1

Vision and Attention

Laurence R. Harris
Michael Jenkin

The term “visual attention” embraces many aspects of vision. It refers to processes that find, pull out and may possibly even help to define, features in the visual environment. All these processes take the form of interactions between the observer and the environment: attention is drawn by some aspects of the visual scene but the observer is critical in defining which aspects are selected.

Although this book is entitled *Visual Attention*, none of the processes of “visual” attention are exclusively visual: they are neither driven exclusively by visual inputs nor do they operate exclusively on retinal information. In this introductory chapter, we outline some of the problems of coming to grips with the ephemeral concept of “visual attention.”

1.1 What Is Attention?

Attention implies allocating resources, perceptual or cognitive, to some things at the expense of not allocating them to something else. This definition implies a limit to the resources of an individual such that they cannot attend to everything at once, all the time. In one sense this is obvious in that the senses already provide a filter. The visual system, for example, does not tell us about what is happening behind us or in the infrared part of the spectrum. But attention refers to selection from the array of information that is arriving at the brain and is potentially available.

The term “attention” suffers from the fact that it is a word in both common and scientific usage and the common and scientific meanings only loosely overlap. Furthermore, both common-parlance and scientific “attention” each have several different meanings. The common usage has an implication of urgency and alertness well - illustrated by the single-word sentence substitution “ATTENTION!!.” In scientific usage, we do not necessarily want to incorporate this sense of the unusual, although we cannot avoid the tint.

Although closely related to “active vision” (see Aloimonos et al., 1988, and also Harris and Jenkin, 1998), attention describes the requirement that the visual system attends at least at a computational level (although not necessarily a physical level) to different events in the visual field. Active systems may require attentional mechanisms in order to direct their sensors at different salient events, but attentive systems are not necessarily active in a physical way.

The word “attention” has ancient roots that link it with concepts of general alertness and conscious receptivity on the one hand and with concentration and focusing on the other. Like other words imported from common parlance, for example “stress,” “attention” can perhaps only be meaningfully considered if accompanied by an adjective. Just as there is no such thing as unqualified “stress” physiologically, it might be useful to start from the stance that there is no such thing as unqualified attention. What types of “attention” are there? We suggest four distinct types:

- **Selective attention**. At any one time we seem to attend to and be aware of only some aspects of the sensory input. What is selected is sometimes determined by the demands of a particular task, such as concentrating on a tool while it is being employed. But selective attention is not only activated by interest in a goal. Although it is obvious that we are selectively attending while actively searching for those lost keys, selection is a feature central to the act of seeing anything. It is impossible, with our meagre brain equipment, to process the whole of the retinal image. Selectively attending to something implies that the feature being attended to has already been defined, whether this is a basic attribute such as a colour, a visual direction, or an actual object.

- **Parsing attention**. Attention might be a part of the process of recognizing objects and separating them from their backgrounds. It has been suggested that the act of attending to something is critical to the binding together of the the various features that define a perceptual object.

- **Directing attention**. When something happens, a primitive reflex system instinctively orients us towards it. This is an emergency, interrupting system, sometimes called attentional capture, which overrides normal behaviour when something potentially important or dangerous demands immediate analysis. But directing attention is also a more gentle, omnipresent drive used for exploring the environment or for maintaining attention on an object while carrying out a task. There might be a continuum between emergency interrupt, normal exploration, and maintaining attention or perhaps these behaviours represent quite different control systems.

- **Alertness attention**. It seems intuitive that a certain level of arousal is necessary for normal perceptual processes to operate. Perception is normally associated with a state of being awake and responsive to sensory input: a state in which behaviours are planned and carried out and in which we are interacting with the environment. But how much undirected arousal is in fact necessary for perception? Some perception can occur while in a daydream or even while asleep (Mack and Rock, 1998). Alertness attention may operate by modulating other forms of attention.

1.1.1 Should "attention" be regarded as a discrete behaviour?

Should attention be regarded as a discrete behaviour that can occur independently of other behaviours? Can the act of attending to something be treated analogously to the discrete act of picking something up between the thumb and forefinger or directing the eyes to converge on a given point? It is possible for finger control or auditory localization to occur quite independently of any other behaviour. Can "attending" be regarded similarly or is it unable to exist independently of other behaviours such that it cannot be regarded as having an independent existence? Perhaps attention is just a modulation of other behaviours.

This is a significant question when seeking to understand attention because if attention can be regarded as a discrete behaviour, then it might be expected to have special brain mechanisms and perhaps brain sites, devoted to it. If, on the other hand, it exists only as a modulation of other behaviours, then no such specialized sites would be expected. Should we seek control systems in the brain that specifically underlie the allocation of attention in the same way as we expect to find systems devoted to auditory localization or locomotion control?

