Transsaccadic Integration. The fact that people explore a complex visual scene by making a sequence of many fixations on different informative parts amply demonstrates that normal perception is a dynamic, temporally extended process. But the perceptual result of this sequence of saccades does not resemble even remotely the piecemeal series of images it produces on the retina. Rather, the perception seems to be that of a single, unified scene. This perception must therefore be constructed by integrating the information extracted from the sequence of exploratory eye movements into some coherent internal representation of a single scene (Hochberg, 1970). Moreover, this perception must involve some brief form of memory to bridge the gap between fixations.

One possibility for integrating the contents of sequences of saccades is that they are mapped into a larger, spatially organized memory array according to their positions. The result would be an integrated, composite representation of the visual environment, each fixation being superimposed in its proper location of the array, as illustrated in Figure 11.1.11. As plausible as this **spatiotopic fusion hypothesis** might sound, experimental results show that the visual system does not make use of it (Irwin, 1992), as we will discuss in some detail in Section 12.1.3. Rather, the integration appears to be performed at the level of more abstract representations of objects within the scene.

In addressing this problem of saccadic integration, Hochberg (1970) postulated what he called a **schematic map**: a representation consisting of possible samplings of a spatially extended scene together with contingent expectancies of what will be seen as a result of those samplings. Hochberg never described the internal structure of his schematic maps in detail, but it seems quite likely that structural descriptions, as described in Chapters 8 and 9, would fit the requirements. They explicitly encode the spatial relations among the various parts of an object, effectively specifying the direction and distance at which various configurations of features would be found. In Palmer's face schema (Figure 8.2.16), for example, vectors specify the information required to get from one facial feature to another. This information could be used to support recognition via overt eye movements or covert shifts of visual attention to appropriate places in the image. This information is also encoded redundantly enough to support a variety

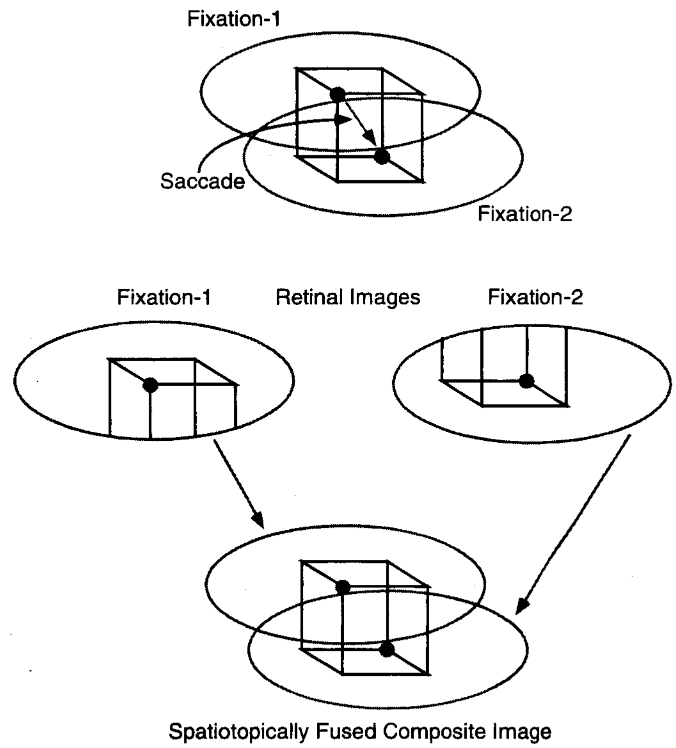


Figure 11.1.11 The spatiotopic fusion hypothesis. When the same scene is viewed by using more than one fixation, the retinal images from the multiple fixations might somehow be fused spatially into a single image in a visual memory.

of different scan paths. Thus, structural descriptions appear to be well suited to serve as the internal structure that accumulates information gathered from a number of different fixations and that integrates it into a coherent, unified whole. This description could then be used to recognize the same object or scene despite a different sequence of eye fixations, provided the proper components and spatial relationships were verified in the input.

11.2 Visual Attention

Even while our eyes are fixated on a particular location, it does not appear that the visual system passively processes all the information available within the image. Rather, we selectively attend to different aspects of it at different times. Sometimes we attend globally to the whole scene; at other times we attend to a selected object or set of objects; at still other times we attend locally to a specific object part. We may even concentrate on a particular property of a particular object, such as the color or texture of a shirt we are considering buying. Our ability to engage in these flexible strategies for processing different information within the visual field—generally referred to as *attention*—is therefore an important component of vision. Indeed, recent experiments suggest that attention may be required for us to consciously perceive anything at all (Mack & Rock, 1998).

Overt eye movements determine what optical information is available to the visual system; covert selective attention determines what subset of this information gets full processing. Attention is such a complex process (or set of processes) that it is difficult to define adequately. For the purposes of this book, however, we will consider **visual attention** to be those processes that enable an observer to recruit resources for processing selected aspects of the retinal image more fully than nonselected aspects.

