CHAPTER 3

VISUAL SELECTION, COVERT ATTENTION AND EYE MOVEMENTS

Our visual environment is crowded with multiple objects, however at any one time we tend to be aware only of a limited part of this array of information (James, 1890). William James noted that the object that we are paying attention to appears to receive more processing and is more richly represented in perception. In addition, paying attention to an object is clearly linked to being able to act on that object—for example, reaching out to pick it up. These then are the central phenomena of visual attention: selection of only part of the visual array and the link between selection and action (see Allport, 1993). The central question in this research area is how to best characterise the mechanisms of visual attention that support these phenomena.

3.1 Covert and overt attention

What happens when we pay attention to some part of the visual environment? With some effort we can fix our eyes straight ahead while at the same time paying attention to some part of the periphery of vision (Helmholtz, 1866). This ability, to pay attention to part of the visual array without moving the eyes, or *covert attention*, has become a cornerstone observation of research on visual attention.

With no effort at all we can move our eyes to align the fovea with an object in the visual array. This ability, to saccade and foveate part of the visual array, or *overt attention*, appeared for many years to be a question that attracted far less interest.

At the heart of our discussion of work on attention is a belief that this emphasis on covert attention is wrong. We argue that spatial selection is best achieved by fixating an item so that it can be processed by the fovea; the processing advantage gained by fixating in this way is substantially greater than the covert attentional advantage. We believe that understanding visual selection primarily has to be about understanding *overt* attention. It is only with this new perspective that we can begin to understand what might be happening when we pay attention. An important issue here is the relationship between covert attention and overt attention or movements of the eyes and we discuss this issue in detail below (Section 3.3 and Section 3.7).

Visual attention appears to have a spatial character; attention after all is most often paid to a location. So in discussing attention we will start with a discussion of spatial attention and more specifically covert spatial attention.

3.2 Covert spatial attention

In a now classic series of experiments, Michael Posner (Posner, 1978, 1980; Posner *et al.*, 1978; Posner *et al.*, 1980) demonstrated that reaction times to visual targets were faster for spatial locations that had been previously cued. Typical experimental displays and results are shown in Fig. 3.1. Subjects were instructed to maintain fixation on a central fixation point. After some interval, an arrow *cue* appeared at fixation indicating the possible location of a subsequent *target*. On the majority of trials the arrow correctly indicated the location of the target: so called *valid trials*. However, on a minority of trials the target appeared in the opposite direction to the cue: so called *invalid trials*. In addition to these two conditions, on some trials no arrow appeared; because these trials gave no prior indication of where the target might appear they were called *neutral trials*. By comparing reaction time across these conditions Posner and colleagues were able to investigate the *costs* and *benefits* of the spatial cues. When compared to the neutral condition, the valid condition led to faster reaction times. In addition, when the cue was invalid, reaction times were slower compared to the neutral condition. These results showed that, even when the eyes did not move, prior information about the possible location of a target led to a benefit, and incorrect information led to a cost.

Such costs and benefits could also be observed when the cue consisted of a peripheral visual event such as a flash. Again, there was a benefit if the peripheral event was in the same spatial location as the target and a cost if the peripheral event was in the opposite location.

These two types of cueing effect share many similarities. However there are a number of reasons to suspect that these two types of cues are, at least in part, functionally different. First, the time course of the cueing effects was different. (e.g. Müller and Rabbitt, 1989). The peripheral cue gave the largest advantage when it occurred around 100 ms before the target and the central cue gave a maximum benefit around 300 ms before the target. Second, when the cue was uninformative—indicating the position of the target on only 50 per cent of the trials—the costs and benefit were still present for a peripheral cue but were absent for a central cue. Posner and others suggested that peripheral cueing was *automatic* and that central cueing was *voluntary*. These two types of cue also map onto the distinction between *exogenous* attention, as demonstrated by the peripheral cue, and *endogenous* attention, as demonstrated by the central arrow cue.

In subsequent experiments Posner and Cohen (1984) demonstrated that for much longer intervals between the cue and the target, there was actually a cost

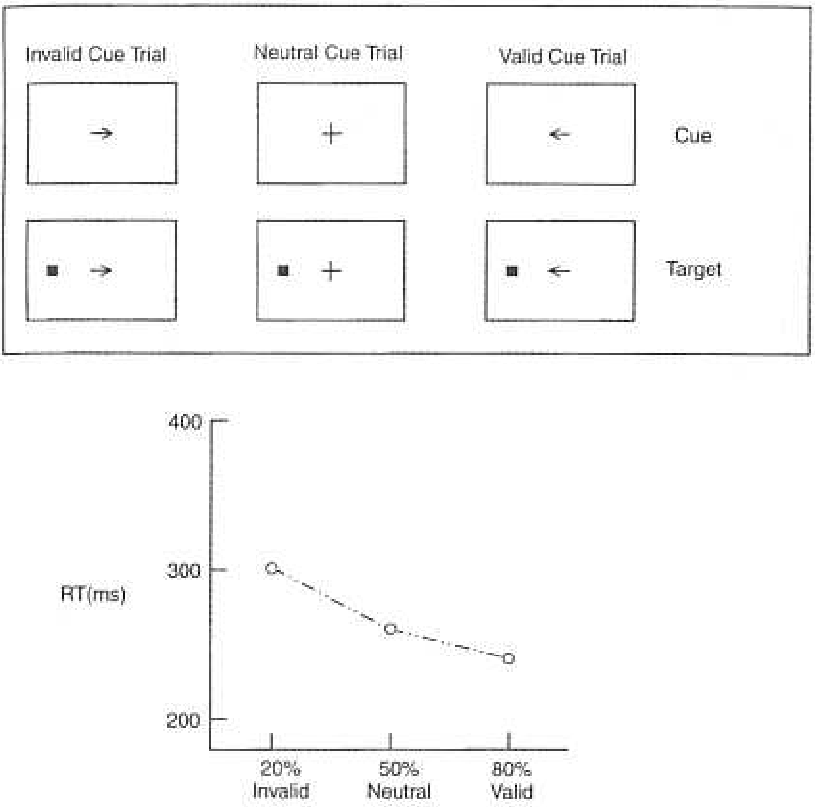

Figure 3.1 The Posner cueing paradigm. The upper panel shows the display sequence for the three trial types. The target is preceded by a cue that indicates the probable location of the target, this cue can be invalid, valid or no cue can be present (so called neutral cue). The cue is followed by a target that the participant is required to respond to as quickly and accurately as possible. The lower panel shows a graph of typical results from this type of experiments. Participants are both faster than the neutral condition for valid trials and slower than the neutral condition for invalid trials.

