition, a growing body of work (Eriksen & Hoffman, 1972, 1973; Hoffman & Nelson, 1981; Hoffman, Nelson, & Houch 1983; Posner, 1980; Yantis, this volume) shows that observers can direct an internal visual attention mechanism to different areas of visual space even while the eye remains fixed. Thus the relationship between attention and eye movements is one of partial interdependence. Attention is free to move independent of the eyes, but eye movements require visual attention to precede them to their goal.

ATTENTION AND SACCADIC EYE MOVEMENTS

The most heavily researched area involving attentional control of eye movements is in the area of saccades, which are rapid, ballistic changes in eye position that occur at rate of about 3–4 per second (Becker, 1991). This means that we make approximately 230,000 saccades during each waking day! The eye is essentially blind during these movements and information is acquired during the relatively long fixations (approximately 250 msec) that intervene between saccades. Saccades are important during reading and scanning of scenes which require the high-acuity form vision provided by the fovea. As Yarbus (1967) pointed out, the location and sequence of saccades is not random. Subjects scanning the same picture show highly replicable “scan paths” from one day to the next. This sequence, however, can be changed by asking subjects to report on different aspects of the picture such as people’s ages, in which case, fixations tend to be restricted to faces.

What is that guides the eye from one fixation to the next? A possibility is that sometime during the course of a fixation, visual attention is allocated to the periphery to determine the location of the next fixation. This location information, in turn, is transmitted to the neural machinery responsible for actually moving the eyes. This proposal, although plausible, is not universally accepted (see for example Klein, 1980; and Klein, Kingstone & Pontefract, 1992; Klein & Pontefract, 1994) and raises several questions regarding the precise role of visual attention in the planning and execution of saccades. First, it may appear to beg the question of how eye movements are made, by postulating an intelligent controller (visual attention) whose operation itself must then be explained, leading to an “infinite regress” of explanatory mechanisms. Luckily, we will see that this is not entirely the case because a fair amount is known about the mechanisms that guide attention and we can make use of this knowledge in understanding saccades.

Second, this proposal needs to be fleshed out with a number of details. For example, when during fixation does attention begin to “search” the periphery for candidate saccade locations and what happens to processing of fixated material during this period? What determines whether and when the eye will actually begin its movement to the new destination? And finally, how can the role of attention in this process be verified? After all, saccades can now be measured with great accuracy using modern eye-trackers but how can “covert” shifts of attention be detected? We will explore these issues in some detail by examining the control of saccades during reading.

Attention and Reading

Adult readers tend to fixate most of the “content” words (nouns, verbs, adjectives, etc.) in a text although short “function” words (articles, conjunctions, etc.) may be skipped (Just & Carpenter, 1980; Rayner & Pollatsek, 1989). For example, Crowder (1982) points out that readers make about 75 fixations in reading a 100-word text, which indicates that they are reading about 1.33 words per fixation. The duration of each fixation lasts about .25 seconds but varies widely for different words as a function of several factors such as the word’s length and frequency in the language, how well it fits into the context of the sentence, and its syntactic and thematic roles (Just & Carpenter, 1980). Just and Carpenter (1980) suggested that readers fixate each word until processing at perceptual, linguistic, and conceptual levels has been completed. According to this model, fixation durations provide a direct estimate of the total time to process each word in the text and therefore provide a useful metric of the cognitive processes involved in reading.

Preview Effects. The fact that function words located in the periphery of the currently fixated word are skipped, however, indicates that at some point during fixation, readers are obtaining information, not just from the fixated word, but also from the about-to-be-fixated word located in the periphery (also known as the “parafovea”). Subjects are apparently “previewing” words before they are fixated. When these words are short and
familiar (such as of or the), this preview is sufficient to identify the word and allow it to be skipped. However, even when subjects do not skip the word, a preview may be beneficial in reducing the subsequent fixation on the previewed word.

This predication is supported by an experiment reported by Rayner (1975). Consider a reader fixating the first word in the phrase ship's chart. During fixation of ship's, subjects may preview the upcoming word chart. This preview might provide a saving in processing during the subsequent fixation of chart. Now consider a situation in which subjects view the first word in the phrase ship's chyft. When a saccade is initiated toward the nonword chyft, the computer changes it to chart. (subjects rarely notice such changes when they occur during the time the eye is in motion). In this case, there shouldn’t be any benefits associated with parafoveal preview and, indeed, fixation durations on chart were longer relative to the case when chart was previewed.

Parafoveal preview effects have also been obtained in an experiment using pictures. Henderson, Pollatsek, and Rayner (1989) presented subjects with a series of displays containing four pictures, each one portraying a common object. Subjects were instructed to fixate each picture in the display in anticipation of a memory test. In the one object condition, pictures were presented for inspection one at a time, thus preventing any preview effects. In the one + next condition, while the subject fixated one drawing, the next drawing to be fixated was present. This preview shortened the subsequent fixation on the previewed picture. In the all objects condition, all four objects were presented simultaneously and remained on the screen during the entire trial. This condition provided no additional preview benefits suggesting that previews are restricted to that object about to be fixated.

Why do preview effects occur? One answer is suggested by examining factors that influence the landing position of saccades. Readers do not simply make a fixed-size saccade after fixating each word but instead adjust the size depending on several aspects of the material to the right of fixation (Rayner & McConkie, 1976). Readers appear to have a preferred viewing location in words, which lies between the beginning and middle of the word (O’Regan, 1992; Rayner & Pollatsek, 1989). This position will, of course, depend on the length of the word, and several studies have shown that readers increase the length of their saccades with increases in the length of the peripheral word. In addition, they rarely fixate spaces between words so that if extra spaces are inserted next to the currently fixated word, saccades will be lengthened to skip over these uninformative areas (Abrams & Zuber, 1971).

These considerations suggest that programming of the landing point for a saccade depends on locating the first nonblank character to the right of fixation (indicating the beginning of the word) and then the next blank character (indicating the end of the word). This information can then be used to estimate the length of the word and the corresponding optimal viewing position. The location information can then be passed to the mechanism responsible for programming and executing the saccade. This view suggests that preview benefits may occur simply as a byproduct of the need to pay attention to the periphery to control saccades.

A different answer as to why preview effects occur is offered by Epelboim, Booth, and Steinman (1994, 1996). They reasoned that if spaces between words are critical for guiding the eye during reading, removing spaces should disrupt eye movements and severely impair reading. They presented sentences in which the spaces between words were removed, like this one:

```
Readingsentenceswithoutspacesmayseemdifficultatfirstbutyoumay
finditsurprisinglyeasy.
```

They found that although this manipulation did slow reading speed by about 30% on average, some subjects showed no slowing at all. In addition, many characteristics of the saccades were similar with spaced and unspaced text. For example, the same subjects that showed a preferred viewing location (PVL) with normal or unspaced text, did so with unspaced text as well. Their data raise the question: If space information is not used to guide the eye to its landing position, then what is? Epelboim et al. suggest that it is recognition of the peripheral word that guides the eye. That is, while fixated on one word, subjects are also processing the peripheral word for the purpose of word recognition. Even when spaces are not present, letter strings can be recognized as words (note that spoken language does not have spaces either!) and this information in turn, may draw attention and the eye to the proper location in the word. This recognition stage is more difficult with filled spaces because errors in segregating the letters into word strings will sometimes occur and therefore reading is a bit slower, but still possible.