Notice that this definition implies two different but related functions of attention: *recruiting* resources and *focusing* them on selected aspects of visual information. These correspond to two different properties of attention that theorists often distinguish:

1. *Capacity.* Capacity is the amount of perceptual resources that is available for a given task or process. Attentional capacity can vary with a number of factors, such as alertness, motivation, and time of day (Kahneman, 1973).
2. *Selectivity.* Even at a given moment, when the total capacity is fixed, the amount of attention paid to different subsets of visual information can be allocated flexibly to some degree. This ability allows attention to be selective in terms of what gets processed and what does not.

Of these two aspects, selectivity has been more intensively studied by vision scientists. In the remainder of this chapter, we will mainly be discussing selectivity, although capacity issues will sometimes arise as well.

Complex visual scenes like the ones that we normally look at contain a staggering amount of information, far more than we can be aware of at one time. As a result, we have to sample visual information over time in a series of distinct perceptual acts, each of which is inherently selective. As we discussed in the previous section of this chapter, voluntary eye movements are the first line of visual selection. But even when our eyes are stationary and we are processing a single retinal image, we selectively sample the information it contains for further processing.

This second level of selectivity does not consist of overt, physical acts of orienting such as turning our heads or eyes toward objects of interest, but of covert, internal acts of orienting toward different information available within the retinal image. You can demonstrate the selective effect of attention without eye movements simply by focusing on some small object in your field of view and then attending to (or noticing) various nearby objects—without moving your eyes. This is not particularly easy because eye movements and shifts in attention are normally performed together, but they can be separated with effort. The fact that they can is evidence for the existence of selective attention in vision, independent of eye movements.

A more compelling and rigorous demonstration of our ability to attend to different things without fixating them can be achieved with the help of a camera flash. At night or in a windowless room, turn off the lights and adapt to the dark for a few minutes. Face the flash toward the room (away from yourself) and press the button. This brief burst of intense light will "paint" a single unmoving image of the environment on your retina. It cannot be selectively sampled by foveal fixations due to eye movements, because the afterimage necessarily moves with your eyes whenever you make an eye movement. Even so, you will be able to direct your attention to a number of different objects within the scene before it fades from view. This covert sampling is the work of **spatial selection**, the process of concentrating attentional resources on information from a restricted region of the visual field. At least to a first approximation, such attentional changes seem to be like movements of an internal eye—the so-called "mind's eye"—that can sample different locations within the stationary afterimage. Given the retinotopic organization of much of visual cortex (see Section 4.1.3), spatial selection can be thought of as internally sampling information from a restricted portion of a cortical map.

Spatial selection is only one aspect of visual attention, however. Attention is also at work when we selectively perceive different properties or features of the same object. Keeping a steady gaze on a complex object, for example, you can focus your attention sequentially on its color, its shape, its texture, its size, and so on. This demonstration illustrates **property selection**. It is at first difficult to understand property selection in terms of attention being an internal "mind's eye," for eyes have no physical structure that enables them to select among properties other than space. If different properties have spatially distinct representations within the brain, however, as suggested by the discovery of many different retinotopic maps in visual cortex to encode different visual properties (see Section 1.3.3), then property selection may also be understood in terms of covertly sampling information in different locations of these maps.

11.2.1 Early versus Late Selection

Why do we have the ability to selectively attend to different aspects of visual information? A plausible answer is that it protects the visual system from being overloaded by the massive amount of information available in the visual field. To be effective, however, attention must somehow manage to focus on the most important information given the organism's current goals, needs, and desires. Otherwise, selective sampling would be essentially random, and random selection is not very useful to an organism. Attention is therefore likely to have some means of selecting the most relevant information to process further so that only irrelevant information is rejected.

But how can the visual system choose the most important information without first processing all the information to determine what is most important? This is the *paradox of intelligent selection.* If attention operates very early in the visual system, before much processing has been done, it is unclear how the attentional system can determine what is important. If attention operates relatively late, after a good deal of processing has already been done, it is easy to determine what is important, but much of the advantage of selection would have been lost because most of the irrelevant information has already been processed to perform the selection.

Selective attention to important information is possible by using heuristics based on either innate principles or ones learned through individual experience. It is evolutionarily advantageous to attend to some kinds of information before others. Moving objects are generally important for survival, for example, especially objects that are coming toward you. It therefore makes sense for moving objects to attract your attention for further analysis and for objects that are headed in your general direction to have priority. This attentional heuristic might even be hard-wired at birth through evolutionary processes of natural selection, and it seems plausible that it could be performed very early in visual processing.

In other cases, however, the importance of information is highly specific to an individual. Most people have had the experience of seeing their own name "pop out" of a page of text, for example, and grab their attention before any of the other words. This kind of selection clearly cannot be innate; it must be learned through experience by the individual. By the same token, it seems unlikely that it could be selected at a very early stage of processing, for it seems to presuppose the identification of the letters that make up the name. These theoretical considerations are suggestive, but whether (or how much) attentional selection takes place at early versus late stages of processing is an empirical question toward which many experiments have been directed.

Auditory Attention. The first research on whether attention operates early or late in human perception was conducted in the auditory domain. We will describe it briefly because many of the key questions and theoretical issues were originally explored there.