This debate often takes place in the context of whether there is an executive controller. The controller would not be directly involved in any sensory processing but only in the control or modulation of such target systems when it applied attention. Many parts of the visual sensory system show clear modulation in response to attention (e.g., parietal cortex, superior colliculus, even lateral geniculate and striate cortex - see chapter 11 by Bischot), but these areas also show stimulus-related activity when attention is not involved. One area, the cingulate cortex however seems to show activity exclusively when attention is required. Hence, it is a strong candidate for the location of an executive controller (Badgaiyan and Posner, 1998; Carter et al., 1998). When a task that previously required attention becomes automatic through practice, activation of the cingulate cortex is no longer found (Frith et al., 1991).

This therefore gives us a model for what attention is and how it is implemented by the brain. The model is driven by a central executive controller, located at least partially in the ancient limbic association areas, that either makes or at least administers decisions. The decision about what is to be attended is then implemented by selective modulation of the activity or sensitivity of the sensory area or areas that process the desired attribute. The modulation is as specific or fine-tuned as required for the task that the controller is executing. The resulting increased activity in the processing system then gives that coded attribute a competitive advantage or even a flag that leads to preferential treatment. Why might selected attributes need preferential treatment?

1.2 Selective Visual Attention

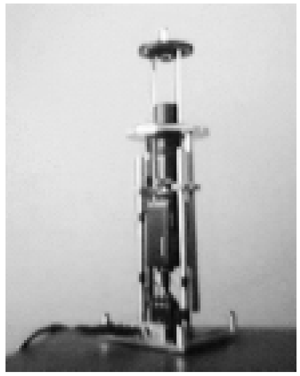
The amount of information falling on the retina at any one time is truly vast and most of it is of no survival value at all (Fig. 1.1). Look at the heavens and all the

FIGURE 1.1. Bookshop scene. The amount of information falling on the retina at any one time is enormous. A selective process is required not only to acquire useful information but in order to be able to see at all. Often the selection will depend on the task at hand. A person just looking at this photo is unlikely to be able to report on the type of lighting, for example, unless attention was specifically drawn to that type of information.

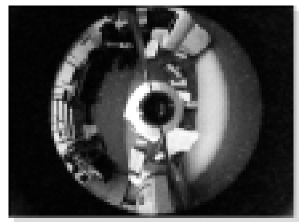
visible stars within the eyes' visual fields are imaged on the retina at once. Look at a tree with its complex and detailed branching pattern, much of which can be resolved at the same time. Look at this page where all the letters are focused on the retina. Most of the potentially smothering information arriving at the retina is lost by the limited resolving power of the retina outside the central few degrees of the fovea. But even with this blesséd filter, there remains an enormous amount of information that could, theoretically, be extracted. Normally, the gaze only stays at one point for about 300 msecs (Yarbus, 1967), implying that the information needed is extracted in this small time before the gaze is shifted again.

Visual attention is an essential component of machine-vision systems. Figure 1.2 shows the Eyes 'n Ears sensor (Kapralos et al., 2000). This is an omnidirectional sensor based on the Paracamera. Consisting of both audio and video components, the Eyes 'n Ears sensor's visual input is generated by a video camera mounted vertically and directed at a semispherical mirror. The resulting view (shown in Fig. 1.2b), provides a 360^o view of the environment. This view can be unwarped to provide a perspective panoramic view (Fig. 1.2c). As in the biological example of Figure 1.1, it is not practical for a machine to attend to each and every pixel in the view. Many are uninteresting for the specific task at hand. For the Eyes 'n Ears sensor, locations in which the image changes are of interest, so the sensor uses an attention-directing mechanism based on attending at image locations that correspond to image differences. Figure 1.3 shows the system in operation. Figure 1.3c is a retinotopic difference image between the image frames shown in Figs. 1.3a 1.3c. Brighter locations correspond to more interesting image locations.

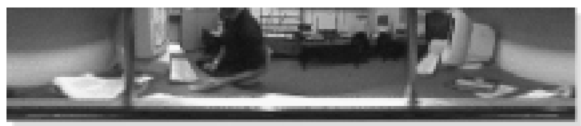
At any one time, the brain just does not need to know very much about the