Are there any critical observations that point to a resolution of this debate? Rayner and Pollatsek (1996) point out that word-skipping data seem compatible with their view. Short, familiar words are often skipped during reading which is consistent with the idea that peripheral words are recognized only sometimes as opposed to being the usual state of affairs as predicted by the Epelboim et al. model. Nonetheless, the very fact that at least some readers can read normally with space information removed shows that eye guidance in reading can be based on more than just perceptual information about word length. It may be that both kinds of information are used to guide saccades during normal reading with perceptual information or word identity information being more or less useful depending on the circumstances. In any case, we should not lose sight of one principle that transcends the particular models being debated: words to the
right of the one being fixated are often identified and this fact suggests that visual attention precedes the eye to its destination.

Note that we have assumed that the preview effects observed in these experiments are due to an attentional mechanism directed to the right of the fixated word. This asymmetry to the right is presumably due to the left-to-right pattern of saccades that maps onto the direction in which our language is written. Pollatsek, Bolzky, Well, and Rayner (1981) investigated this claim using bilingual readers fluent in English and Hebrew. These readers showed the usual “asymmetric to the right” span when reading English and an opposite pattern for Hebrew, which is read from right to left. This flexibility strongly implicates visual attention as the mechanism mediating the preview effects.

Models of Saccade Control in Reading. At this point, one might agree that attention is important in controlling the eye during reading but the details of how this might work remain to be specified, and here we need a model that makes testable predictions. A good starting point is the sequential attention model of reading that was originally proposed by Morrison (1984) and expanded on by Rayner and Pollatsek (1989) and later by Henderson (1992). According to this model, fixation begins with attention allocated to the fixated word. Processing of this word continues until it has been identified and integrated into the context of the text (Just & Carpenter, 1980). Attention is then disengaged (Posner, 1980) from the fixated word and reallocated to the periphery, in order to locate the position for the upcoming saccade. This location can presumably be determined using low-level visual features such as spaces to determine the length of the next word. The engagement of attention in the periphery provides direction and amplitude parameters necessary for constructing a motor program, which when executed will move the eye to the attended location. In addition, allocation of visual attention to the periphery, will result in preview benefits for the next fixation. Finally, if the peripheral word is identified before the saccade is executed, the current saccade can be canceled and a new one programmed resulting in a longer saccade that skips the already identified word.

This model highlights several questions about the time course of attention allocation in saccade programming. First, it suggests that the allocation of attention to the periphery occurs late in fixation. This can be tested by manipulating when information necessary for programming the saccade becomes available in the periphery. If peripheral information is only accessed at the end of fixation, there should be a period of time, starting at fixation onset, in which readers are not affected by whether or not peripheral information is available. Morris, Rayner, and Pollatsek (1990) tested this prediction by having subjects read text in which words and inter-word spaces beyond the fixated word were initially filled in with Xs. Consider the following schematic displays:

| (0 ms) | Magic XXXXXXXXXXXXXXXXXXXXX | Spaces Released |
| (50 ms) | Magic XXXXXXXXXXXXXXXXXXXXX |
| (0 ms) | Magic XXXXXXXXXXXXXXXXXXXXX | Letters and spaces |
| (50 ms) | Magic Wand XXXXXXXXXXXXX | released |

In the spaces released condition, subjects initially fixated on the word “magic”. Fifty msec later, the length of the adjacent word was revealed by “releasing the space information” (accomplished by removing the appropriate Xs). The length but not the identity of the adjacent word was now available and, as pointed out earlier, this length information may be one factor used in programming the upcoming saccade. If this information is used late in the fixation, for example, later than 150 msec after fixation onset, then a 50 msec delay should not be any different than a 150 msec delay as long as the information in both cases arrives before programming is initiated. However, Morris et al. found that word length information was particularly effective when it was available during the first 50–100 msec of the fixation, although it continued to have some effect even late in fixation. This suggests that either saccade programming or the information useful for saccade programming is accessed early in fixation. According to Morris et al. (1990), information arriving near the end of the fixation is apparently used to modify an already existing program rather than building one from scratch.

In the letters and spaces released condition, both the length and identity of the adjacent word are provided during fixation. Not surprisingly, Morris et al. found a preview effect in this condition. That is, having the word “wand” present while fixating “magic” shortened the subsequent fixation on “wand”. This preview was effective only when identity information was available during the first 50 msec of the fixation on “magic”. Like the saccade length data discussed earlier, the fixation duration data suggest that it is primarily early in the fixation when the location of the upcoming saccade is determined, perhaps because some time is required for attention to be fully “engaged” on the fixated word.

In any case, we are still left with the question of the state of attention during the fixation. There are several possibilities. The location information acquired early in fixation may be stored in some way to be used later for saccade programming; saccade programming may be completed immediately, with the “execution” or “go” command coming later; or attention may be maintained on the location of the upcoming saccade until saccade programming and execution of the saccade commence at the end of fixation. This latter possibility seems a bit puzzling because we are assuming that
attention needs to be simultaneously allocated to the fixated word as well as
to the location of the upcoming saccade, requiring attention to be in two
places at once. We return to this possibility later.

Complementing the aforementioned study, which sought to determine the
temporal course of information processing of peripheral information during a
fixation, is a study by Blanchard, McConkie, Zola, and Wolverton (1984)
which examined processing of the foveal word. Their subjects read short
texts while their eye position was monitored. During selected fixations, the
word being fixated was briefly masked and then replaced by the same or a
different word. This change was made at various times after fixation in an
attempt to track the time course of identification of the fixated word. The
two possible words differed by a single letter and both fit the context. For
example, subjects might read “The underground caverns were meant to
house hidden (bombs, tombs) but then the construction stopped because of
lack of funds.” The masking stimulus prevented subjects from detecting the
letter change on the basis of apparent movement or other physical cues.
Following the sentence, subjects had to choose which of four possible words
had been present in the sentence. Two of the choices corresponded to the
two critical words.

What would the sequential attention model predict for this experiment?
Once the initial word has been on long enough to be “identified”,
attention should switch to the periphery to program the next saccade and
the subject should report only the first word. This identification time has been
estimated to be about 100 msec (Rayner et al., 1981) so we might expect that
once the word change is delayed beyond 100 msec after fixation, the subject
will simply fail to “see” the second word and will always choose the first
word. Blanchard et al. (1984) reported that when the arrival of the second
word was delayed by as much as 120 msec, subjects generally did report the
first word but also frequently reported only the second word or both words.
Remarkably, 12% of the times when the second word was reported, it had
been presented only during the last 30 msec or less of fixation. These results
are complex and open to several interpretations but they do suggest that
subjects continue to process information in the fovea throughout the
duration of fixation, even quite late when attention might be expected to be
allocated to the periphery.

So far the results do not fit easily within the sequential attention model.
Information regarding the peripheral word appears to be acquired
throughout the fixation but primarily at the beginning, not at the end, as the
model assumes. In addition, information regarding the fixated word appears
to be acquired throughout the fixation interval, even at the very end when
attention should now be firmly fixed on the location of the soon to be
executed saccade. Such results are at least suggestive that both the fixated
and peripheral words are processed in parallel and that a full commitment of
attention to the new fixation location is not made until the very moment
prior to the actual fixation location. Before entertaining specific versions of
such a model, one additional finding needs to be considered.

Henderson (1992) derived the following prediction from the serial
attention model. Suppose that the difficulty of processing the fixated
stimulus is increased (referred to as an increase in foveal load), for example, by
using a low-frequency word. This should prolong the processing of the
fixated word but what will be the effect on information acquired from the
periphery? Recall that one way to evaluate this is by looking at preview
benefit, the amount of savings one gets during a fixation by having pre-
viewed that word in the preceding fixation. According to the serial attention
model, the amount of preview benefit one gets from a peripheral word
should be independent of the difficulty of the fixated word. Increased load
should prolong the fixation but once processing of the fixated stimulus is
completed, attention will be shifted to the periphery for a constant average
time prior to the initiation of the saccade, producing a constant preview
benefit. In other words, the sequential nature of processing of fovea and
periphery prevents the difficulty of one process from influencing the
duration of the subsequent process. In contrast, if processing of peripheral
and foveal words are proceeding in parallel, and both processes are competing
for a limited “supply of attention”, it easy to see how increases in difficulty
of processing of the foveal word may call on resources that could be used to
process the peripheral word.