associated with the cue indicating the target location. They argued that this reflected a process which actively inhibited attention returning to previously visited location. They called this process *inhibition of return* (IOR). As discussed in Section 6.6, IOR may provide an important mechanism to structure scanpaths and prevent rechecking in visual search (see Klein, 2000 for a thorough review of IOR).

A number of models have been proposed to account for these and related attentional enhancement effects, as we shall see in the following sections. One dominant approach is to conceptualise an internal spotlight that has a spatially restricted extent and enhances processing in part of the display; this will be discussed in detail in Section 3.2.1. An extension of this type of model is a zoom lens account in which the spotlight can have a variable spatial extent and so can be distributed over a large part of the display or focused on a small section; this type of account is discussed in Section 3.2.2. Alternatively, as we shall see later on in this chapter, one account of the facilitation effects is that attention is allocated to specific objects rather than parts of space (see Section 3.6.1).

None of these models makes any reference to the fact that the eyes are mobile. An alternative model, proposed by Rizzolatti *et al.* (1987), is the *premotor* model (Section 3.3.3), suggesting that covert attention arises out of the mechanisms of eye movement preparation, although with the actual movement withheld. We propose that such an approach can account for the results as effectively as that of the other approaches and can also, as we discuss in Section 3.7, allow incorporation of the phenomenon of covert attention into our account of active vision.

3.2.1 Spotlights

Posner and colleagues explained their results by suggesting that the costs and benefits resulted from the action of a spatial attentional mechanism that could best be conceptualised as a *spotlight* which was spatially localised and 'moved' across the visual display to facilitate processing in a restricted part of the visual scene. The benefits in the cueing experiments arose when the attentional spotlight had moved to the location where the target appeared and the costs were a result of attention being located at the incorrect location so that attention had to be reallocated in the display to the target.

Within this framework the action and behaviour of the attentional spotlight was characterised by a number of operations. When a location was being attended to, the spotlight was *engaged* at that location. In order to pay attention to a new location, the spotlight had to be *disengaged* from the current location, *moved* across the display and *engaged* at the new location. This cycle of disengage-move-engage characterised the action of the attentional system; each process was assumed to take time and it was these processes which lead to the reaction time costs and benefits in the cueing tasks. So when the cue appeared, attention would disengage from the central fixation point and move to the cued location; if the cue was valid, then detection of the target would occur. However, if the cue was invalid then attention would have to disengage from the cued location and move across the display to the target location; these additional processes took time and led to the costs of invalid cues.

The framework was supported by studies of patients who apparently had a disorder in one of these functions. Patients with damage to the parietal lobe appeared to have a deficit in disengaging attention (Posner *et al.*, 1984): damage to the midbrain, as a result of supranuclear palsy, appeared to result in a deficit in the ability to move the covert attentional spotlight (Posner *et al.*, 1985), and damage to the thalamus resulted in deficits consistent with an inability to engage the spotlight of attention (Rafal and Posner, 1987). Work on visual search (see Chapter 6) initially also supported this spotlight metaphor for attention and it was integrated into one of the dominant models of search, Treisman's Feature Integration theory (Section 6.2.1).

The spotlight metaphor for attention bears many detailed similarities to the properties of the saccadic system. First, like the fovea, the spotlight gives preferential processing to a small area of the visual input. Second, movements of attention, like saccades, take time to initiate and carry out. And third, like saccades, covert attention appears to be slowed down by activity at the fixation point (Mackeben and Nakayama, 1993). Why should covert attention have such similar properties to the saccadic system? Different reasons might be suggested. First, the spotlight model of covert attention might have been inspired by the properties of overt attention; indeed such a similarity may be a great part of its intuitive appeal. Second, the two may share such similar properties because they share some common underlying neural mechanism; this second idea is developed in Section 3.3.3.

3.2.2 Zoom lens accounts of attention

One alternative to the spotlight model of covert attention was proposed by Eriksen and St James (1986). They suggested that rather than having a fixed spatial extent, attention could be allocated over a variable area. This *zoom lens* account of covert attention fitted well with a range of empirical findings (e.g. Egeth, 1977; LaBerge, 1983). If attentional resources are limited and finite then one consequence of increasing the size of the spotlight should be a reduction of the amount of attention allocated to any given location. In support of this, Castiello and Umiltà (1990) used a Posner cueing paradigm and showed that the size of the area cued influenced the extent of a cueing advantage found.

Intuitively, it might be expected that visual attention would allow selection of different visual scales, given the enormous range of scales that objects of interest in the visual environment can take. We will discuss instances elsewhere in the book where it appears that overt attention can also be allocated to spatial regions of different sizes. As discussed in Section 6.8, Zelinsky *et al.* (1997) monitored eye movements during a search task and found that the eye movements were directed to the geometric centres of progressively smaller groups of objects rather than being accurate fixations to individual objects in a display. This *zooming in* of the saccades to progressively smaller units of the display is reminiscent of the zoom lens model. In the search situation the focus is initially broad and so saccades are directed to the centre of groups of items. With subsequent narrowing of the search region, smaller groups are selected until only the target is fixated. The global effect (Section 4.4.3) in which initial saccades to pairs of items are often directed to the centre of mass of the two items also shows the importance of processing at a broad spatial scale. However, McPeek *et al.* (1999) have shown that focused attention may be a requirement for saccade programming. Therefore, it may be that the pattern of saccades observed by Zelinsky *et al.* (1997), and the global effect, reflect the focusing of attention on the centre of mass of a number of elements rather than the action of a more dispersed attentional spotlight.