Auditory researchers began studying attention by asking subjects to perform a **shadowing task** in which they had to repeat aloud the message coming through either the left or right channel of a pair of headphones. The question of interest was what information subjects perceived about the other, unattended channel while performing this shadowing task. Initial results showed that they could perceive gross sensory features without attention, such as whether the unattended channel contained speech sounds or not and whether the voice was male or female. More specific features, such as what was being said or even whether the message was in English or French, were not perceived unless attention was diverted to the unattended channel (Cherry, 1953; Cherry and Taylor, 1954).

On the basis of such findings, British psychologist Donald Broadbent (1958) proposed that auditory attention operated early, analyzing the input to both ears only for gross sensory features and then selecting one ear for further processing of higher-level features to reach the level of meaning. This theory was called **filter theory** because it assumed that selection was due to an all-or-none blocking mechanism (or filter) that passed only the selected channel (see Figure 2.2.5).

Subsequent studies showed that auditory attention was not quite this complete or simple, however. Moray (1959) found that subjects were very likely to hear their own name if it was presented in the unattended channel. This phenomenon may be familiar to you from personal experience. If you are at a party talking with one person, for example, and someone nearby says your name, you are very likely to notice it and to shift your attention to find out why your name was mentioned. Note that this fact causes problems for an early selection theory of auditory attention because it suggests that recognition of your name occurs before selection, not after it, as Broadbent's filter theory would predict.

This difficulty was overcome by supposing that selection operates both early and late in auditory processing (see Figure 11.2.1). According to the most widely accepted theory, often called **attenuator theory** (Treisman, 1960), the initial phase of selection based on gross physical properties is only partial. That is, in contrast to Broadbent's filter theory, early selection merely attenuates (or reduces) the signals in the unattended channels rather than blocking them completely.[2] Attenuator theory can therefore be thought of as a "leaky" version of filter theory.

The second phase of attentional selection in attenuator theory operates during the process of identifying auditory events. Input information first activates **dictionary units**: internal representations of meaningful words and sounds that enable them to be identified. According to attenuator theory, words from both channels activate their corresponding dictionary units, but to different degrees. The units whose input comes from the attended channel are strongly activated, whereas those from the unattended channel are more weakly activated, owing to prior attenuation in the early selection phase. Many dictionary units can thus be active to different degrees at the same time.

The mechanism of late selection within attenuator theory is that dictionary units have dynamic thresholds that must be exceeded for conscious perception to occur. Highly salient items, such as one's own name, have dictionary units with permanently lowered thresholds. Even weak activation from the unattended channel will therefore be sufficient to exceed its threshold and attract attention. This enables attenuator theory to account for Moray's (1959) finding that subjects often hear their name when it occurs in the unattended channel. Less meaningful items will have higher thresholds, so they will tend to be perceived only if they arrive on the attended channel. Treisman also suggested that the thresholds of dictionary units could vary dynamically over time according to context. This would enable contextually expected words to be more easily identified than unexpected words, as reported in other auditory attention experiments (Gray & Wedderburn, 1960).

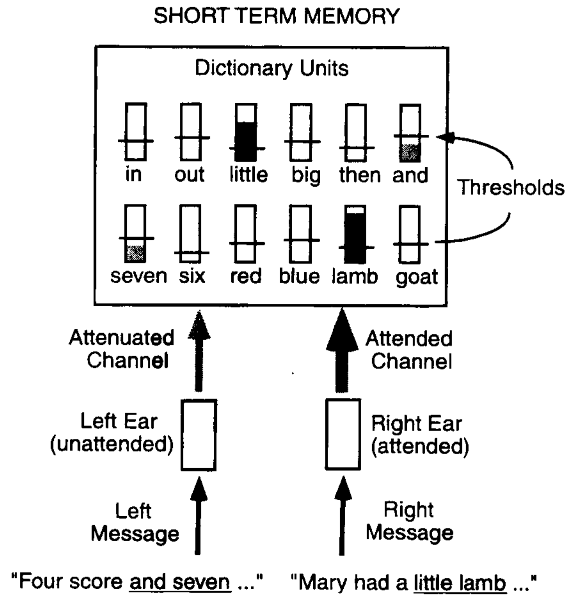


Figure 11.2.1 Treisman's attenuator theory of auditory attention. The message coming from the unattended left ear is attenuated, but both messages activate central dictionary units, only one of which typically exceeds threshold to become conscious. This theory contains mechanisms for both early and late selection.

The Inattention Paradigm. Psychologists Arien Mack and Irvin Rock have recently begun investigating

[2] Equivalently, the information arriving from the attended channel might be amplified rather than those from the unattended channels being attenuated. It is also possible that both amplification of attended information and attenuation of unattended information might occur. The important proposal is that some attentional mechanisms result in relatively more activation from the attended channel than from the unattended ones, regardless of how this effect is achieved.

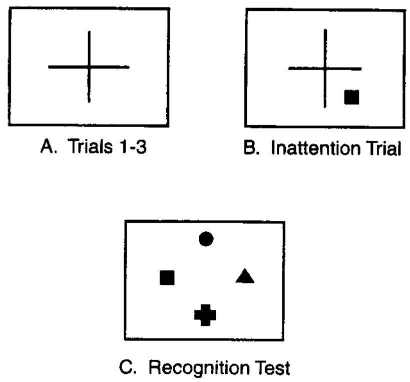


Figure 11.2.2 The inattention paradigm. Subjects are instructed to determine whether the horizontal or the vertical line of a briefly presented cross is longer (A), but on the inattention trial, an extra unexpected element is presented (B). Subjects are asked whether they saw anything besides the cross and are then given a recognition test (C) to evaluate their perception of the extra element.