(a) Sensor



(b) Camera view



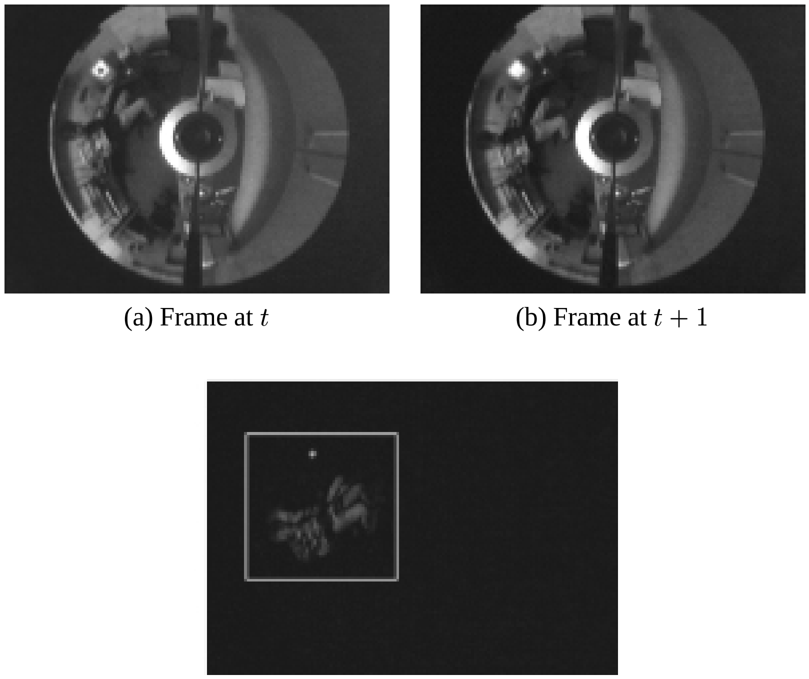
(c) Perspective (unwarped) image

FIGURE 1.2. Eyes ’n Ears sensor.

world - the details of the branching structure of a tree are usually of no use and therefore of no interest. At first glance, it might seem useful to take in the entire content of a page of text at once, like Commander Data of *Star Trek*, but understanding that information is a serial task requiring breaking down, or parsing, the retinal pattern into words and interpreting each word in its correct sequence. This requires selecting the important features (features such as cross-strokes and uprights that define letters, for example) and discarding others (such as the colour or details of the font or the particular layout of the words on the page or paper).

A selective mechanism saves the brain valuable time and processing capacity, allowing it to dedicate its limited resources toward doing something with survival value. It also gives certain information a “hotline” so that it is processed preferentially, resulting in faster and more accurate reactions (Posner et al., 1978) and better sensitivity (Carrasco et al., 2000 – see Chapter 3 by Hawken et al. in this volume). This is not just a general alertness, but a specific visual process because only objects that appear at the expected visual location enjoy enhanced reaction times (Posner, 1980).

At any one time, some information in the sensory array is important and some is not. What is and is not important are by no means simple for the experimenter, or a naturally behaving brain, to determine. Furthermore, what is significant will vary from species to species, from individual to individual, and from moment to moment, even for a given individual and a given scene. If survival is benefitted by filtering salient information from a potentially overwhelming sensory bombardment, what exactly is it that is extracted? How is that done? And what is the connection between “selection” and “attention”?



(a) Frame at t (b) Frame at $t+1$

(c) Difference image

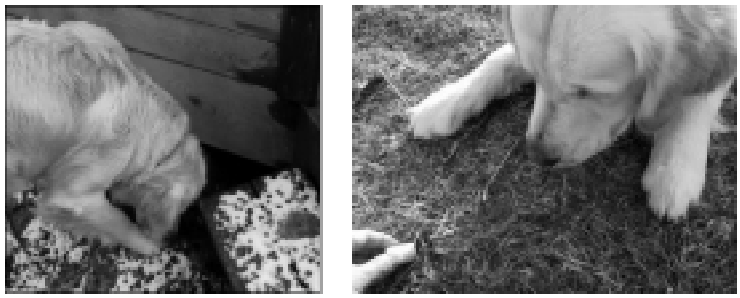
FIGURE 1.3. Attending with the Eyes 'n Ears sensor.

1.2.1 *What is selected?*

People and animals can choose to attend to different things. They can attend to a modality (vision, hearing, touch, taste, smell) or to a colour, or shape or location. Intuitively it seems that the act of attending to one thing makes other things less noticeable - even an entire modality can be ignored in this way.

Outside the laboratory, a specific decision to concentrate on a single attribute is not usually required - the demand arises naturally from the task at hand. When trying to thread a needle, less attention is paid to the radio or the colour of the thread than to the lining up of the thread with the hole in the needle. This does not mean that other things are not processed at all. If something significant happens in an unattended modality (such as hearing one's name), it will not go unnoticed (Moray, 1959; McCormick, 1997). This means that unattended things must be processed to a high enough level to identify them but with a disadvantage or attenuation that normally keeps them unnoticed (Treisman, 1960).

When attending, what actually is being attending to? Is attention directed to a particular zone of space in which interesting things are occurring? This is the implication of the spotlight (Posner, 1980) or zoom-lens (Eriksen and St. James, 1986) models of attention. Under these models, anything appearing in the attended zone enjoys the advantages of attention. Alternatively, perhaps a particular



(a) (b)

FIGURE 1.4. Berkeley can attend either to (a) a location, in this case a hole, or (b) an object, in this case an attractive doggie treat.

object or feature is selected? This would be appropriate for a task-based role for attention. In this case, even if two objects appear close together, only the attended one will receive the full benefits of selection. These viewpoints are reviewed in Chapter 2 by McFadden and Wallman and their differential predictions are illustrated in Fig. 1.5. If attention is object based, then points on the same object (e.g., Point b in Fig. 1.5) should receive more benefits than points on other objects (e.g., Point c) once any part of that object has been selected (e.g., Point a). On the other hand, a spotlight-like, spatially defined attentional beam would radiate symmetrically from its focus, indiscriminately selecting whatever is within the zone and thus giving equal advantage to points b and c (Fig. 1.5).