3.2.3 Late vs. early selection models of attention

One fundamental issue within studies of covert attention is at what stage in information processing attentional selection occurs. Posner's spotlight model of covert spatial attention, like most models of attention, implicitly incorporated the idea of two stages of information processing (e.g. Broadbent, 1958). The first pre-attentional stage is computed in parallel, prior to selection taking place. The second stage involves the more in-depth processing of only a restricted part of the input. This later post-attentional stage has a limited capacity and so only a few items, or a small part of space, can be processed at one time. It is this dichotomy, which pervades research on attention, that results in the questions of how much processing is performed in the first parallel stage, and, conversely, what processes require attention. And models of attention have been classified on this basis. Models are characterised as either *early selection* accounts—attention is required to extract all but the simplest of visual information—or *late selection* accounts—complex object properties are computed from the stimuli before the attentional selection. A similar contrast appears in the models of visual search discussed in Chapter 6. One of the underlying principles that drives this debate on the location of selection is the extent to which one of the functions of attention is to compensate for the limited capacity of the nervous system. Put simply, the argument is that there is insufficient neural capacity to process in-depth all the properties of every item in the visual input. To compensate for this shortcoming, attention is required to select the limited number of items, or part of the display on which in-depth processing will occur. Such a position is more consistent with the early selection models. In the case of vision, it is important to note that a degree of selection occurs simply as a result of the structural distinction between he fovea and the periphery (Section 2.2).

In contrast to the early selection viewpoint, Allport (1993) argued that selection has no single locus and can occur throughout the visual system, at different stages of processing. Within this framework, attention gates the output from multiple visual areas that can drive a response, only allowing a limited part of the information to activate motor areas. This allows outputs from multiple visual areas to drive a single action and ensures that action sequences are produced in a co-ordinated manner. Work on the neuropsychology of attention has provided additional support for multiple anatomical and functional sites for selection, and inspired a number of models (e.g. Humphreys and Riddoch, 1993; Desimone and Duncan, 1995). These multiple sites for selection may reflect functional differences in visual input. Under some circumstances selection will be object-based (see Section 3.6.1) and, as we have seen, selection can be spatially based. In addition, the multiple sites for selection may reflect the involvement of different action systems. For example, the motor systems required to act in reaching space are different from those required to act in the space beyond reaching space (Cowey *et al.*, 1994; Halligan and Marshall, 1991).

3.2.4 The visual benefits of covert spatial attention

There is clearly a huge wealth of evidence that it is possible to allocate attention covertly. However, one of the central questions is what function does covert attention alone serve? Covert attention appears not to be a useful means to accommodate the limited capacity of the system as selection appears to occur throughout the system (Section 3.2.3). A further possible argument that covert attention can scan displays more rapidly than the eyes will be considered (and rejected) in Section 3.4.

One additional, and often neglected, problem facing a model of covert attention is that the magnitude of the effects tends to be relatively small. For example, spatial cueing often leads to a reaction time advantage no greater than 40 ms. A number of studies have attempted to measure which perceptual properties of the stimulus are facilitated and to what extent. The allocation of covert spatial attention lowers orientation thresholds far more than contrast thresholds, and bi-directional vernier thresholds are far more affected than unidirectional thresholds (Lee *et al.*, 1997). In a similar manner, Carrasco *et al.* (2000) found variable levels of facilitation dependent on the nature of the judgement (see also Downing, 1988; Müller and Findlay, 1987). Across all of these studies, the facilitation is relatively small compared to the huge differences in detection thresholds across the visual field (Anstis, 1974). This difference between overt and covert benefits is made even more extreme when the greater effect of lateral masking in the periphery is taken into account (Bouma, 1970). In addition any covert attentional benefits may result in part from differential setting of decision criteria (Downing 1988; Müller and Findlay, 1987).

The magnitude of the benefit associated with covert attending and the apparent lack of some independent function for covert attention suggest that covert attention only makes sense when considered as part of an integrated attentional system that includes both covert and overt attention. To return to the question at the beginning of this section: what function might covert attention *alone* serve? Our answer would be 'very little'. Instead we would strongly argue that covert attention is an integral part of the active vision cycle of fixating items that are of interest. *The fixation act is the process of paying attention* and is supported by covert processes that result in peripheral preview for the next fixation location (see Section 5.3.3, Section 7.2.3 and Section 9.4). In this way covert and overt selection are intrinsically linked. Unfortunately, although the acuity benefits accrued by fixating an item far outweigh the advantage gained by selective attention, such overt attentional explanations for visual selection are often disregarded. However, this approach with its focus on the central role of overt orienting has been embodied in a few models of attention including the Sequential Attentional Model (Section 3.3.2) and the pre-motor theory of attention (Section 3.3.3), which are discussed in detail below.

3.3 The relationship between covert and overt attention

Covert attention leads to processing advantages for localised parts of the visual field without any overt movement of the eyes. As a concept it provides an important organising principle to understand a large body of research findings. One of the central questions of this chapter and others that follow is, what is the role of covert attention when eye movements are not prevented? A critically important point must be to establish the relationship between these two forms of orienting. The relationship between these two processes has been the topic of extensive research and debate. Three clear positions can be identified. The first is that the two processes are independent and co-occur only because they happen to be driven by similar visual input (e.g. Klein, 1980). The second is that the two are closely coupled, with saccades being directed by the location of covert attention. Here covert attention takes the lead and hence can be allocated without a saccade being prepared but not vice versa (e.g. Henderson, 1992). In the third position, developed by Rizzolatti and colleagues (Section 3.3.3), the two phenomena arise out of the action of a single motor system: covert attention is achieved by preparing to generate a saccade. Here, covert attention comes closer to being a by-product of the overt scanning system.