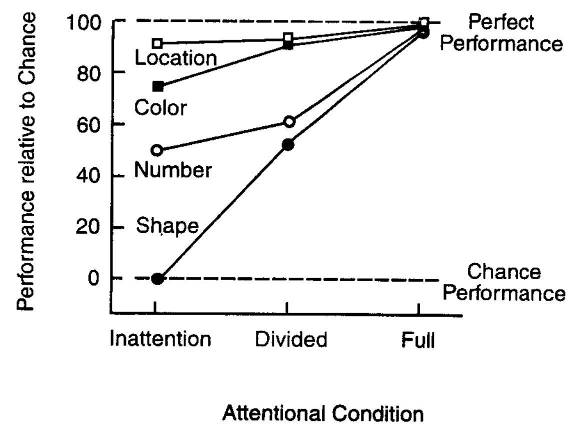


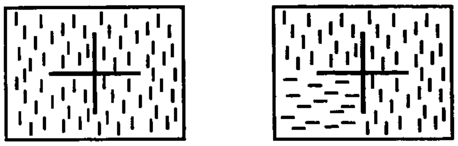
Figure 11.2.3 Results from the inattention paradigm. Subjects perform better than chance at recognizing location, color, and number of elements but not shape. (Data replotted from Rock et al., 1992, to equate chance levels.)

similar questions about early versus late selection in the visual domain (Mack & Rock, 1998). They began by asking what visual features could be perceived without attention. To find out, they had to develop a procedure—which they called the **inattention paradigm**—in which attention would not be focused on the object whose properties they wished to study, even though the object would be clearly visible within the visual field if it had been attended. The task they used was a relatively difficult discrimination in which subjects had to determine whether the vertical or horizontal line of a large cross was longer (see Figure 11.2.2A). The cross was presented briefly (200 ms) and followed by a mask. Several trials of this task were presented with nothing except the cross on the computer screen so that subjects would not expect anything else to be presented.

Then, on the third or fourth trial, the experimenters presented an *inattention trial* in which an additional object was displayed near the cross (Figure 11.2.2B). After the subject said which line of the cross was longer, the experimenter asked whether the subject had seen anything in addition to the cross. Perception of various properties of the unattended object was assessed in this inattention trial through a recognition test in which subjects were asked to pick out the alternative that corresponded to the extra object. An example is given in Figure 11.2.2C for an experiment studying shape perception under conditions of inattention.

After a few more trials of just presenting the cross task, subjects were again shown an extra object and again asked whether they had seen anything. This *divided attention trial* was included because subjects would have been alerted to the possibility of the extra object by the questioning that they received after the inattention trial shortly before. Subjects were given a final trial in which they were told to forget about the cross task entirely and to focus on perceiving anything else that might be present in the display. This *full attention trial* was designed to determine the perceptibility of the extra object under the same presentation conditions as the inattention and divided attention trials but without having to divide attention.

Initial results suggested that simple sensory properties were perceived without attention but that more complex ones were not. Specifically, the results indicated that color, position, and approximate number of objects could be perceived without attention but that shape could not (Rock, Linnett, Grant, & Mack, 1992). This conclusion was based on the fact that subjects performed no better than chance at picking out the correct shape alternative after the critical inattention trial but were almost perfect at doing so after the divided attention trial (see Figure 11.2.3). With properties such as color and position, however, subjects were nearly as good at choosing the correct alternative following the inattention trial as the divided attention trial. These results are reminiscent of early selection phenomena in auditory



A. Trials 1-3 B. Inattention Trial

Figure 11.2.4 Texture segregation requires attention. When background texture elements unexpectedly differed in one quadrant on the inattention trial, subjects did no better than chance at reporting which quadrant was different. Conscious experience of grouping therefore appears to require attention.

attention, much like being able to tell whether the unattended voice was male or female without being able to understand what was being said (Cherry, 1953). Further experiments, to be described shortly, led Mack and Rock (1998) to believe that late selection was also occurring, however.

Other questions that have been examined to clarify the role of attention in perception were whether perceptual grouping or texture segregation occurs under conditions of inattention (Mack, Tang, Tuma, Kahn, & Rock, 1992). In the latter case, for example, the displays contained a texture of randomly placed vertical lines in the background of the cross for all of the initial trials (see Figure 11.2.4A). Then, on the crucial inattention trial, the orientation of all the lines in one quadrant was changed (Figure 11.2.4B), and subjects were asked whether they saw anything different. They did not. On the full attention trial, however, all subjects correctly reported the different quadrant. This indicates that the conscious perception of texture segregation requires attention, contrary to Julesz's (1984) claim that it is preattentive. Gestalt grouping by proximity and lightness were also found to be absent in the inattention trial but clearly present in the full attention trial.