Curiously enough, the answer does not seem to be straightforward. Whether attention is location- or object-based probably depends on what the subject is trying to do. While attention can be object based (Duncan, 1984; Egly et al., 1994; Weber et al. 1997; Moore et al., 1998), sometimes, as when Laurence's dog, Berkeley is hoping something will come out of a hole (Fig. 1.4), or when a subject is cued to attend to a particular part of a CRT screen (Posner, 1980), it seems possible to attend to an empty location (Egeth and Dagenbach, 1994). On the other hand, when Berkeley attends to a hole, it is possible that other stimuli are attracting or guiding his attention rather than just a volume of space. The visual cues that define the hole are clearly relevant, although they are not the cues attended to as such - they are merely guiding the attention such that a speedy response to the emergence of something interesting can be guaranteed. For the subject viewing the CRT screen, there are also visual cues specifying the location and distance of the attended point on the screen.

We can conclude that visual items can be selected, although the details of what exactly is being selected are still not settled. But how does either a either particular object or a location come to be selected from all its competitors?

1.2.2 How is selection achieved? How much salience is due to the sensory input itself and how much to higher processes?

Of course there is nothing salient about the environment at all. Salience, like colour and beauty, is in the eye of the beholder. Even the most dramatic and attention-grabbing visual events seen on a TV screen pass seemingly unnoticed by Laurence's dog, Berkeley, and fire hydrants, although of great interest to Berkeley, scarcely attract human attention at all. Salience implies some kind of "biological significance".

How is such biological significance to be defined and determined? There is probably only one general rule for stimulus-determined biological significance: "Has something changed?" Although change is neither necessary nor sufficient to make something interesting it clearly indicates something that might not have been previously explored or that might have moved and therefore altered its status as something of interest. One way of assessing the significance of features is simply to present stimuli and see what grabs human attention!

When we look at an image, certain patterns, shapes, or areas defined by various parameters seem to pop out instantly and effortlessly, without the need for searching (Fig. 1.6). In a simple pattern, such as the characters in Figs. 1.6a and 1.6b, this feature cannot be missed or ignored. In a more complex scene, such as a crowded bookshop scene (Fig. 1.1), what is immediately apparent is not any particular pattern but a gist: that it is a bookshop. Neither of these perceptions require careful examination of the scene but instead seem to indicate a global, pre-attentive, parallel processing of the visual information (see Chapter 4 by Wilkinson and Wilson) all at once applied to the entire scene before detailed conscious attention, of the type discussed in the previous section, is deployed (Treisman and Gelade, 1985). These are examples of what perception can do without attention. They are what the next stage, attention proper, must work on. When attention is applied, presumably the popped out features are the most likely to be selected. That would imply that the route of attention is determined by the stimulus.

The question of the relative significance of intrinsic stimulus-based features as opposed to cognitively imposed salience has been explored experimentally by artificially attaching a rather tepid salience to inert little pictures – not biologically but by request. Subjects were asked for example, to search for inverted T's hidden in a mass of distractor items (e.g., Eriksen and Hoffman, 1972, 1973; Posner, 1980). In such tasks, the intrinsic salience is in the feature content that distinguishes the target from its surroundings (Treisman and Gelade, 1988), and the cognitively added salience is that it had been consciously selected as a target.

Change is probably the most reliable, stimulus-based attention attractor, especially when it indicates a new object (Yantis and Jonides, 1984; Yantis and Hillstrom, 1994; Yantis, 1998). Most other cues (Jonides and Yantis, 1988; Theeuwes, 1990), and, surprisingly, even object motion (Hillstrom and Yantis, 1994), turn out not to be reliable stimulus-based attractors. The conclusion is that when looking around a fairly homogeneous scene or photograph, there is usually very little in the scene itself to which everyone would have their attention drawn.

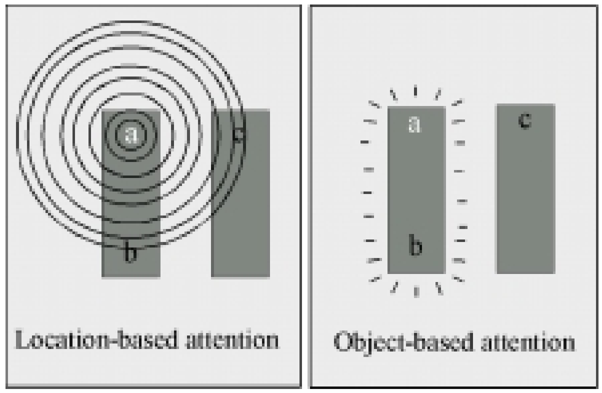


FIGURE 1.5. Spread of attention around a location (left) or around an object (right). The points b and c are equidistant from a, but a and b are on the same object, whereas a and c are on different objects. If attention is object-based, then attending to point a should be beneficial to point b because it is part of the same object. If there is just an indiscriminate spatial spread of attention around location a however, the points b and c should benefit equally.