Henderson and Ferreira (1990) tested these predictions by having sub-
jects read sentences like the following:

1. Mary bought a chest despite the high price.
2. Mary bought a trunk despite the high price.

When subjects were fixated on the word chest or trunk, the word despite
could either be present in the periphery (same preview condition) or it could
be replaced by a random letter string (different preview condition). The
difference in fixation durations on despite between these two conditions
provides a measure of preview benefit. Notice that trunk is a lower-fre-
quency word than chest and should therefore be more difficult to identify.
As it turns out, trunk also led to a smaller preview benefit than did chest,
contrary to the predictions of the sequential attention model.

The finding that increased difficulty with the foveal word decreases
"knowledge" about the peripheral word suggests an explanation in terms of
competition for attention. For example, subjects may determine early in
fixation the location of the next saccade. A partial allocation of attention to
this location may then be established with primary attention remaining on the
foveal word. Partial allocation of attention to the periphery provides a
convenient way to maintain information about the saccade goal (similar to what Pylyshyn, 1989, 1994, calls a “FINST” mechanism for spatial indexing). It also accounts for why information about the peripheral and foveal words are both available throughout fixation. On those occasions when peripheral attention is adequate to identify the peripheral word (a likely occurrence for short, high-frequency words), attention may be reallocated to the following word, resulting in word skipping. Finally, the amount of attention available for processing of peripheral information would depend on the difficulty of the foveal word and hence the amount of attention it required.

Henderson (1992) pointed out that postulating divided attention introduces its own set of difficulties. At the time he wrote his chapter, there was little in the way of convincing evidence that subjects can divide their attention between two separated spatial locations and this continues to be an area of controversy in attention research. Pylyshyn (1989, 1994) showed that subjects can track multiple objects simultaneously without including areas between the tracked objects. Yantis (1992) provided evidence that this occurs only when subjects see the dots as the vertices of a single object moving on the screen. Kramer and Hahn (1995) however, provided direct evidence for divided visual attention in a paradigm requiring subjects to simultaneously compare two shapes in different locations. Subjects were able to do this while excluding irrelevant shapes located between the two comparison shapes. So the possibility of divided spatial attention cannot be excluded.

Second, in the sequential allocation model, the eye movement was triggered by a clearly defined criterion, namely identification of the foveal word. The parallel model would need some sort of control structure spelled out so that eye movements could be triggered at appropriate times. Third, the parallel model would predict that increasing difficulty of peripheral processing should affect processing of the foveal word and this apparently does not occur (Henderson, 1992). This latter difficulty can be handled by simply assuming that it is the “primary” locus of attention that determines how attention is allocated among competing inputs.

**Summary.** What can we conclude about the role of attention in guiding the eye during reading? One clear finding that has emerged from this line of work is that during fixation, readers attend to and acquire information about the word lying to the right of the currently fixated word. This information includes length, shape, and sometimes identity. These characteristics affect the subsequent saccade but exactly how and when attention plays a role is still a matter of speculation. Evidence favors the idea that attention is divided between the fixated and peripheral words throughout fixation but much more research will be required to determine if this is the case. We now turn to the question of whether the coupling between attention and saccades that we have observed in reading is one of convenience or necessity.

**Can Attention and Saccades be Directed to Different Locations?**

_Evidence Favoring the Dissociation of Attention and Saccades._ The research reviewed earlier suggests that attention normally precedes the eye to its destination. However, the possibility remains that this relation is one of convenience and that under suitable conditions, subjects could be induced to send their eyes and their spatial attention to different locations. This possibility has been investigated by several different researchers using some variant of a dual-task paradigm. Subjects are given two tasks: they are to move their eyes to one location as quickly as possible and, in addition, try to detect or identify a visual target presented in close temporal proximity to the saccade. The target can occur in the same location as the saccade goal or at varying distances. Accuracy on the target detection task serves as an indirect measure of attention allocation because accuracy and speed of target detection and identification is generally superior at the locus of attention and drops with increasing distance (Eriksen & Hoffman, 1972, 1973; Hoffman & Nelson, 1981; Yantis, this volume). If saccades and attention can be dissociated in space, then subjects should be able to attend to one location (verified by superior target performance at that location) while making a saccade to another.

Klein (1980) reported one of the first attempts to experimentally dissociate movements of the eyes and attention. We will extensively review one of his experiments, both because his results are often cited as showing that this dissociation is possible and because it will make it clear that subjects’ strategies in the dual-task paradigm play a powerful role in the pattern of results. Klein proposed a particular version of the relation between attention and eye movements known as the _oculomotor readiness theory_ which holds that movements of attention are mediated by activity in brain areas that are responsible for moving the eyes. For example, attention to a particular location is accomplished by constructing a program suitable for moving the eyes to that location. This saccade program, in turn, produces the enhancement effects at the attended location. This theory makes two predictions which are the basis of his experiments. First, subjects who are preparing to make a saccade to a location should also be better at detecting signals at that location relative to other locations in the visual field. Second, attention to a location should lead to preparation of a corresponding saccade program resulting in fast saccades to that location.

The attentional enhancement predicted to accompany saccades was examined in Klein’s first experiment. Subjects faced a computer screen and were told that on each trial, they would have to perform one of two tasks. In the saccade task, they were to move their eyes in a particular direction (for example, right) if an asterisk was presented on the screen. In the detection
task, they were to release a response key if a dot brightened. The display consisted of three evenly spaced dots, one in the center and the other two falling 8° on either side of the center. After trial initiation, one of three events could occur: (1) the left or right dot brightened for 100 msec, (2) an asterisk replaced either the left or right dot, or (3) no change occurred, which constituted a catch trial. Note that this is not strictly a dual-task paradigm because subjects make only one response per trial. However, the important point is that they must prepare for both tasks because they do not know which one they will be performing until the display appears.

Consider the predictions of the oculomotor readiness theory. Suppose that the subject is instructed to favor the saccade task over the manual task, and is told to move to the right as soon as the asterisk is detected (remember that the asterisk can occur on either the left or right and, in both cases, signals a saccade to the right). They should be attending to the right before the trial begins because there is an oculomotor readiness to move to the right and this produces allocation of attention in this direction. Therefore, detection of the brightening of the dot (reflected in manual RTs) should be faster when it occurs on the right side.

Contrary to the oculomotor readiness theory, manual detection latencies were the same regardless of their relationship to the direction of the saccade. Subjects were no faster at detecting a signal on the right side of the display when they were preparing a saccade in that direction than when they were instructed to saccade to the left. Single-task control conditions verified that when subjects were instructed to attend to the left or right in the absence of a saccade task, they showed the usual attentional effects found previously in this paradigm (Posner, 1980), verifying that the detection task was sensitive to the direction of attention.