One surprising result from the early studies with single objects was that about 25% of the subjects reported *not perceiving anything at all* on the inattention trial (Mack & Rock, 1998; Rock & Mack, 1994). Mack and Rock refer to this phenomenon as **inattentional blindness**. It cannot be attributed to sensory factors because virtually everyone reported seeing the target on the divided attention and full attention trials. These trials were optically identical to the unattended trial but differed in terms of the subject's expectation. On the inattention trial, they expected only the cross, but on the divided and full attention trials, they were also monitoring for anything else that might be presented. The much higher incidence of missing the target on the inattention trials therefore strongly suggests that expectation is an important component of inattentional blindness.

Subsequent studies demonstrated a number of even more surprising effects concerning the degree of inattentional blindness. For instance, the amount of attentional blindness was actually greater (typically 50–75%) when the extra object was presented foveally at fixation than when it was presented about 2 degrees off center, as in the usual procedure. Most surprising of all, however, was the finding that the degree of attentional blindness depends greatly on the personal meaningfulness of the extra stimulus. As in the auditory domain, Mack and Rock found that only about 5% of subjects were blind to their own name when it was presented under conditions of inattention. Presenting someone else's name under the same conditions led to 35% inattentional blindness, and presenting letter strings with only one different letter—e.g., "Kon" instead of "Ken" or "Jeck" instead of "Jack"—led to about 60% inattentional blindness. Similar, but weaker effects of superior perception for meaningful visual stimuli under conditions of inattention were obtained for certain words (such as RAPE and STOP) and for a standard cartoon "happy face" (but not a sad, neutral, or scrambled face).

Clearly, these results from the inattention paradigm suggest that some form of late selection must be at work in visual as well as auditory attention. Unattended objects must be receiving fairly detailed visual processing for inattentional blindness to be so sensitive to the difference between one's own name and a slight modification of it. Notice that this conclusion actually seems to contradict the earlier empirical finding that shape information is not perceived without attention (Rock et al., 1992). It is not yet clear what the resolution of this conflict will be. Perhaps shape does get *processed* without attention but does not become *consciously perceived* unless attention is then drawn to the object because of its high salience, as would be the case with one's own name. Meaningless shapes would seldom attract attention and therefore fail to become conscious. If their activation dissipates over a matter of seconds, no trace would remain when subjects were asked to report whether they saw anything different. (We will consider some further evidence from the inattention paradigm supporting this view in Chapter 13 when we tackle the topic of visual awareness.) It is puzzling from this hypothesis that properties such as color and position of meaningless shapes are sometimes perceived even when subjects are not expecting them.

Although these and other results suggest that late selection is possible in vision, the early/late question may not have a single solution. Lavie (1995) has recently proposed that both early and late selection occur, but under different conditions. When the task places a high load on visual processing, she finds evidence that selection operates at an early stage of processing, effectively blocking out stimuli other than those within the current focus of attention. When perceptual load is low, however, selection appears to operate at a later stage, allowing the processing of stimuli outside the focus of attention. The task and stimulus conditions in the inattention paradigm are consistent with a low perceptual load, which may explain why Mack and Rock's results largely conform to the predictions of late selection.

Mack and Rock's hypothesis that people are literally blind to unattended visual information, even though such information may be processed extensively at a nonconscious level, is a radical one. Is there any other evidence that bears on it? One compelling source of evidence comes from patients with certain kinds of brain damage who appear not to consciously perceive some of the objects in their visual field because of an inability to attend to them. We will discuss these conditions, known as *unilateral neglect* and *Balint's syndrome*, a bit later in this chapter when we consider neurological mechanisms of attention. But there is also relevant evidence from other studies of normal perceivers who fail to see what would otherwise be clearly perceivable under conditions in which their attention has been captured by some other object or event. We will now examine two of these phenomena, known as the *attentional blink* and *change blindness*.

The Attentional Blink. The **attentional blink** refers to the fact that perception of a second target item is greatly reduced if it is presented within a half second of a first target item (Raymond, Shapiro, & Arnell, 1992; Shapiro, Raymond, & Arnell, 1994). It is typically studied in a **rapid serial visual presentation (RSVP)** search task, in which subjects are shown a very rapid sequence of visual stimuli, all at fixation where acuity is greatest, and are asked to report targets of a specific type (Forster, 1970). For example, subjects might be shown a series of 15 alphanumeric characters at fixation in a period of only 1.5 s (100 ms per character), 13 of which were digits and 2 of which were letters. Their task would be to report the identity of any letters that they saw in the RSVP stream.

If the rate of presentation is no faster than about 11 items/s, the first target can almost always be correctly identified. If the second target is presented more than about 500 ms after the first target, it too is well perceived and reported. But if the second target is presented within about 200–500 ms of the onset of the first target, subjects are very likely to miss the second one completely. They appear simply not to see it.

This phenomenon has been dubbed the *attentional blink* because one interpretation is that after the first target captures the subject's attention, there is a period during which no attention is available for processing the incoming items that immediately follow it, much as blinking keeps visual information from being perceived while the eye is shut. If attention is indeed completely absorbed by processing the first target, then subjects' failure to identify or even detect a second target can be counted as further evidence for Mack and Rock's theoretical interpretation of inattentional blindness: People do not see the second target because it cannot be attended to at the same time the first target is being processed. Other interpretations are possible, however, including ones based on failure of memory rather than perception (e.g., Wolfe, 1999). We will discuss this proposal shortly.