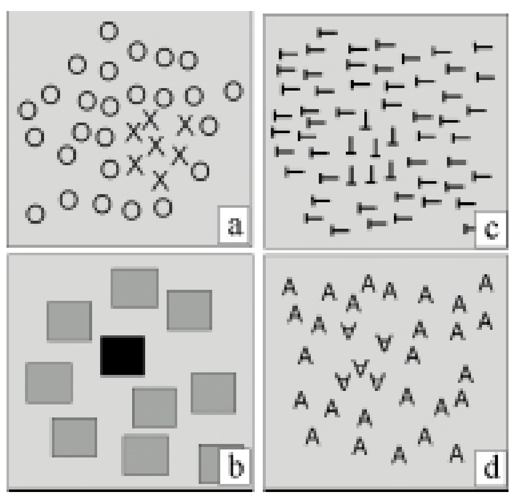


FIGURE 1.6. Examples of pop-out stimuli. Areas within each figure can be detected with different degrees of ease. The areas here are defined by shape (a), contrast and luminance (b), or orientation (c and d). The areas defined are visible immediately (pop out) in some cases (a and b). Not all orientation changes are sufficient to define a form, however, as illustrated in (d).

Salience is thus a varying property of features in the environment determined by an interaction between them and the observer. Salience is determined neither wholly by the properties of the object nor by the state of the observer but by the interaction between them. A food item is only a food item if you are hungry.

1.2.3 What is the connection between selection and attention?

The retinal input is used by many quasi-independent systems for a variety of tasks that are carried out simultaneously. Visual information is used, for example, for pupil control, directing and maintaining fixation, controlling head and body orientation, and for guiding reaching, locomotion, and many other interactions with the environment. Although research in the last few years has elevated these behaviours above the status of mere reflexes and shown them to be quite sophisticated and responsive to different contexts, nevertheless they largely take place without our knowledge or attention. Each of these tasks has particular needs, which involve extracting different features from the retinal input and elsewhere. Mostly, the systems need only to know some special pieces of information and those only very transiently. The fact that your hand is presently three inches from the coffee cup with a time-to-contact of 0.7 sec. is just not something that you need to know at any time other than the present. It is not necessary to store information of this type beyond the time that it is useful (see Chapter 4 by O'Regan). In addition to these various, essentially subconscious, visual functions, however, the existence and details of some visual objects reach perceptual awareness.

Although all tasks in which a person or animal is involved require some selection, only some need and receive attention. Attention implies awareness. Selective attention thus refers only to those selection processes that present their selections to awareness. Why should we become aware of anything? What could be the biological survival value of awareness? Why might some filtered aspects of the sensory information pass into conscious awareness while most of what is not rejected out of hand is dealt with transiently by autonomous mechanisms? The essence is that awareness implies continuity. It is only possible to have an awareness of the present in the context of a past and an anticipation of a future. Awareness then becomes intimately associated with memory (see Dennett, 1991). Memory links the awareness of the present moment to past experience in a way that is necessary for true, context-laden awareness to occur. Further, awareness of the present moment viewed in the context of previous experience allows prediction of the future. Awareness can thus be seen as an emergent property of a selective memory process. Things that are selected then are able to interact with previous memories and might themselves be laid down as memories.

This argument does not outline an exclusive connection between attention, memory, and awareness. It is not necessary that everything must be selected and go through awareness to be remembered or for predictions to be made.

Awareness and memory allow us a degree of flexibility that is difficult to incorporate into an algorithm. It enables us, for example, to acknowledge but to not respond to the repeat of an event that was previously deemed interesting. Awareness allows the application of history and an anticipation of the future. Interest and importance can be added to and removed from the environment at will: sticks can become interesting and scarecrows can be ignored. By this argument, attention becomes the tool of a selective awareness: a high-level process that guides perception. Research into attention thus becomes an exploration of the limits and implementation of awareness. An excellent review of many of these issues can be found in Palmer (1999).

1.2.4 Mechanisms of selective attention

Selective attention is controlled by mechanisms on at least four levels:

1. The mechanism that leads a high-level executive controller to decide to attend to (or to ignore) something. This may be a location, a class of objects, a modality, or a single object, depending on the behaviour in which the subject is engaged.

2. The mechanism by which neural activity related to the target items is given a selective advantage.

3. The mechanism that detects the appropriately enhanced or otherwise flagged activity and passes it to awareness.

4. How awareness uses that information.

Imaging techniques have suggested brain sites for parts of mechanisms 1 and 4, although the details of how they might work are unknown. Most research has concentrated on mechanisms 2 and 3 with exciting results that allow us to relate neural behaviour to perception.