Although this result seems to offer a clear and simple test of the theory, there are several reasons to be cautious in accepting the conclusions. First, there were large delays in the speed of both responses in dual-task conditions compared to the single-task controls (approximately 90 msec for the saccade task and 125 msec for the manual task). If subjects were primarily preparing the saccade task, it should have been delayed only a small amount or not at all (see Kowler, Anderson, Dosher, & Blaser, 1995, and Hoffman & Subramaniam, 1995 reviewed later in this section for evidence supporting this assertion). But notice that in the dual-task blocks, unlike the controls, one of two kinds of peripheral visual events could occur unpredictably: a brightening of the dot or an onset of the asterisk. Subjects may well have had to attend to the peripheral event in order to determine which signal had occurred. In single-task conditions, simple detection of any visual event in the periphery was enough to initiate a response. Therefore, it is possible that subjects were not prepared to make a saccade in a particular direction in dual-task trials but chose a strategy of first determining which task was required (manual or saccadic) before programming the appropriate response. Without this saccade preparation, the predictions of the oculomotor readiness theory cannot be tested.

Second, the use of a peripheral visual event (the onset of the asterisk) to signal the subject to make a saccade would be expected to automatically attract attention to that location (Yantis, 1996). Therefore, even if subjects were preparing a saccade to the left location, and were attending to that location as part of the saccade preparation process, the onset of an asterisk on the right would essentially abolish this attentional set by reorienting attention to the right. The same difficulty holds for the second experiment which also failed to find any link between attention and saccades. Other investigators (e.g., Crawford & Mueller, 1992; Remington, 1980) have also used peripheral visual signals to indicate the direction of saccades. Even in cases where target detection does appear to be enhanced at the saccade goal, one can not conclude that attention normally precedes saccades. Such an enhancement effect of a peripheral visual signal would be expected even if subjects were not making a saccade and therefore is not relevant to the question of whether attention and saccades can be dissociated.

Evidence for a Role of Attention in Saccades. The preceding analysis suggests several methodological requirements for research addressing this issue. First, it is probably best to signal the initiation of a saccade with a nonvisual signal. Visual signals will tend to draw attention automatically (see Yantis, this volume) and may disrupt the attentional allocation established by instructions. Second, subjects may adopt various strategies for combining the two tasks and this possibility needs to be assessed.

Shepard, Findlay, and Hickey (1986) used an improved methodology to study this issue. They used a central arrow cue pointing to a box on the left or right of fixation to indicate the target of the saccade, avoiding the capture effects associated with peripheral cues. Attention was manipulated by varying the probability of a target occurring in the left or right box. Thus, subjects could be instructed to move their eyes to the left while target
probabilities favored attending to the right. In these conflict situations, subjects detected targets more quickly when they occurred in the saccade target location, not the position favored by the probability manipulation. When saccades were not required, the probability manipulation had the expected effect of speeding responses to signals on the probable side. Shepard et al. concluded that making a saccade requires that attention be allocated to the saccade target location. Although this experiment is an improvement over previous methods, it too has a flaw that suggests caution. The target was left on the screen until subjects responded, so that targets located in the direction of the saccade were foveated. Because average saccadic latencies were shorter than manual RTs, at least part of the advantage for targets appearing in the direction of a saccade could have been due to faster processing of foveal signals compared to peripheral signals.

Hoffman and Subramaniam (1995) used the central cueing procedure of Shepard et al. together with short target durations to eliminate the target foveation problem. Their method is shown in Fig. 3.1. On dual-task trials, subjects were instructed to make a saccade to one of the four fixation boxes as soon as they heard a tone. The direction of the saccade was constant for the entire session. The saccade task was emphasized and subjects were encouraged to achieve saccade latencies comparable to those obtained in a control condition in which this task was performed alone. Subjects were able to do this, indicating that they were optimally prepared to make a saccade in the instructed direction at the beginning of each trial. Prior to each trial, they were also shown a central arrow pointing to one of the four boxes which indicated the likely position of a target letter. One of two targets (L or T) could occur and nontarget positions were filled with distractors (an E or an F). Subjects made their choice as to which target had been presented after the saccade. Target presentation was brief (14 msec) so that the letter display was gone by the time subjects made their eye movement.

The data are shown in Table 3.1, separated according to whether the target occurred in the location indicated by the cue (cue–target match) or the location corresponding to the saccade goal (saccade–target match). Also shown are data for a control condition in which subjects performed the detection task by itself. In the detection-only block, target detection was best in the cue–target match condition, indicating that subjects used the arrow to attend to the cued position with a resulting enhancement in target discrimination. In the dual-task condition, subjects sometimes faced a conflict. The arrow cue directed them to allocate their attention to one location while their saccades were directed to another. If visuospatial attention is required to execute a saccade, targets should be discriminated best when they occur at the location of the saccade goal and the arrow cue should be ineffective in directing attention. The data in Table 3.1 show this pattern. Targets were discriminated 13% better when they occurred in a location to which subjects were preparing to move their eyes. In contrast, the attention arrow was ineffective even though it continued to provide valid information about the likely location of the target (Schneider & Duebel, 1995 report similar results).

FIG. 3.1. Procedure used by Hoffman and Subramaniam (1995) in Experiment 2. Subjects were to make an eye movement to a specified peripheral box on each trial. After trial initiation, they were presented with a central arrow cue (duration = 1000 msec) indicating the likely position of a target (L or T). After a variable warning interval (500–2000 msec) a tone was presented as a signal to initiate the saccade. 20 msec later a letter array was presented containing the target and three distractors. After completing the eye movement, subjects have to indicate which target had been presented. (From “Saccadic eye movements and visual selective attention” by J. E. Hoffman & B. Subramaniam, 1995, Perception and Psychophysics, 57, 787–795.) Reprinted with the kind permission of the Psychonomic Society, Inc.

<table>
<thead>
<tr>
<th></th>
<th>Detection-only Block</th>
<th>Saccade–Target Match</th>
<th>Saccade–Target Mismatch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cue–Target Match</td>
<td>79.3</td>
<td>86.6</td>
<td>73.4</td>
</tr>
<tr>
<td>Cue–Target Mismatch</td>
<td>66.93</td>
<td>86.4</td>
<td>70.4</td>
</tr>
</tbody>
</table>

Percent correct target discrimination in the detection-only and dual-task blocks. On saccade–target match trials, the target was located at the goal position for the saccade. On cue–target match trials, the target was located at the position indicated by the central arrow cue. (From “Saccadic eye movements and visual selective attention” by J. E. Hoffman & B. Subramaniam, 1995, Perception and Psychophysics, 57, 787–795.)
Subjects did not simply ignore the arrow cue indicating the likely target location. Saccade latencies were slightly shorter when the directions of attention and saccade agreed than when they disagreed, suggesting that subjects were allocating some attention to the arrow cue. This experiment probably represents one point on a trade-off function relating saccade latency and target discrimination performance. Emphasis on the saccade task resulted in attentional allocation that was dominated by the location of the saccade goal, and saccades were only slightly delayed when there was a conflict between the two tasks. Increasing attention to the target discrimination task should be accompanied by increases in performance on that task at a cost of slower eye movements.

Kowler et al. (1995) have recently evaluated the nature of this trade-off function in terms of the attention operating characteristic or AOC (Sperling & Dosher, 1986; Sperling & Melchner, 1978). In their experiment, subjects were presented with a circular array of eight letters and a simultaneous central arrow cue pointing to one of the eight locations. The arrow cue indicated the goal of the saccade that was to be initiated at display onset. Throughout the block of trials subjects were to report the identity of the letter occurring at a fixed position in the display. The saccade arrow sometimes agreed with this location, but on conflict trials it indicated a different position. In the random saccade condition, the cue indicating the direction of the saccade changed from trial to trial. In the fixed saccade condition, the arrow pointed to the same location for all trials in a block. In the latter condition, subjects knew both the locations of the letter target and saccade goal prior to target onset. If it is possible to attend to one location and make an eye movement to another, this condition should provide the optimal conditions for its occurrence.