3.3.1 Klein's independence account

Klein (1980; Klein and Pontefract, 1994) argued that two clear predictions could be made if there was a close link between saccade programming and attention. First, that if a subject attends to a location, then saccades to that location should be facilitated. Second, if a subject is preparing a saccade to a location, visual performance at that location should be facilitated. For endogenous cues, Klein (1980) found no evidence for either effect. Klein argued that, although an exogenous cue may attract both the programming of a saccade and an attentional shift concurrently, this correlation is no evidence for a causal link. For endogenous cues, covert attention can be allocated without the programming or preparation to make a saccade. Klein's result has been controversial since a number of subsequent studies have found evidence for coupling effects under various cueing conditions (Deubel and Schneider, 1996; Hoffman and Subramanian, 1995; Kowler *et al.*, 1995; Shepherd *et al.*, 1986). One important factor in determining whether such effects can be detected may be task difficulty. Dual task interference may only be measurable when the task is more difficult (see McPeek *et al.*, 1999).

Klein does not deny that under normal circumstances the processes of generating a saccade and shifting covert attention will occur together or that the two systems are related. His claim is simply that the two responses are not the differential manifestations of the same system. Endogenous covert orienting is accomplished independently of eye movement programming. Remington (1980) also argued for a loose relationship between an attentional and saccade system, suggesting that the saccadic system and the attentional system are both drawn concurrently to peripheral events, but by different mechanisms.

3.3.2 The sequential attentional model

Henderson (1992) developed a *sequential attentional* model in which there is a closer relationship between covert and overt attention. The model is based on four basic assumptions. First, at the beginning of each fixation, attention is allocated to the stimuli at the centre of fixation. Second, attention is allocated to a new stimulus when the fixated stimulus is understood (or identified). Third, the reallocation of attention is coincident with the commencement of saccade programming to the new location that becomes the target for the next saccade. Fourth, the allocation of attention to the new location gates higher level processing at the new location. Within this model eye movements are necessarily led by an attentional shift. Support for such a conjecture comes from work by Shepherd *et al.* (1986) who found that subjects were unable to attend to one location when they were required to programme a saccade in another direction (see also Deubel and Schneider, 1996).

There is considerable evidence that more than one saccade can be programmed concurrently (Section 4.4.4). However, the model is based on a strictly sequential allocation of attention and then the saccade (although this condition is relaxed in the E-Z Reader model that developed from it—see Section 5.7.3). The model Henderson has been developing allows for the parallel programming of saccades by allowing attention to move at a faster rate than saccades and so to be allocated sequentially to two locations before the programming of a saccade to the first location is complete. We shall argue below that the evidence is at best weak that covert attention can move faster than overt attention in the manner that would be required for covert attention to be allocated sequentially in this manner. Henderson (1992) considers the possibility of parallel processing as an alternative to sequential attentional allocation. He rejects the idea that parallel processing might occur across the fixation location and the to-be-fixated location because a benefit is not found for the intermediate locations. In addition he rejects the possibility of the concurrent parallel processing of the fixated location and the spatially independent parallel processing of the to-be-fixated location, in part because this undermines the idea that attention can only be allocated to one location at a time (Eriksen and Yeh, 1985; Posner *et al.*, 1980).

The model also contains a fixation cut-off property; fixation will only be maintained up to a fixed time. After the fixation cut-off point is reached, a saccade is generated regardless of whether a candidate location for the next fixation has been generated by covert attention. Such a property is consistent with results from visual search. In difficult search conditions saccades are often initiated before a peripheral discrimination has been made even though such a discrimination is possible when the eye movement is delayed by training (Brown *et al.*, 1997).

3.3.3 The pre-motor theory of attention

Rizzolatti and colleagues (Rizzolatti *et al.*, 1987; Rizzolatti *et al.*, 1994; Sheliga *et al.*, 1997) have argued for the strongest link between saccades and shifts of covert attention. In the *pre-motor theory of attention*, covert attention effects are a result of activity within the motor systems responsible for the generation of a saccade. Spatial facilitation of the type reported by Posner and colleagues occurs as a consequence of the motor system preparing to generate a saccade. Within this framework, attention is a by-product of the action of motor systems, and attentional effects can be associated with different motor systems or spatial co-ordinates (e.g. in visual neglect, see Chapter 8).

There are two distinct experimental phenomena that provide strong support for a pre-motor theory, in which covert spatial attentional phenomena are a result of processing within the motor systems responsible for overt orienting. These will be considered in turn below.

The first of these is an experimental observation that arises out of experiments using the Posner cueing paradigm as discussed in Section 3.2. A number of groups (e.g. Downing and Pinker, 1985; Reuter-Lorenz and Fendrich, 1992; Rizzolatti *et al.*, 1987) have demonstrated that the reaction time cost in the invalid condition was greatest when the invalid cue was in a different quadrant to the target. This would suggest that there was an additional cost of attention 'moving' across the horizontal or vertical meridian to the target; this effect has been referred to as the *meridian crossing effect*. It is difficult at first to see how this effect could be explained. However, if attentional facilitation is a component of motor preparation, as suggested by the pre-motor theory, then the cost of shifting attention will be a function of the extent of motor reprogramming required to attend to the new location. When reprogramming is required to a different quadrant as opposed to the same one, reprogramming of both direction and amplitude is required; this is not the case for shifts within a quadrant. The meridian crossing effect is well documented for endogenous cueing although it does not appear to occur for exogenous orienting whether the orienting is covert (Reuter-Lorenz and Fendrich, 1992; Umiltà *et al.*, 1991) or overt (Crawford and Müller 1992). Rizzolatti *et al.* (1994) argue that this difference is due to differences in the nature of the motor programme generated in the two cases. In contrast, Klein and Pontefract (1994) argue that the difference is in the nature of the cognitive representation that serves the two types of attentional orienting.