Visual sensory neurones are tuned to particular properties of the stimuli to which they respond. Often, these cells show a spatial tuning, which takes the form of a receptive field: an area of the retina outside which stimuli will evoke little or no response (see Chapter 3). They may in addition be tuned to orientation, wavelength, direction and speed of motion, spatial frequency, or binocularity - all the parameters that describe the image. Different cells in different areas are selective for different attributes (Zeki, 1978). Some cells are tuned for complex combinations of features that provide a high-level description of the scene, such as faces in particular orientations (Perrett et al., 1985). In response to the presentation of a stimulus with a particular combination of features, a given cell will have a particular firing rate. The presence or absence of a cell's firing above the background level is an indicator that those features that it is tuned for, are present (Barlow, 1972; Rieke et al., 1997).

Mechanism 2, the mechanism by which neural activity is given a selective advantage, seems to involve an enhancement of a cell's normal response. For a given stimulus configuration, a larger response than usual is evoked when attention is being paid to those features coded by that cell. Visually responsive cells in many areas show an enhancement if a stimulus that normally excites them is attended (Motter, 1998; Colby and Goldberg, 1999; see Bischot's Chapter of this volume). The enhancement is the signature of the neuronal workings of attention, presumably making the activity easier to detect by the next stage of processing (see Chapter 8 by Tyler). Desimone and Duncan (1995) suggest a competitive neural network for extracting this enhanced activity. Their model has recently received considerable support and is further explored by Wilkinson and Wilson in Chapter 4. A more general model is given by Tsotsos in Chapter 6. One of the problems the enhancement-detection mechanism must face is how to spot an enhanced response in the presence of other responses evoked by other stimulti that, although not attended, might still be powerfully activating their own detecting systems. For example, when searching for a blue object, imagine the brain state when a small, only slightly blue object is spotted among many other brightly coloured, high contrast items that, even unenhanced, would be activating many neurones strongly. How is the weak response enhanced enough to be chosen?

Notice that various other selection procedures referred to in the previous section operate simultaneously and in parallel with the selective attention process outlined here. These other selection processes might share some of the mechanisms of stage 2 and lead to enhancements of their own. Neural activity for targets that are going to be used for the next saccade will be enhanced (Wurtz and Mohler, 1976a, b) but may not necessarily be selectively attended.

1.3 Parsing Attention. Is Attention Central to the Act of Seeing or is it Merely a Servant Carrying its Master to the Right Place?

We have argued that selective attention is employed in the act of bringing things to conscious awareness. Is the act of attention actually a necessary part of defining perceived objects (Yantis, 1998; Regan, 2000)? The various features of the visual world are carried by a whole population of different cells, each involved in coding a different feature of the image. Starting from the retina itself where each cell responds only to features appearing within its discrete receptive field, the processing of the image is spread out over many cells. At higher levels, cells become more specialized in the attributes that they are involved in processing (Zeki, 1978). Different systems encode colour, motion, depth, etc. How this scattered information is put back together to produce a coherent percept of each individual object is called the binding problem. Parsing attention might be an integral part of the solution to the binding problem. Under this theory, the perceptual world is not parsed into individual objects until the act of attending to them somehow binds their features and separates them from their background (Treisman, 1998). This possibility is developed by Regan and Kohly in Chapter 13.

If attention is needed to bind features together to create perceptual objects, then this produces a logical problem for those theories in which selective attention is attracted not to locations but to objects (Duncan, 1984; Egly et al., 1994; Weber et al. 1997; Moore et al., 1998). It would seem under these theories, that objects need to be defined before they can attract attention, but that attention is required to define the object. This conundrum can be resolved by having multiple types of attention as we postulate in this chapter. Directing attention (see section 1.4) works with primitive object definition using only the visual analysis available with the poor resolving power of the peripheral retina. Subsequent application of parsing-attention can aid the object-definition process.

One way to help understand the role of attention in vision is to explore what, if anything, can be perceived without it (see Mack and Rock, 1998). Recently the theme of "change blindness" has highlighted the need to attend consciously to something if it is to be seen at all. If, during some minor distraction or break in attention, a scene is modified, even quite dramatically, that change is often difficult or impossible to spot without many repetitions of the change (see Chapters 9, 10 and 14). Change blindness has been interpreted as illustrating the need for directed attention in order for perception to occur. A representation of the first scene must be retained (memorized) to compare with the second. Change blindness demonstrates how little of a scene is actually memorized normally. But the change blindness paradigm is unnatural in two ways: First, normal vision must be disrupted (by a flash, mud splash, or brief temporal interval) and secondly attention guidance (mechanism 1 in section 1.2.4) is severely hampered because the target for attention is not specified.