In different blocks of trials, subjects were given instructions emphasizing either the saccade or letter task. For example, in one condition, subjects were instructed to emphasize the letter task, prolonging the saccade latency only as much as needed to achieve perfect identification performance. In another condition, saccade latencies were to be kept as short as possible. In a third condition, both tasks were to be performed together with equal emphasis. Performance of each task was also evaluated in single-task control conditions. Evaluating performance in each condition as task emphasis is changed should reveal the nature of the trade-off that exists when subjects make a saccade to one location while trying to attend to another.

The results are shown in Fig. 3.2, which graphically shows the trade-off in performance between the two tasks. The dotted rectangles provide a reference for evaluating dual-task interference. Points located on the axes represent single-task performance levels. For example, points on the y axis represent performance in the letter identification task when subjects are not required to make saccades. Similarly, the open rectangle on the x axis represents the average saccade latency when letter identification was not required. The intersection of these points (upper right corner of graph) represents the “independence point”; the expected dual-task performance if both tasks can be performed together as well as each task is performed alone. If the two tasks are incompatible, we should see dual-task interference, in which performance on one or both tasks is reduced compared to their single-task levels. This would be reflected in dual-task data that lie “inside” the rectangle defined by the single-task performance levels. Each condition in the experiment (fixed or random) produces a set of points that constitute an AOC function.

Several notable features of the AOCs can be seen in Fig. 3.2. For both subjects, the AOCs in both the random and fixed conditions lie inside the rectangle defining “interference-free” dual-task performance. For example, consider the random condition for subject EK (solid circles). The upper left point represents performance when the perceptual task (letter identification) was emphasized, and indeed accuracy is close to that achieved when this was the sole task. The saccade latency here is about 265 msec. In contrast, when the saccade task is emphasized, letter performance drops by 43% and saccade latency is improved by about 30 msec. This trade-off holds for both conditions.
(random and fixed) for both subjects. The nature of this trade-off is illuminated by two additional observations. In separate blocks, subjects were given a “final opportunity” to try to perform both tasks together as well as they were performed separately. Subjects were encouraged to try to combine both tasks to achieve performance comparable to single-task conditions. The upward-pointing triangle represents performance in this block when the letter and saccade goal were in the same location (“look, attend same”). The downward-pointing triangle shows corresponding performance when these locations were different (“look, attend different”). The “same” condition leads to joint performance close to the independence point. When different locations were to be attended, performance of one or the other task suffered. This shows that dual-task interference is not due to some general inability to do both tasks together, because subjects were quite good at combining tasks when attention and saccades were to be directed to the same location. The difficulty appears to be due to competition for a specific limited resource, a spatial attention mechanism that can be allocated to one location at a time.

One additional aspect of the AOCs shown in Fig. 3.2 is striking. Notice that as the emphasis is shifted from favoring the saccade task to equal emphasis on both tasks, there is a substantial improvement in letter identification with little or no cost in terms of saccade latency (the AOCs in this region appear as vertical lines). In other words, there are diminishing returns of allocating attention to the saccade task. Some attention is beneficial but more does not help. In contrast, the letter task seems quite sensitive to attention, with large changes in performance accompanying each change in task emphasis. Kowler et al. account for these features in terms of a temporal model of the effects of attention on saccades. Suppose that there is a critical period, late in the fixation in which attention plays an important role in saccade generation. For example, attention may have to be switched to the saccade goal just before a “go” signal is issued to initiate saccade execution. Attending to the saccade goal before this critical period does not produce any benefits in terms of saccade latency, and reduces the discriminability of information at other locations. When subjects are trying to do both tasks together (equal emphasis condition), a good strategy would be to keep attention on the relevant letter location until the critical period for saccade initiation arrives, at which time attention would be switched to the saccade goal. Increasing emphasis on the saccade task might cause subjects to start attending to the saccade goal prior to the critical period, with a resulting loss in letter identification accuracy and little improvement in saccade latency.

Summary. Let us pause at this point and try to draw some conclusions about the relationship between visual attention and saccadic eye movements. The section on attention and reading pointed to two important conclusions. First, saccades appear to be guided by information in the periphery of the fixated word. Either the word’s physical characteristics, its identity, or both help the reader program the saccade to fall on an optimal viewing position in the word. Second, the data were suggestive that readers divide their attention between the fixation and the saccade goal and consequently, continue to process information about the fixated and peripheral words throughout fixation.

Research with the dual-task paradigm shows that the link between attention and saccades is mandatory. Attention must be allocated to the saccade goal at some point prior to saccade execution. The time course of attentional allocation is not known but Kowler et al. (1995) offer an interesting speculation. They suggest that attention must be shifted to the saccade goal during a critical time period that may occur late in fixation. In addition, they speculate that this attention shift will initiate a saccade only if it is accompanied by a “go signal”; otherwise the eye is inhibited from moving.

If this go signal can be set to trigger automatically with shifts of attention, one might account for the smooth and effortless coupling of attention and saccades that seems to characterize many ordinary activities such as reading.

This model is similar to the divided attention model in that both models propose that a shift of attention to the saccade goal is one of the critical events that triggers the execution of a saccade. Each model also faces a difficulty. In the Kowler et al. model, how does attention know where to shift? In other words, what mechanism finds the critical location for the upcoming saccade and preserves that information until saccade execution? This question is even more puzzling in the case of reading, where there is no attention cue to indicate the landing point of the eye; instead this must be gleaned from various aspects of the peripheral word, such as its identity and length. Is this information delivered by “preattentive mechanisms” or is attention allocated to the periphery for this purpose? The parallel model proposes that attention is divided between the fixation point and the periphery throughout the fixation interval. Saccade initiation would correspond to a rapid shift of attention from fixation to the peripheral location being indexed by a partial allocation of attention. The problem here is that only certain shifts of attention should produce saccades and it is clear that subjects can shift attention to different locations in the visual field without an accompanying saccade (Yantis, this volume).

Clearly, the parallel model requires an additional mechanism that acts as a switch, allowing a shift of attention to produce a saccade in one case but not another, similar to the “go signal” proposed by Kowler et al. A possible candidate for this mechanism is to be found in the idea of “attentional disengagement” (Posner, 1995). Posner (1995) suggests that before attention can be switched to a new location it must be actively disengaged from its current location. It may be that subjects maintain fixation by engaging attention on the fixated object. The swift “all or none” allocation of
attention that triggers saccades may be possible only when attention has been disengaged from its current focus. Interestingly, the phenomenon of *express saccades* offers direct support for this conjecture.

**Express Saccades**

Express saccades have a mean latency in the range of 100 msec compared to about 225 msec for "normal saccades" (Fischer & Weber, 1993). Express saccades are found in the so-called *gap paradigm* (Saslow, 1967) in which the fixation stimulus is turned off at various times prior to the appearance of a peripheral target which serves as the goal for the saccade. According to Fischer and Weber, the termination of the fixation point allows subjects to disengage attention before the appearance of the target. Once attention is disengaged, the appearance of the peripheral target produces a shift of attention and an express saccade. It is the elimination of the "disengage operation" from the normal saccadic RT which results in the short latency characteristic of express saccades. Fischer and Weber (1993) review a large number of findings supporting the attentional disengagement theory. Their view has, however, generated a vigorous critique. Some investigators have failed to observe express saccades in the gap paradigm (Kingstone & Klein, 1993a). Others have suggested that removing the fixation point provides a warning signal that speeds RTs according to well-known principles governing all speeded reactions (Reuter-Lorenz, Hughes, & Fendrich, 1991). Still others have accepted that the gap paradigm does result in faster RTs over and above warning effects but have pointed out that there is little or no evidence of a role for attention in producing these effects (Kingstone & Klein, 1993b).