The second piece of evidence that provides support for the pre-motor theory comes from studies of saccade trajectories. Rizzolatti *et al.* (1987) asked subjects to make a vertical downward saccade to a box in response to a cue that could appear in a horizontal row of boxes above fixation (Fig. 3.2). They measured the horizontal deviation in the saccade trajectory and showed that saccades were curved away from the horizontal location of the cue (see also Sheliga *et al.*, 1995). The allocation of covert attention thus has a direct spatial effect on the motor response. Such a direct interaction between the location of covert attentional allocation and a saccade trajectory provides strong support

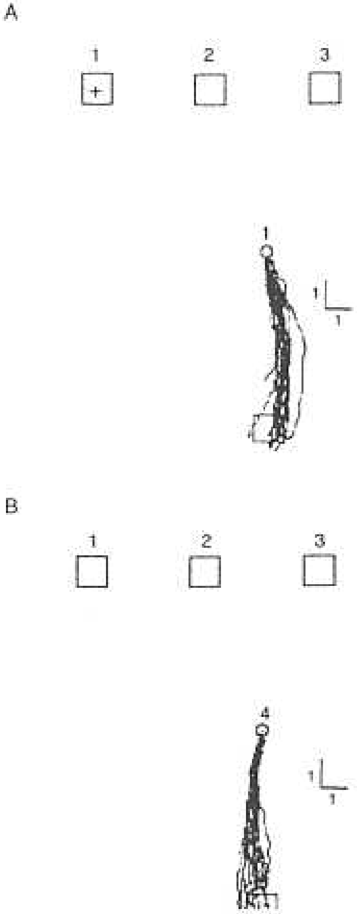

Figure 3.2 The influence of covert attention on saccade trajectory. In this experimental task the participant's covert attention was allocated to the location of the cross that could be located in one of the four boxes in the upper portion of the display. The appearance of the cross was the signal for the participant to generate a saccade to the lower box. For the saccades reproduced in panel A the cross (and covert attention) was allocated on the extreme left hand side and the saccades curved to the right. In contrast when attention was allocated to the right hand side saccades curved to the left. (Reproduced from Rizzolatti *et al.*, 1994, Fig. 9.2).

for the idea that the two responses share a common neural substrate and provides additional support for the pre-motor theory of attention (see also Kustov and Robinson, 1996 for similar effects in monkeys).

3.4 Speed of attention

One of the functional properties often suggested for covert attention is its ability to move around items of a display more rapidly that overt attention. If this were the case then covert attention would allow more items to be scanned in a given time than would be possible with overt attention alone, conferring an obvious functional benefit for the use of covert attention. Thus, a critical question is how quickly can covert attention move around a visual display.

With this motivation, a number of studies have attempted to measure the speed of attention. One indirect method to measure the speed of attention is to calculate the speed on the basis of visual search slopes (Chapter 6). Serial search, when the target is absent, often shows search slopes around 40–60 ms/item. If it were assumed that only one item a time is processed with the attentional spotlight, such a figure would be an estimate of the speed of redeployment of covert attention. This estimate is a great deal quicker than the speed of a series of saccades where the maximum scanning rate would be about 200 ms/item. However, alternative accounts (Section 6.4) propose that multiple items can be processed in parallel even during apparently serial search (Duncan and Humphreys, 1989, 1992; Treisman and Gormican, 1988). Thus it may be argued that serial search slopes are generated by parallel mechanisms, either wholly (Müller *et al.*, 1994; Townsend, 1971) or partially (Wolfe *et al.*, 1989; Wolfe, 1994). These points highlight the serious limitations of using search rates to estimate the speed of attention.

A number of other studies have attempted to measure the speed of attention more directly. Saarinen and Julesz (1991) asked subjects to report letters presented sequentially around fixation. Even with an interval between presentation of the letters as short as 33 ms, performance was above chance. They concluded that covert attention could be moved at a rapid rate. Some doubt has been cast on the exact way that chance performance should be calculated for such tasks (Egeth and Yantis, 1997) and it remains unclear the extent to which performance could be accounted for by parallel mechanism. This paradigm also addresses exogenous attention (Section 3.2), which may show different properties to endogenous attention.

One alternative direct method to measure the speed of covert attention is to use rapid serial presentation (RSVP) techniques. In a majority of these experiments, attentional speed is assessed for responding to two stimuli presented at the same location; these studies are reviewed in detail elsewhere (Egeth and Yantis, 1997) and are not central to the current discussion. However some studies have been carried out using an RSVP type task in which items are spatially separated. Duncan and co-workers (Duncan *et al.*, 1994; Ward *et al.*, 1996) measured the time course for shifting attention by asking subjects to perform two temporally and spatially separated tasks (Fig. 3.3). The two tasks interfered with each other up to and beyond an interval of at least 200 ms. These experiments provided a more direct estimate of the speed of attention. And because of their experimental simplicity they are less likely to suffer from the confound of memory capacity (Egeth and Yantis, 1997). The issue of the speed of attention is clearly still a contentious one and a matter for further detailed experimental study. Given that parallel processing of stimuli may occur in a number of these paradigms, the balance of evidence suggests that the longer estimates more correctly reflect the speed of attention. Estimates of around 200 ms are very close to the time course of overt eye movement based scanning. If these direct estimates