1.4 Directing Attention

The metaphor of attention as a spotlight scanning around the visual world (Posner, 1980) naturally links it to the act of directing the gaze. Although the spotlight metaphor does not explain all aspects of selective attention (see Driver and Baylis, 1989 and Chapter 13, for example) it is an intuitively attractive idea and has been very influential in studying attention. It seems that attention does move as a spotlight at least sometimes (Tsal, 1983; Posner and Peterson, 1990; see Chapters 2 and 5 of this volume). Although attention movement is normally linked to gaze control, the spotlight can be moved covertly, independently of actual gaze movement.

Clearly, to be useful, visual attention must be directed. But does the mechanism that does this qualify as a distinct type of visual attention in itself or is it just a transport mechanism? The act of orienting does not need to be powered by vision. Events involving touch or sound often demand orientation even more urgently than vision, and orienting movements do not have to be exclusively visual. An itch on the neck can only be reached with the hand or foot (in Laurence's dog Berkeley's case) and is never available to direct visual scrutiny. This suggests an orienting system independent of the awareness-defined selective visual-attention system and its concern with objects.

Looking around a static scene uses directing attention to select different parts of the scene. These mild, nonurgent orienting movements of the gaze will be driven by a mixture of internal and external stimuli. Internal drives might include a task, such as reading or long-term plans. Attention might be drawn, for example, to a particular book that was of interest weeks earlier but that was spotted while gazing around aimlessly. External stimuli include dramatic, attractive things such as movement or flashes of light (as well as sounds, smells, and tickles) but also features such as edges or corners. This interaction between internal and external contributors to what drives the orientation system is similar to the discussion of salience in the visual-attention system (Section 1.2.2). Are they the same? This question is addressed by McFadden and Wallman in Chapter 2. They conclude, on the basis of adaptation studies, that they are closely linked, whereas Findlay and Gilchrist (Chapter 5) suggest they are more independent.

1.5 Conclusions

The enormous interest in the topic of visual attention reflects the willingness of researchers to consider vision in real life (see Chapter 12 by Bülthoff and van Veen). Rather than treating people as psychophysical machines who always have the same responses to a given stimulus, scientists now freely acknowledge the role of attention and the predominant role of mind-set and higher-level cognitive factors in even very simple visual tasks. How attention contributes to and interacts with perception takes us close to some of the big questions of the nature of consciousness and the connection between neural activity and perception: mind and brain.

References

Aloimonos, J., Weiss, I. and Bandopadhyay, A. (1988). Active vision. *Int. J. Comp. Vis.*, 1:333-356.

Badgaiyan, R. D. and Posner, M. I. (1998). Mapping cingulate cortex in response to selection and monitoring. *Neuroimage,* 7:255-260.

Barlow, H. B. (1972). Single units and perception: a neuone doctrine for perception. *Perception*, 1:371-394.

Carrasco, M., Penpeci-Talgar, C. and Eckstein, M. (2000). Spatial covert attention increases contrast sensitivity across the CSF: support for signal enhancement. *Vis. Res.*, 40:1203-1216.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinik, M. M., Noll, D. and Cohen, J. D. (1998). Anterior cingulate cortex, error detection and online monitoring of performance. *Science*, 280:747-749.

Colby, C. L. and Goldberg, M. E. (1999). Space and Attention in Parietal Cortex. *Ann. Rev. Neurosci.*, 22:319-349.

Dennett, D. (1991). *Consciousness Explained*. Boston: Little Brown.

Desimone, R. and Duncan, J. (1995). Neural mechanisms of selective visual attention. *Ann. Rev. Neurosci.*, 18: 193-222.

Driver, J. S. and Baylis, G. C. (1989). Movement and visual attention: the spotlight metaphor breaks down. *J. Exp. Psychol.: Hum. Percept. Perform.*, 15:448-456.

Duncan, J. (1984). Selective attention and the organization of visual information. *J. Exp. Psychol. Gen.*, 113:501-517.

Egeth, H. E. and Dagenbach, D. (1994). Shifting visual attention between objects and locations: evidence from normal and parietal lesion subjects. *J. Exp. Psychol. Gen.*, 123:161-177.

Egly, R., Driver, J. and Rafal, R. D. (1994). Shifting visual attention between objects and locations: evidence from normal and parietal lesion subjects. *J. Exp. Psychol. Gen.*, 123:161-177.

Eriksen, C. W. and Hoffman, J. E. (1972). Temporal and spatial characteristics of selective encoding from visual displays. *Percept. Psychophys.*, 12:201-204.

Eriksen, C. W. and Hoffman, J. E. (1973). The extent of processing noise elements during selective encoding from visual displays. *Percept. Psychophys.*, 14:155-160.

Eriksen, C. W. and St. James, J. D. (1986). Visual attention within and around the field of focal attraction: a zoom lens model. *Percept. Psychophys.*, 40:225-240.