The research presented earlier showing a relationship between attention and saccades is consistent with the occurrence of the gap effect and the interpretation in terms of attentional disengagement. The notion that disengagement must precede a change in the direction of attention is an important component of Posner's (1980) theory of orienting. He suggests that a peripheral signal triggers a disengagement of attention from a fixation stimulus, movement of attention to the new stimulus, and a subsequent engagement at the new location. Duncan, Ward, and Shapiro (1994) have recently shown that the time to disengage attention (what they call "dwell time") can be surprisingly long, in some of their experiments as long as 500 msec. (see Hoffman, 1978, 1979 for similar arguments).

The preceding arguments suggest that time to disengage attention from fixation is at least a reasonable explanation for the reduction in saccadic latency that is associated with removal of fixation prior to onset of the target. What is needed is some direct evidence that these same conditions also result in rapid shifts of attention to the target, otherwise we cannot be sure that attention is involved in the gap effect. Mackeben and Nakayama (1993) provided such a demonstration. Their paradigm is shown in Fig. 3.3. Subjects viewed a central fixation point which in the gap condition went off 200 msec prior to the appearance of a circular cue in the periphery. The cue

![Fig. 3.3](image-url)
was followed at variable intervals by a display containing a target in the cued position together with 17 distractors. The target was a line with a vernier offset while the distractors were vertical lines. The subjects' task was to determine the direction of offset in the target. The target display was followed by a mask and the measure of interest was discrimination accuracy. In the no-gap condition, the fixation stimulus remained on throughout the trial.

Results are shown in Fig. 3.4. It can be seen that providing a gap before the onset of the cue improved performance with short cue–target SOAs, presumably because attention was in a “disengaged state” and could be quickly captured by the cue. Additional experiments varied the gap duration while holding cue–target SOA constant and found that the optimal gap was approximately 200 msec, which agrees with results by Fischer and colleagues for the optimal express saccade gap. In addition, a variety of control conditions ruled out the possibility that the gap effect was due to warning effects. This is a fairly direct demonstration that a gap has the predicted effect on the visual attention system. A more direct approach would be to try to measure attention shifts in the gap paradigm when subjects are actually making saccades, much as in the dual-task experiments reviewed earlier. In the meantime, debate continues on the occurrence and meaning of express saccades (see Kingstone & Klein, 1993a, 1993b; Reuter-Lorenz et al., 1991; Tam & Stelmach, 1993)

![Diagram](image)

**FIG. 3.4.** Results from Experiment 1 of Mackeben and Nakayama (1993). Percent correct discrimination of the target as a function of cue lead time for the gap and no-gap conditions. (From “Express attentional shifts” by M. Mackeben & K. Nakayama, 1993, *Vision Research, 33*, 87.) Reprinted with kind permission of Elsevier Science Ltd, The Boulevard, Langford Lane, Kidlington OXS 1GB, UK.

3. ATTENTION AND EYE MOVEMENTS

The Premotor Theory of Attention

The research reviewed to this point shows clearly that spatial attention is a critical component of the programming and execution of saccadic eye movements. There are many possible mechanisms by which this link could be achieved. For example, there may be a “general” spatial attention mechanism that is capable of indexing locations in visual space. This mechanism can be used to provide coordinates to any system that needs spatial location information including perception, reaching, locomotion, eye movements, etc. This model is consistent with the finding that there are cells in the parietal cortex that increase their firing rate when an animal attends to a location independent of the task being carried out (Wurtz, Goldberg, & Robinson, 1980). In this model, there is no special relationship between visual attention and the eye movement system.

An alternative view is illustrated by the oculomotor readiness hypothesis proposed by Klein (1980) and a related theory, the *premotor theory of attention* proposed by Rizzolatti et al. (e.g. Rizzolatti, Riggo, Dascala, & Umita, 1987; Rizzolatti, Riggio, & Sheliga, 1994). According to these theories, there is no response-independent representation of space being activated by an attentional mechanism. Instead there are many different representations of space, each responsible for particular motor actions such as reaching, eye movements, etc. Attending involves activating motor routines in the area appropriate for the response system being employed.

Because of the close association between visual attention and saccades (we usually look at the object of our attention) visual attention should be strongly related to activity in those brain mechanisms involved in eye movements. In particular, these theories hold that shifts of attention are accomplished by preparing an eye movement program to execute a saccade to the attended location via activity in oculomotor areas such as the superior colliculus (SC). The SC is a subcortical brain area that has both visual and motor functions and is particularly important in generating saccades. For example, the SC has indirect outputs to the motor areas that actually move the eyes and stimulation of SC cells can produce saccades. It also has cells that are visual rather than motor. These cells receive inputs from the retina along a pathway separate from the set of neurons that proceed from the retina through the geniculate to the visual cortex. In addition, the SC and several areas of cortex are heavily interconnected, allowing for two-way communication.

This anatomical arrangement suggests the possibility that the visual cells in the SC could perform certain visual functions, such as detecting the sudden onset of a stimulus in the visual field, and directly program a saccade to that location. In cases where saccades are generated by an intention to move the eyes to a location without any abrupt onsets, cortical mechanisms
could send the necessary information to the SC so that, once again, an appropriate saccade program could be compiled. Finally, when subjects are attending to a location in the absence of a saccade, information may still be sent to the SC, resulting in a saccade program but, in this case, the program is inhibited from being executed. The preparation of this program, however, enhances the sensitivity of purely visual cells in the SC, perhaps by feedback from the motor neurons to the visual cells. Presumably, it is the enhanced response of these visual cells that produces enhanced detection and identification of attended objects, although this enhancement would also have to be fed back to the visual cortex, because SC cells would not, by themselves, be able to accomplish processes such as object recognition, which are affected by visual attention.

As we have seen, recent experiments (Hoffman & Subramaniam, 1995; Kowler et al., 1995) have shown that the two predictions derived from the premotor hypothesis are supported: saccades to a location entail attending to that location, and attention to an area speeds saccades to that area. It was not necessary, however, to assume that attending without saccade generation involved any activity specifically connected with saccades (although this is certainly a possibility). This suggests that one needs to look for other aspects of behavior that might be uniquely related to the oculomotor system and see whether these hold true of covert orienting as well. The meridional effect has been proposed as one such signature of oculomotor activity (Rizzolatti et al., 1987).

Rizzolatti et al. (1987) presented subjects with a linear string of four boxes (oriented either vertically or horizontally) so that two boxes were located in each visual field. A central cue pointed to one of the boxes, indicating the likely position of a target flash to be detected. Subjects were to press a key as soon as they detected the appearance of the target regardless of its location. The focus of the analysis was on the costs associated with invalid trials, that is trials in which the target was presented in an uncued box. Rizzolatti et al. found that costs were greater when the invalid position was in the opposite hemifield from the cue, compared to the same-hemifield case, even when distance between attended and unattended locations was held constant across the two conditions. According to the authors, target detection was especially slow when attention had to cross either the vertical or horizontal meridians to locate the target. Note that the meridians separate the visual field into left vs right and top vs bottom. The muscles that move the eyes are also arranged in pairs so that one set moves the eyes left and an opposing set moves them right. An oculomotor program needs to specify both the direction (left vs right, for example) and amplitude of the movement (how far left or right).