Subsequent results indicate that during an attentional blink an unperceived target nevertheless receives nonconscious processing to the level of meaning. This fact has been demonstrated both behaviorally and electrophysiologically. Behaviorally, the target that appears during the blink has been shown to prime (facilitate) a semantically related third target item that occurs after the blink is over (Shapiro, Driver, Ward, & Sorenson, 1996). Electrophysiologically, the target during the blink has also been shown to influence a component in the evoked potential (called *N400*) that is known to be sensitive to semantic factors (see Shapiro & Luck, 1999). It therefore appears that the attentional blink, like inattentional blindness, operates at a fairly late stage of selection.

Change Blindness. Another phenomenon that may support Mack and Rock's hypothesis that lack of attention to an object causes failure to perceive it comes from a series of recent studies on what has come to be called **change blindness**. The basic finding is that people are surprisingly poor at detecting even gross changes in a visual stimulus if they occur in objects that are not the focus of attention (e.g., Rensink, O'Regan, & Clark, 1995a, 1995b).

A basic change blindness experiment goes like this. Subjects are alternately shown two complex scenes that are identical except for one object or feature that changes. If the two pictures are spatially aligned and presented right after one another with no black interval or distracting event, this task is extremely easy. Attention is immediately called to the stimulus change, which can then be accurately reported. But if a brief blank interval is inserted between the presentations, the task becomes extremely difficult. The same difference that was found effortlessly without a blank interval can now take 20 s or more of repeated alternations while the subject laboriously searches, object by object, for the change.

To experience a version of change blindness for yourself, look back and forth between Figures 11.2.5A and 11.2.5B until you notice the difference between them. If you are like most people, you will find this task surprisingly hard. Once you find the change, you will probably be amazed at how long it took you to spot such a big difference. In this case, there is no blank interval, but the eye movements you have to make between fixations on the target pictures appear to serve the same function and cause blindness to such changes (Blackmore, Brelstaff, Nelson, & Troscianko, 1995; Grimes, 1996). Other abrupt stimulus modifications, such as the sudden appearance of a distracting visual "mudsplash" will produce the same effect (Resnink, O'Regan, & Clark, 1997).

Further experiments show that this insensitivity to change over a short period of time is not a mere laboratory curiosity, but can occur even under normal perceptual conditions. In one study, a subject was approached by a stranger with a map who was asking directions to a location on campus (Simons & Levin, 1997). During their conversation, two workmen, carrying a door lengthwise, walked between the subject and person asking directions. In the few seconds during which the subject could not see the questioner, one of the workmen carrying the door deftly switched places with the person so that after the door passed, the subject was talking to a different person wearing different clothes. Only about 50% of the subjects noticed that any change had occurred when they were asked about it moments later.

A

B

Figure 11.2.5 Change blindness. There is a fairly obvious difference between these two pictures, but people have a hard time detecting it unless they attend directly to the relevant area. (Courtesy of Ron Rensink.)

One interesting interpretation of this literature on the various conditions under which we fail to see things that are quite clearly visible—including inattentional blindness, attentional blinks, and change blindness of various sorts—is that our impression of normal conscious perceptions of our environment as being rich, complete, and detailed is just a grand illusion (O'Regan, 1992). In fact, it is claimed, we experience only the things to which we specifically attend for whatever purposes we currently have in mind, and the rest is simply not perceived because of the lack of focused attention. In this view, the unattended portion of the world *seems* to be there in our perceptions, at least under normal circumstances, because when we examine any given object to see whether it is fully represented in our perception, we necessarily attend to it. Once we have done so, the richly detailed information becomes part of our conscious perception, and seems as though it must have been there all along—even though it hasn't been. The piecemeal nature of our perceptions can therefore be revealed only under ecologically unusual circumstances when the target object changes quickly.

A different interpretation is possible, however, and in some ways preferable. Wolfe (1999) has argued that these phenomena are evidence of **inattentional amnesia** rather than inattentional blindness. He claims that all of these supposedly unseen objects and changes are actually experienced perceptually, albeit very briefly. But without the benefit of focussed attention, he suggests, there is absolutely no memory of them, even over very short time intervals. This account has the advantage of not having to explain away the fact we have conscious perceptual experiences everywhere in our visual field rather than just where we are attending: It is simply because we do have conscious perceptions of unattended objects. The problem comes in reporting these fleeting and fragile perceptions when any sort of memory is required. The unattended perceptions are simply gone by the time we can divert attention to them from either the cross-task in the inattention paradigm, the first target in the attentional blink paradigm, or the blank interval, mudsplash, or interrupting door in the change blindness experiments.

The proper interpretation of these intriguing and important findings is not yet clear. What is clear is that attention somehow plays a very important role in our conscious perception of visual events, by enabling nonconscious visual processing to reach consciousness and/or by creating durable representations in memory that can be used to report fleeting conscious perceptions that would otherwise disappear without a trace.

Intentionally Ignored Information. Thus far, we have concentrated on what happens to objects that are not attended because of lack of expectation (as in the inattention paradigm), lack of resources (as in the attentional blink) or some sort of distraction (as in change blindness). But what about objects that are *intentionally* ignored? What properties, if any, are perceived under these circumstances, and what is their fate?