Frith, C. D., Friston, K., Liddle, P. F. and Frackowiak, R. S. J. (1991). Willed action and the prefrontal cortex in man: A study with PET. *Proc. Roy. Soc. Lond. B.*, 244:241-246.

Harris, L. R. and Jenkin, M. (1998). Vision and action. In L. R. Harris and M. Jenkin (Eds), *Vision and Action*, (pp. 1-12). New York, USA: Cambridge University Press.

Hillstrom, A. P. and Yantis, S. (1994). Visual motion and attentional capture. *Percept. Psychophys.*, 55:399-411.

Jonides, J. and Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Percept. Psychophys.*, 43:346-354.

Kapralos, B., Jenkin, M., Milios, E. and Tsotsos, J. (2000). Eyes 'N Ears: A System for Attentive Teleconferencing. Proc. 139th meeting of the Acoustical Society of America, Atlanta, GA.

Mack, A. and Rock, I. (1998). *Inattentional Blindness*. Cambridge, USA: MIT Press.

McCormick, P. A. (1991). Orienting without awareness. *J. Exp. Psychol. Hum. Percept. Perf.*, 23:168-180.

Moore, C. M., Yantis, S. and Vaughn, B. (1998). Object-based visual selection: evidence from perceptual completion. *Psychol. Sci.*, 9:104-110.

Moray, N. (1959). Attention in a dichotic listening: affective cues and the influence of instructions. *Q. J. Exp. Psych.*, 11:56-60.

Motter, B. C. (1998). Neurophysiology of visual attention. In R. Parasuraman (Eds), *The Attentive Brain*. (pp. 51-70). Cambridge, Mass: MIT Press.

Palmer, S. E. (1999). *Vision Science*. Cambridge, MA: MIT Press.

Perrett, D. I., Potter, D. D., Smith, P. A. J., Milner, A. D., Head, A. S., Jeeves, M. A. and Mistlin, A. J. (1985). Visual Cells in the Temporal Cortex Sensitive to Face View and Gaze Direction. *Proc. Roy. Soc. Lond. B.*, 223:293-317.

Posner, M. I. (1980). Orienting of attention. *Q. J. Exp. Psychol.*, 32: 3-25.

Posner, M. I., Nissen, M. J. and Ogden, W. C. (1978). Attended and unattended processing modes: the role of set for spatial locations. In H. L. Pick and B. J. Saltzman (Eds), *Modes of Perceiving and Processing Information*. (pp. 137-158). Hillsdale, NJ.: Erlbaum.

Posner, M. I. and Petersen, S. E. (1990). The attention system of the human brain. *Ann. Rev. Neurosci.*, 13:25-42.

Regan, D. M. (2000). *Human Perception of Objects*. Sunderland, MA: Sinauer.

Rieke, F., Warland, D., de Ruyter van Steveninck, R., and Bialek, W. (1997). *Spikes: Exploring the Neural Code*. Cambridge, USA: MIT Press.

Theewes, J. (1990). Perceptual selectivity is task-dependent: evidence from selective search. *Acta. Psychol.*, 74:81-99.

Treisman, A. (1960). Contextual cues in dichotic listening. *Q. J. Exp. Psych.*, 12:242-248.

Treisman, A. (1998). Feature binding, attention and object perception. *Phil. Trans. R. Soc. Lond. B Biol. Sci.*, 353:1295-1306.

Treisman, A. M. and Gelade, G. (1980). A feature integration theory of attention. *Cog. Psychol.*, 12:97-136.

Tsal, Y. (1983). Movement of attention across the visual field. *J. Exp. Psychol.*, 9:523-530.

Weber, T. A., Kramer, A. F. and Miller, G. A. (1997). Selective processing of superimposed objects: An electrophysiological analysis of object-based attentional selection. *Biol. Psychol.*, 45:159-182.

Wurtz, R. H. and Mohler, C. W. (1976a). Organization of monkey superior colliculus: enhanced visual response of superficial layer cells. *J. Neurophysiol.*, 39:745-765.

Wurtz, R. H. and Mohler, C. W. (1976b). Enhancement of visual responses in monkey striate cortex and frontal eye fields. *J. Neurophysiol.*, 39:766-772.

Yantis, S. (1998). Objects, attention and perceptual experience. In R. Wright (Ed), *Visual Attention*. (pp. 187-214). New York: Oxford University Press.

Yantis, S. and Hillstrom, A. P. (1994). Stimulus driven attentional capture: evidence from equiluminant visual objects. *J. Exp. Psychol.: Hum. Percept. Perform.*, 20:95-107.

Yantis, S. and Jonides, J. (1984). Abrupt visual onsets and selective attention: evidence from visual search. *J. Exp. Psychol.: Hum. Percept. Perform.*, 10:601-621.

Yarbus, A. L. (1967). *Eye Movements and Vision*. New York, USA: Plenum.

Zeki, S. M. (1978). Functional specialisation in the visual cortex of the rhesus monkey. *Nature*, 274:423-428.