Consider the case in which subjects are cued to attend to a location in the left visual field. Attending is accomplished by compiling a saccade program specifying the direction and amplitude parameters required to actually move the eyes to the attended location. On invalid trials, in which the target does not occur at the cued position, this saccade program must be modified so that its parameters reflect the new location. In the case of invalidly cued targets occurring in the right visual field, the direction parameter must be changed, whereas in the case of same-hemifield targets, it is the amplitude that must be modified. There is evidence that saccades are programmed in a hierarchical fashion, with direction being specified first, followed by an amplitude parameter (Becker & Jurgens, 1979). Once programmed, changes in direction require a new program to be constructed from scratch, whereas a change in amplitude requires only a change in a parameter. In the Rizzolatti et al. (1987) study, invalid locations in the same visual field would involve changes in amplitude (which might be accomplished quickly) whereas crossing a meridian involves a change in direction and hence a change in the entire program (which might take longer than a change in a parameter value). This scheme provides a clever explanation of the meridional effect based on changes in saccade programming.

One additional finding regarding the meridional effect is relevant. It is only observed with exogenous cues; costs associated with exogenous cues do not show a meridional effect (Reuter-Lorenz & Fendrich, 1992). According to Rizzolatti et al. (1995), this is compatible with their theory. Exogenous cues may exert their effects by directly stimulating cells in the SC and causing a saccade program to be set up in a “holistic” manner; that is, the exogenous cue can program a saccade to its location without separately specifying direction and amplitude. Therefore, the costs associated with modifying such a program (i.e., constructing a new one directed to the location of the invalidly cued target) should be independent of same vs different hemifields. Although this explanation might seem post hoc and convenient, there is some evidence that exogenously cued saccades are programmed in a “holistic” fashion compared to the hierarchical specification of direction and amplitude associated with endogenously cued saccades (Abrams & Jonides, 1988).

Reuter-Lorenz and Fendrich (1992) provided an alternative explanation of the meridional effect and its dependence on exogenous cues. They pointed out that exogenous cues tend to produce fairly narrow attentional fields compared to endogenous cues (Shepard & Mueller, 1989). This makes some sense in that exogenous cues specify the attended area “directly” whereas endogenous cues (such as an arrow pointing to another location), require the observer to estimate where attention should be allocated (Logan, 1995). This estimation process would be expected to be approximate and error-prone. Therefore central cues would tend to activate large areas and, if the size of these areas was limited by anatomical divisions such as horizontal and vertical meridians (Hughes & Zimba, 1985; but see Egly & Homma, 1991).
for qualifications on this result), the meridian effect would be obtained for central or endogenous cues. Notice that in this explanation, the costs associated with invalid cues are not due to realignments of attention, but rather the amount of enhancement that unattended signals receive. Exogenous cues, which produce a tight focus of attention on the cued area, will not produce enhancement of large areas and therefore will not lead to meridian effects.

Summary. Findings reviewed earlier showing that saccades are preceded by shifts of attention are compatible with predictions of the premotor theory. Such predictions are not unique to this theory; they are also compatible with the assumption of a central attentional control mechanism that sends spatial coordinates to the saccade generation system. One does not need to add the further assumption that shifts of attention in the absence of saccades necessarily involved the saccade system. A similar situation holds for phenomena such as the meridian effect which can be accounted for by both theories.

More direct tests of the premotor theory would examine whether damage to the SC results in impaired covert orienting as well as impaired eye movements. Rafal et al. (1988) investigated this possibility by examining patients with progressive supranuclear palsy, a degenerative disorder that attacks the superior colliculus (SC), among other structures. These patients displayed a deficit in making voluntary saccades, particularly in the vertical direction. Rafal et al. found a similar deficit in shifting attention in the vertical direction in response to exogenous cues, suggesting a common role for the SC in exogenous orienting and saccades. Rafal, Smith, Krantz, Cohen, and Brennan (1990) studied a group of hemianopic patients who were functionally blind in one hemifield due to destruction of part of their striate cortex. The retinotectal pathway leading to the SC is intact in these patients and we might expect that orienting to exogenous cues might also be intact. Rafal et al. demonstrated that this was the case by presenting a saccade target in their good field together with a visual signal in the blind field. Although these subjects were unaware of the signal in their blind field, it slowed their saccade to the signal in their good field, probably because it competed with the target for exogenously triggered attention. These two experiments together make a strong case for the proposition that the geniculostriate pathway may play a key role in mediating awareness of visual stimuli while the older retinotectal pathway, which includes the SC, plays a role in orienting to exogenous signals. Desimone, Wessinger, Thomas, and Schneider (1989) directly tested this proposition by showing that focal lesions in the SC of monkeys produced deficits in covert orienting to stimuli occurring in the receptive fields of the lesioned cells (see also Gattass & Desimone, 1992).

This suggests that, as the premotor readiness hypothesis predicts (Rizzolotti et al., 1987), brain structures directly concerned with saccade generation may also be important in some aspects of covert orienting in the absence of eye movements. However, this appears to be true only for exogenous cuing. There is no evidence that endogenously directed attention depends on the SC. The claim that the SC is involved in orienting to exogenous cues but not endogenous cues is supported by the phenomenon of inhibition of return which we review in the next section.

Inhibition of Return

Exogenous cues produce an initial enhancement of information at the cued location, followed a few hundred msec later by a period of inhibition in which detection of visual signals is impaired. Posner and Cohen (1984) called this phenomenon inhibition of return (IOR). They suggested that attention is initially drawn to the location of the exogenous cue and then withdrawn after the contents of the location have been determined. It is this withdrawal of attention that generates IOR and it represents a bias against attending to the same location more than once within some time period. Such a mechanism would be useful in visual search, for example, where having searched a location, one should avoid immediately searching it again.

The inhibition that follows withdrawal of attention from a location increases the latency of both manual and saccade responses. A close link between IOR and saccade programming is suggested by the finding that IOR occurs following exogenous cues regardless of the task; endogenous cues produce IOR only when the subject prepares or executes a saccade to the attended location (Rafal, Calabresi, Brennan, & Scioli, 1989). This difference in susceptibility to IOR has been taken as evidence that exogenous and endogenous attention may be mediated by different mechanisms (Klein et al. 1992). The findings are consistent with the claim that exogenous cues automatically initiate saccade programming in the superior colliculus. Endogenous orienting does not invariably produce saccade programming but can when this is the appropriate response.

What is the basis for IOR? Abrams and Dobkin (1994) considered two different loci for the IOR effect. The slowing of signal detection at previously attended locations could be the result of inhibition of early perceptual processing at the previously attended position, or it could be due to effects on the response system. They separated these different contributions with the following method. The subject’s attention was attracted to a peripheral location by a brief flash and then it was drawn back to fixation by another transient. Withdrawal of attention from the peripheral location should have produced an IOR for that location. This was tested by requiring
subjects to make a saccade, either to the previously cued position or to a new one. The direction of the saccade was indicated either by an arrow at fixation (endogenous cue) or a peripheral flash (exogenous cue). The endogenous cue does not involve presenting a signal at the peripheral location, so any inhibition found here would be attributed to the response system rather than perceptual processes. In contrast, the exogenous cue involved both perceiving the peripheral signal as well as responding. They found a small IOR for the endogenous cue, consistent with a role for purely response level processes, and a larger effect for the exogenous cue, indicating an additional contribution to IOR by factors important for signal detection.

There is also evidence that part of the IOR is defined in "object-based" coordinates rather than retinotopic or environment-based coordinates. Tipper, Driver, and Weaver (1991) had subjects view two boxes on either side of a fixation box. Attention was drawn to one box by an exogenous cue (a brightening of the box) and drawn back to fixation by another transient. The two peripheral boxes were then rotated about the fixation point by 180°. A signal appeared randomly in one of the two boxes and subjects had to press a key as soon as they detected it. They were slower in detecting signals in the box that had been cued even though the uncued box was at the same retinal location as the previous cue. Apparently, some aspect of IOR is "attached" to the object and moves with it to new locations (see also Tipper, Weaver, Jerecat, & Burak, 1994).