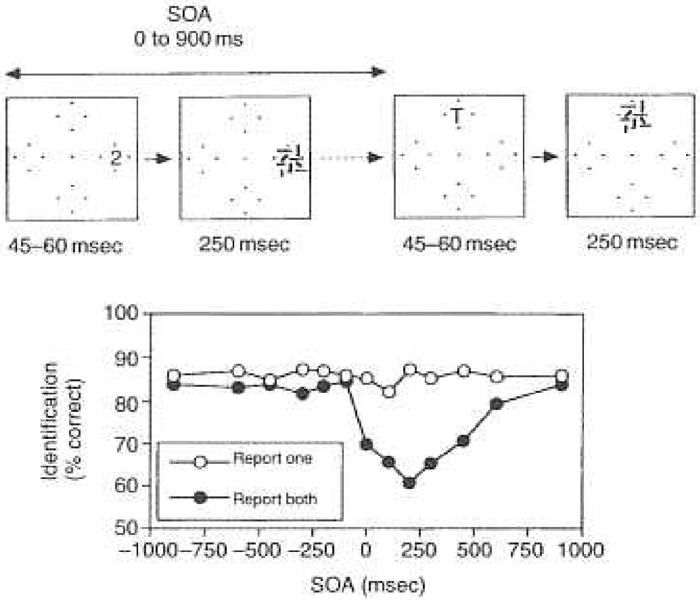

Figure 3.3 The time course of attentional movements. The upper panel illustrates the display sequence for the experiment. A number appears in one of the four locations followed by a mask, this is followed after a variable interval (the SOA) by a letter at a different spatial location. The participant's task is to report both the letter and the number. The graph shows the relationship between the SOA and the percent of items correctly identified. The open circles show the results from a control condition when only a single character has to be reported. The filled circles plot the data when both items have to be reported. It is clearly difficult to attend to (and so report) both items even with a relatively long interval (up to 500 ms at least) between them. This gives a measure of the time it takes attention to 'move' between the two spatial locations. (Adapted from Ward, Duncan, Shapiro 1996, Fig. 2 & 3).

are correct, covert attention simply cannot be a mechanism that allows fast scanning of the visual scene.

Without a good explanation for the purpose of an independent attentional mechanism it seams more probable that covert attention reflects the action of a system closely tied to the overt saccade system, in a manner similar to that proposed by the pre-motor theory of attention.

3.5 Neurophysiology of attention

There have been extensive and detailed studies over the last thirty years, on the neurophysiology of both overt and covert attention. A number of brain

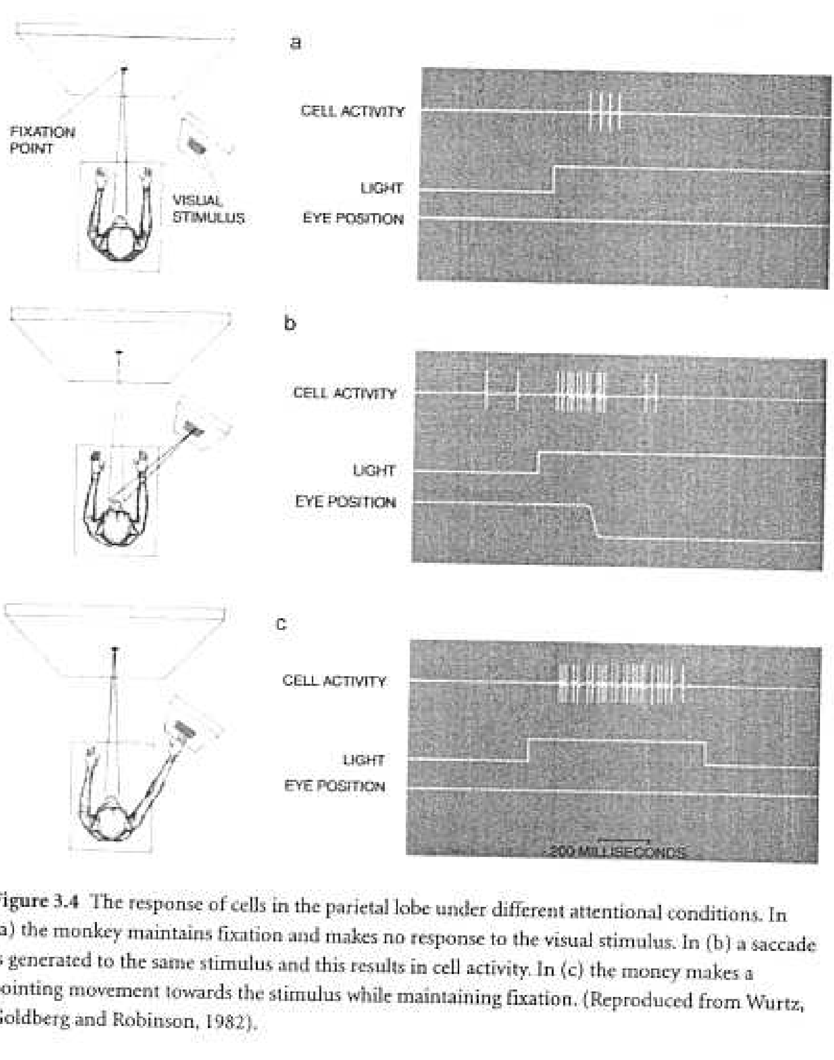

Figure 3.4 The response of cells in the parietal lobe under different attentional conditions. In (a) the monkey maintains fixation and makes no response to the visual stimulus. In (b) a saccade is generated to the same stimulus and this results in cell activity. In (c) the money makes a pointing movement towards the stimulus while maintaining fixation. (Reproduced from Wurtz, Goldberg and Robinson, 1982).

areas have been identified as playing an important role in the control of saccades (Section 4.3). These include the parietal lobe, frontal eye fields and superior colliculus. In each region, covert attentional effects have also been reported.

In a now classic series of studies Wurtz and colleagues (e.g. Goldberg and Wurtz, 1972; Mohler and Wurtz, 1976) recorded from the superficial layers of the superior colliculus in a monkey that had been trained either to make a saccade to a precued light or to attend to the light without making an overt response. Cell activity was linked to the onset of the cue specifically and did not occur when a saccade was generated without sensory stimulation. These cells appeared to be responding as a result of selective attending rather than simply as a result of the subsequently generated eye movement. In contrast, when a manual response was required then the attentional activity from the cells did not occur to the presentation of a cue. This provides strong support for a close coupling between covert attention and saccade generation that both arise from the same basic neural processes.