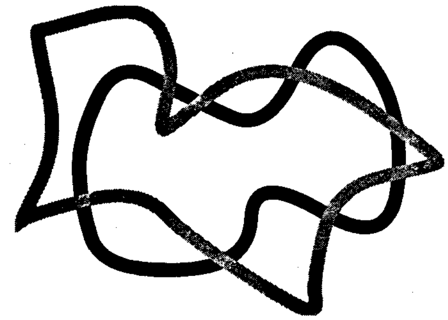


Figure 11.2.6 The overlapping figures experiment by Rock and Gutman (1981). Subjects attended to one of two overlapping novel figures and made a rating of its aesthetic appeal while they ignored the other. Subsequent memory tests revealed good memory for the attended figure but essentially no memory for the ignored one.

The seminal experiment on this topic was performed by Rock and Gutman (1981). They constructed displays containing two novel outline figures that overlapped spatially, one red and the other green, as illustrated in Figure 11.2.6. Half of the subjects were told to attend just to the red figures and to rate them for aesthetic appeal; the other half were told to do the same for the green ones. After seeing a series of such displays, subjects were tested on their memory for the presented shapes using black versions of both sets of figures. Memory for the shape of the attended figures was quite good, but memory for the shape of unattended figures was essentially at chance. Rock and Gutman (1981) concluded that shape was not perceived unless a figure was attended.

There are alternative interpretations of this result, however. Perhaps shape was perceived perfectly well but forgotten so quickly that it was not recognized in the memory test. Or perhaps it was perceived initially but then was suppressed in order to process the attended figure. Although these hypotheses are more complex than assuming that shape was not perceived initially, both are consistent with Rock and Gutman's results.

Using sophisticated experimental procedures, Tipper and his associates have found evidence that the ignored

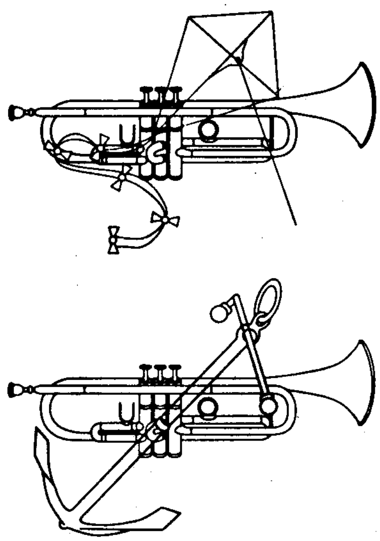


Figure 11.2.7 The negative priming effect. Subjects had to name one of two overlapping familiar figures. On trials in which the target figure was ignored on the previous trial, subjects took longer to name the target than they did when that figure was not present on the previous trial. (From Rock, 1995.)

shape is perceived at some level but not remembered because of active suppression (Allport, Tipper, & Chmiel, 1985; Tipper, 1985; Tipper & Cranston, 1985). The evidence supporting this interpretation is fascinating. Subjects were presented with two meaningful objects using Rock and Gutman's overlapping-contours method (see Figure 11.2.7 for examples in this task) and instructed to name just the ones in a given color (e.g., red), ignoring the other. On *repetition trials,* the target object on a given trial was preceded by a trial in which the same object was presented in the unattended color (see Figure 11.2.7). On *unrelated trials,* the target object was preceded by a trial in which an unrelated object was presented in the unattended color. The surprising result was that naming the target object took significantly longer in the repetition trials (when the same object had been ignored on the immediately preceding trial) than on the unrelated trials (when it had not). This increase in response time has come to be known as the **negative priming effect** because the result is the opposite of what is usually found when a to-be-perceived object has been primed by prior exposure.

The existence of negative priming strongly suggests that subjects must have registered the shape of the unattended object, but suppressed its perception in order to correctly name the target object on that trial. The effect of this suppression is then measured on a repetition trial by slowing the process of naming the same object when it is presented in the attended color. Further experiments showed that the unattended figure appears to be processed at least to the level of meaning, because the slowing of responses also occurred for semantic associates of the suppressed object. For example, if a picture of a dog were presented in the unattended color on one trial, the time to name a cat figure presented in the attended color on the next trial would also be measurably slowed (Tipper & Driver, 1988).

These findings are theoretically interesting for at least two reasons. First, they indicate that intentionally ignored objects receive extensive perceptual processing, at least to the semantic level, because for them to slow responses to a related object on the next trial, they must have been identified on the previous trial. Second, they suggest that attention operates not only by *facilitating* processing of the attended object, but also by *inhibiting* processing of ignored objects. Although it has always been acknowledged that selection can occur either by facilitating the attended object or by inhibiting the unattended ones, Tipper's results have provided strong evidence that both mechanisms are at work, at least in this situation.

What is responsible for the difference between the findings of Rock and Gutman versus Tipper and his associates? There are several possibilities:

1. *Long versus short retention interval.* Rock and Gutman typically assessed perceptual effects by measuring memory after a long delay, after several figures had been presented. Tipper's negative priming paradigm assessed them much more quickly, in the very next trial after only a few seconds had passed.