Abrams and Dobkin (1994) showed that when a previously cued object moves to a new location, subjects are slower to respond to exogenous signals that are part of the object. In contrast, if they are cued to move their eyes to the box by a central arrow, there is no IOR. They concluded that the IOR associated with the saccade system was defined in retinal or environmental coordinates and that IOR associated with stimulus detection was defined in object-based coordinates. They took this as supporting evidence for Klein's (1980) position that selection mechanisms for saccades are not dependent on perceptual attention. However, the perceptual IOR in these experiments could be due to object perception processes that are location-invariant and do not reflect the operation of visual attention. That is, there is no evidence that the slower detection of the signal in object-based coordinates is particularly due to difficulties in attending to the location of the object. A better test would be to see if other signals that are nearby the object but not part of it are also impaired, as would be expected if there is a difficulty in spatial attention. There is nothing in the object-based IOR effect that shows it involves any location-specific processes at all.

Summary. The inhibition of return phenomenon is compatible with the claim that a single spatial attention mechanism is important for both perceptual processes and the programming/execution of saccades. There appears to be a perceptual component of IOR that is object-centered and does not specifically affect the saccade system. A second component of IOR occurs when attention is cued exogenously, and reveals itself in longer latency saccades to the previously attended location. The close relationship between exogenous cuing and the saccade system suggests that exogenous cues may directly activate local saccade programs in the superior colliculus. Saccade programming can also be activated by endogenous processes but only when the subject intends to make a saccade. This program, whether executed or not, slows down subsequent saccades to the same location. In addition, detection of signals associated with the cued object are inhibited and this inhibition moves with the object to new locations.

ATTENTION AND PURSUIT EYE MOVEMENTS

Saccades are not the only variety of eye movement and although they have been the subject of most of the studies on attention and eye movements, there has been important work on pursuit and vergence eye movements as well. Pursuit or smooth eye movements serve to keep a moving target stable on the retina, reducing image blur and maintaining the position of the object in the fovea. Because the world will usually contain a variety of moving and stationary stimuli at any given moment, observers need a mechanism for choosing which stimulus will be pursued and which ignored. Similar to the case with saccades, visual attention provides such a mechanism. Kowler, Van Der Steen, Tamminga, and Collie (1984) presented observers with two identical random dot patterns, one moving and the other stationary. In different conditions, observers were instructed to attend to the moving or stationary pattern. Results showed that the behavior of the eye (moving or stationary) was determined almost completely by the attended pattern, demonstrating that the smooth oculomotor subsystem receives an input from a voluntary attention system.

Later work by Kowler and colleagues (Khorana & Kowler, 1987; Kowler & Zingale, 1985) was aimed at determining whether the attentional mechanism affecting pursuit eye movements was the same one involved in enhancing perception of attended areas. Khorana and Kowler (1987) presented subjects with four horizontal strings of letters moving across the screen from left to right at two different velocities. Subjects were instructed to pursue two rows having a particular velocity. At a point midway in their trajectory, the letters in all four rows were briefly replaced by a target array containing two digits, one in an attended row (the row being pursued) and one in an unattended row. Subjects were to report the identity and locations of both digits. Search performance was strongly influenced by pursuit, with target detection accuracy about 35% better for the target rows than the unattended "background" rows. Of course, this result might be due to the
relatively stable retinal image in the case of pursued letters as opposed to the retinal smear that might accrue to background letters. The authors reject this possibility by considering those cases where tracking was less than perfect, resulting in retinal velocity for attended letters as well. Even when retinal velocity was comparable for target and background rows, attended targets maintained a substantial edge over background rows.

In subsequent experiments, subjects were instructed to try to pay attention to one row, e.g. the slow one, while tracking the fast one. Subjects were unable to fully comply with this instruction, as performance on the attended row was not as good as it was when the row was also the target for pursuit eye movements. Subjects could improve their target detection on the “attended” but untracked rows only by sacrificing pursuit performance on the other row. These results suggest that there is a single visual attention mechanism shared by pursuit eye movements and perceptual tasks.

ATTENTION AND VERGENCE EYE MOVEMENTS

Vergence eye movements occur when both eyes converge to fixate a near target or diverge to fixate a far target. The fixated stimulus will fall on corresponding retinal points (the fovea of each eye) while objects located closer and further from fixation will have disparate retinal images. This disparity, in turn, is processed by the binocular cells of the visual cortex and leads to a perception of depth. As many objects located at various depths are available to serve as targets for a vergence eye movement, the observer must be able to choose the target voluntarily. It seems likely that visual attention is the required mechanism, especially as attention can be deployed to different locations in stereo space (Hoffman & Mueller, 1994).

Erkelens and Collewijn (1991) verified that observers have voluntary control over which stimulus will control vergence. In their experiment, observers fixated a long vertical bar which was flanked on either side by shorter vertical bars, one having crossed disparity (appearing in front of the fixation) and the other, uncrossed disparity (appearing behind fixation). Image stabilization was employed so that eye movements did not change the position of images in either eye. Subjects were instructed to change fixation to either the left or right bar and their vergence changed according to the binocular disparity of the attended bar. Notice here that vergence changes are being initiated by the depth of the attended stimulus and are not due to changes in positions of the retinal images, as they have been stabilized. Thus, the vergence eye movement system, like the saccadic and pursuit systems, appears to be under control of a voluntary attentional system which determines the targets for these various eye movements. It remains to be shown that vergence eye movements are accompanied by an enhancement in perceptual processing of the vergence target.

CONCLUSIONS

The results reviewed here support the claim that visuospatial attention plays an important role in the programming and execution of eye movements including saccades, smooth pursuit, and vergence movements. Eye movements are not random but, instead, are guided by information extracted from the periphery prior to movement. Spatial attention appears to be the mechanism providing this guidance. The evidence favors the parsimonious view that there is a single covert orienting mechanism responsible for both perceptual enhancement and programming of eye movements. This evidence comes in several different guises including: the asymmetric perceptual span in reading, the perceptual enhancement of information at the goal of a saccade, express saccades, and inhibition of return. The details of exactly how attention is used to guide the eye, however, are still a matter of speculation. Here is a short list of issues still outstanding: What peripheral information is used to guide the eye: spaces, word identity, or both? When is this information accessed by attention? When is attention shifted to the saccade goal? How is the location of the saccade goal maintained until saccade generation?

Attentional guidance of eye movements may also help explain how perception and action are coordinated during overt orienting. For example, Currie et al. (1995) recently reported that changes in some of the details of a complex scene that occur during a saccade often go unnoticed. The exception is when such changes involve regions of the scene that are the target of a saccade. They suggest that scene details in the region of the saccade goal are attended and stored in short-term visual memory and are then available for comparison to the perceptual information available at fixation following the saccade. This comparison mechanism offers a novel explanation for the stability of perception during eye movements (McConkie & Currie, 1996). As long as the eye generally lands in the attended area and the world does not move during saccades, the contents of short-term visual memory and current fixation will match. Thus, attention plays a role in guiding the eye to informative areas of a scene, as well as in integrating the separate “snapshots” provided by a moving eye.

ACKNOWLEDGEMENTS

Preparation of this chapter was supported by a grant from the McDonnell-Pew Program in Cognitive Neuroscience. I am grateful to Lou French for helpful comments on an earlier draft of this chapter.
REFERENCES