The response properties of cells in the parietal cortex also show attentional effects but appear to be less dependent on the nature of the response. Wurtz *et al.* (1982) showed that these cells responded equally, regardless of whether the response was a saccade or a manual response to the target without a saccade (Fig. 3.4). At first, this appears to indicate the existence of 'attentional cells' that are not directly linked to the nature of the output. However a manual response is normally associated with a saccade, and so units that allow for the co-ordination of attention across these two may serve an important purpose. Kustov and Robinson (1996) showed that the effects of cueing on cells in the superior colliculus were linked to the presentation of the cue when the response to the target was either a saccadic eye movement or a manual response. This suggests that even when the response does not require a saccade, such spatial cueing results in activity in the SC. When a response is not required or is actively inhibited, attentional allocation may generate concurrent activity in multiple motor systems, particularly if the motor systems act in a co-ordinated manner to control behaviour (see Colby and Goldberg, 1999). Recent evidence from the frontal eye fields provides compelling evidence for a close link between saccades and spatial attention. Moore and Fallah (2001) carried out subthreshold microstimulation in the frontal eye field. Such stimulation was shown to result in improved performance on a covert attention task. Together, these results suggest a close link between spatial attention and saccade generation.

Corbetta and Shulman (1998) review a range of functional anatomical studies in which spatial attention and saccade generation have been studied by means of neuro-imaging (Fig. 3.5). Together, these data indicate that a common set of neural signals in parietal and frontal cortex mediates the covert and overt allocation of attention. The frontoparietal network includes the frontal eye field and supplementary eye field. This anatomical overlap

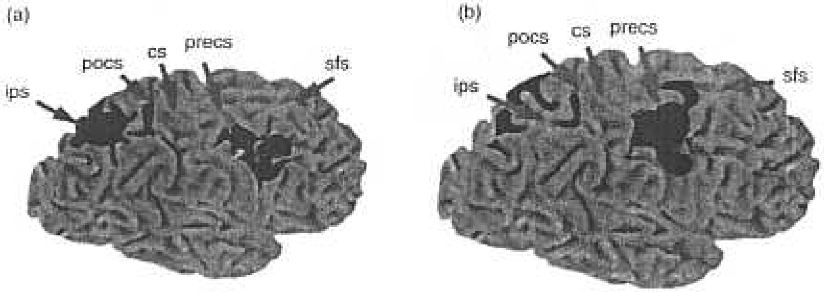

Figure 3.5 Brain activity in the human during covert attention (panel a) and saccadic eye movements (panel b). The areas of activity are shown in black superimposed on a view of the right hemisphere. Results from a number of brain imaging studies are combined, see Corbetta and Shulman (1998) for full details. The following sulci are also indicated: *ips*, intraparietal sulcus; *pocs*, postcentral sulcus; *cs*, central sulcus; *precs*, precentral sulcus and *sfs*, superior frontal sulcus. (Adapted from Fig. 4, Corbetta and Shulman, 1998).

between tasks requiring overt and covert shifts of attention also suggests a close link between these two processes which are more consistent with the pre-motor theory of attention (Section 3.3.3).

3.6 Non-spatial attention

So far this chapter has focussed on spatial selection and we have argued that overt attention to achieve foveation of the target is the primary method by which attention is paid to a specific item and selection of that item occurs. However, selection does not only occur spatially. In this section we discuss non-spatial selection. In the first sub-section we will discuss evidence that selection can occur on the basis of objects. And in the second sub-section we will discuss attention to visual properties.

3.6.1 Attention to objects

The cueing experiments carried out by Posner suggested that attention was allocated to a region of space. Items that were within that part of space received more processing. However, the experiments were equally consistent with a non-spatial, *object based*, allocation of resources, particularly since his design frequently used peripherally located boxes to mark the locations to which attention was to be directed. Duncan (1984) proposed a three-way classification for theories of attention. Object-based theories suggest that processing is limited to a restricted number of objects: objects are the units of selection. Discrimination-based theories propose that a limited number of discriminations can be made. Finally, space-based theories suggest that selection is limited to a fixed part of space. To test these theories, Duncan (1984) asked subjects to perform two simple perceptual tasks concurrently

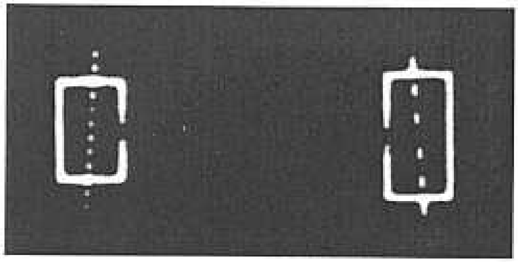

Figure 3.6 Duncan (1984) showed that participants are less accurate when attending to two properties of different objects than two properties of the same object. Participants were presented with a box stimulus superimposed on a line. Two example displays are shown above. In separate trials they were required to report properties of the line (direction of tilt; dotted or dashed) or the box (size of the gap; shape of the box), or one property from each object (Taken from Duncan, 1984, Fig. 1).

(Fig. 3.6). The stimuli consisted of a box and a line drawn through it. The box varied in size (large or small), and had a gap in it (left or right). The line varied in tilt (clockwise or counterclockwise), and pattern (dotted or dashed). The two tasks that subjects were required to do each involved reporting two attributes, either both relating to one object (line or box) or one relating to each object (line and box). Duncan found that there was a large drop in performance when subjects had to attend to two separate objects, rather than to one alone. This was true even if the two objects occupied the same spatial area (as they were overlapping) and were small (so did not require multiple fixations). The parts of the display, relevant to the task, always occupied the same locations, whether part of 1 or 2 objects; so spatial selection factors were held constant across the two tasks, as was the number of discriminatios required. The results could only be explained in terms of an object-based attention account, in which there is an additional cost for attending to two objects over one. Duncan's (1984) result suggests that the selection procedure is not firmly linked to spatial co-ordinates. And in addition, support is provided for a late-selection account of attention, in which there is extensive preattentive processing, with even spatially overlapping items competing independently for selection (see also Baylis and Driver, 1993; Egly *et al.*, 1994; Vecera and Farah, 1994).

The challenge for attentional theory, including the active vision perspective, is to integrate overt attentional allocation into a model of attention in which objects that are behaviourally relevant are facilitated and receive preferential processing via fixation. The contrast between space-based attention and object-based attention will re-emerge in consideration of how to account for eye movement control during reading (Section 5.7). Some influential models take as a start point that attention is deployed at the level of the word unit while others assume a spatial framework.

3.6.2 Attention to visual properties

If a subject is required to select all the red apples from a basket which contains both red and green apples, the red apples are more likely to be selected. This non-spatial facilitation of all red items forms part of the attentional processes that guide selection and clearly cannot be explained by the selection advantage given by fixation alone. Such processes are particularly important in visual search, which is discussed in detail in Chapter 6.

PET studies by Corbetta and colleagues (1991) showed that attending to colour or form or motion increases activity in largely non-overlapping regions of extra-striate cortex. These activated areas corresponded to the areas which, in non-human primates, are known to contain cells that were tuned to these stimulus dimensions. These areas are anatomically distinct from the superior parietal-frontal network that was proposed to be involved in the allocation of spatial attention (Fig. 3.5). These conclusions complement single cell electrophysiology studies carried out by Moran and Desimone (1985). They recorded in area V4 and found that within the classical receptive field, spatial attention could modulate the extent of responding of the cell. These results may build a bridge between how stimuli specific effects—such as attending to the red items in a display interact with spatial selection mechanisms (see also Section 6.7).

3.7 Active vision and attention

An active vision account of attention places fixation as the primary method by which items are selected. As we have seen, fixation confers a large advantage in terms of acuity when compared to the advantage gained by covertly attending. An active vision approach to selection, in which selection occurs via fixation, is in one sense an extreme early selection model. In a single fixation, some items in a display are not selected simply because the sensory apparatus is not sufficient to process them—in some cases visual selection occurs at the retina! However, which items suffer the cost of the poor visual abilities of the periphery can be determined by a range of factors including higher level visual constraints and task demands.

Overt selection becomes more complex when we consider the process by which the next item becomes selected in a series of fixations. Mechanisms drawing on information in peripheral vision determine the next item to benefit from being fixated. These attentional processes themselves operate with reduced visual input. Understanding how the next location for fixation is selected is one of the core questions that this book addresses. And it is clear that these mechanisms themselves are limited. For example, subjects find it difficult to generate a saccade to a face when it is presented amongst a set of jumbled faces (Brown *et al.*, 1997), but saccades can be guided to an item on the basis of colour and shape (Findlay, 1997; see also Chapter 6).

One benefit that covert scanning could confer is that it could allow faster scanning of the visual environment than is possible with overt scanning eye movements. Our review of measures of the speed of attention in Section 3.4 shows that the direct evidence for fast covert scanning is weak. Indeed it appears more likely that covert scanning rates are close to the scanning rates that are possible with overt eye movements. We return to this issue in Chapter 6, where we again conclude that there is no evidence for fast sequential scanning of items with covert attention during visual search.

A rather different argument relating to speed comes from theorists such as Henderson (1992) who believe that there is a close link between covert and overt attention as discussed in Section 3.3.2. This argument suggests that covert attention 'moves' to the target of a saccade some time prior to the eyes themselves moving. The basis for such an argument comes from findings that show improved visual discrimination abilities at the saccade target, immediately prior to a saccade (Kowler *et al.*, 1995; Deubel and Schneider, 1996). This important result is sometimes used to argue that covert attention is primary with saccades being secondary. We prefer a somewhat different approach, since the (potential) dismissal of saccades as secondary is scarcely consistent with their ubiquitous nature. Rather than commenting on the chicken and egg problem raised if the two types of attention are treated as separate, we argue that the two forms of attention are normally inextricably linked and the phenomenon of preview advantage is an important component of active vision.

We should also mention that other possible functions have been suggested for covert attention, although at present the evidence for these functions appears at best weak. One intuitively appealing use would be to allow deception to occur in social situations: covert attention might allow us to pay attention to one person while, because of our point of fixation, appearing to pay attention to someone else. However it remains to be demonstrated that processing facilitation (Section 3.2.4) can be achieved in such a naturalistic situation.

Our argument for the importance of overt attention and the minor role for covert attention is based on some very basic facts about the structure of the visual system. Why then, has research on covert attention dominated for so long? We think there may be multiple possible contributing factors here. First, research in visual attention had its origins in work on audition (e.g. Cherry, 1953). The classic observation from this field is the cocktail party effect. In a crowd where there are many voices speaking at once we can pay attention to just one voice and ignore the others. When we switch auditory attention in these circumstances the process is clearly an internal one although, even in this case, actively orienting the head towards a sound source confers a benefit. When researchers from an auditory background began to study vision they may have taken the auditory system as a model to begin starting to think about visual attention, with the result that they overemphasized covert processes. Second, work on visual attention formed part of the cognitive revolution in psychology; cognitive psychology in general has tended to neglect the motor aspects of task performance and focused on the internal (and so by definition covert) processes. Third, until recently it has been relatively difficult to measure eye movements and so studying covert attention was the most easily achievable goal. Finally, as we have emphasized previously, implicit belief in the passive vision model has been very strong.

3.8 Summary

Attention allows us to select part of the visual information available for further or more detailed analysis. The fovea provides the primary mechanism for such selection to occur. Items that are not fixated receive greatly reduced processing, particularly in terms of acuity. In addition it is possible to attend to items without moving the eyes. This covert attending also confers some processing advantage but these effects are small in comparison to the advantage associated with fixating the item of interest. Both the behavioural and physiological evidence suggests that this covert orienting is closely related to overt saccadic selection. The spatial selection processes works alongside mechanisms that allow feature-based selection; these mechanisms can guide eye movements to behaviourally relevant items.

ACTIVE VISION: THE PSYCHOLOGY OF LOOKING AND SEEING

JOHN M. FINDLAY,
Centre for Vision and Visual Cognition,
Department of Psychology, University of Durham

and

IAIN D. GILCHRIST,
Department of Experimental Psychology,
University of Bristol

OXFORD
UNIVERSITY PRESS

(2003)