2. *Indirect versus direct measures.* Rock and Gutman asked their subjects to make direct assessments of perceptual memory, whereas Tipper assessed perceptual memory indirectly by measuring performance on a task in which the effects of a previously seen stimulus could be measured as an increase in response time. The latter may be a much more sensitive index of visual processing than simply asking what subjects remember.

3. *Novel versus familiar stimuli.* To study negative priming effects on naming latencies, Tipper used figures that could be named, and this required familiar, meaningful

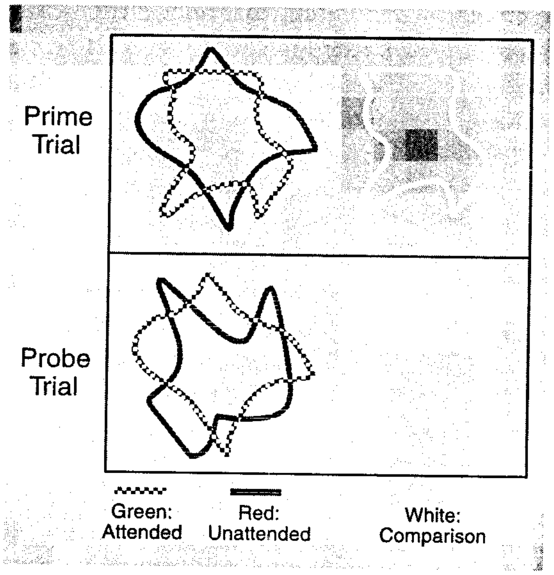


Figure 11.2.8 Negative priming with novel figures. Subjects had to decide whether the green (checkered) and white figures on each trial had the same shape or not. On some sequences of trials, the previous prime trial required subjects to ignore the same figure that had to be attended in the present probe trial. (See text for details.) (After Treisman & DeSchepper, 1996.)

stimuli rather than the novel, meaningless ones that Rock and Gutman typically studied.

Surprising answers to these and other questions have come from a series of experiments by Treisman and DeSchepper (1996; DeSchepper & Treisman, 1996). To address the third possibility, they looked for negative priming effects using novel figures like Rock and Gutman's. They did so by changing the task from naming individual figures to deciding whether pairs of figures were the same or different. Each trial's display contained three figures: an attended figure in the target color (green in Figure 11.2.8) overlapping with an unattended figure in the unattended color (red in Figure 11.2.8), and a comparison figure (always white). Subjects had to indicate whether the green and white figures had the same shape or not, and reaction time was measured. The important question was whether the unattended red figure in the "prime" trial would affect performance on a later "probe" trial in which it reappeared as the attended (green) figure.

The critical experimental conditions were defined by the relation between the attended (green) figure on the probe trial and the unattended (red) figure on a previous prime trial after some number of intervening trials. When they were on consecutive trials (lag = 1), a negative priming effect of 55 ms was obtained for trials in which the same shape was repeated (versus control trials in which a different shape preceded it). This result shows that the difference between Rock and Gutman's findings and Tipper's was not due to the novelty/familiarity of the stimulus materials because negative priming was obtained with novel meaningless figures very much like Rock and Gutman's.

Next, DeSchepper and Treisman investigated the effects of delay. Would the negative priming effect last for only a few trials, as Tipper (1985) had found with his familiar shapes, or would it last as long as the memory delays in Rock and Gutman's experiments? Being careful to show each critical figure in only two trials, they found the same amount of negative priming at lags of 1, 100, and 200 trials. Additional experiments showed that measurable effects of a previous exposure could be obtained up to one month later!

These results indicate that the processes underlying negative priming can be very long-lasting indeed if the figures in question are novel. Moreover, they demonstrate how sensitive indirect measures of memory can be. When explicit memory was tested at comparable delays of 72–104 trials using four-alternative forced-choice recognition procedures—that is, picking the one previously shown figure from among four alternatives—memory for unattended shapes was 26%, no higher than chance (25%). Even attended novel figures were recognized only a bit better than chance (34%) at these delays. The primary reason for the difference between the results of Rock and Gutman and those of Tipper and his associates therefore appears to be the use of direct versus indirect measures of visual memory.

11.2.2 Costs and Benefits of Attention

We now turn to the nature of spatial selection under conditions of explicit attention when the observer is expecting the possibility of some event that contains needed information. We have been presuming that if such an event occurs in a location that is attended, it is processed in ways that are somehow different from how it would be if it were not attended. But precisely what are the consequences of explicitly attending to one object or place rather than another?

Selective attention certainly sounds like a good thing if it enables an organism to focus the bulk of its visual processing capacity on objects, locations, and properties of interest. But this concentration of visual resources presumably comes at a price: Unattended objects and/or properties receive correspondingly *less* processing. This is the "double-edged sword" of selective attention: It may have significant costs as well as significant benefits. For it to be evolutionarily useful, the benefits should outweigh the costs.

The Attentional Cuing Paradigm. The question of how to measure the costs and benefits of selective attention has been studied most extensively by psychologist Michael Posner and his colleagues at the University of Oregon. Posner, Nissen, and Ogden (1978) developed an **attentional cuing paradigm** that has proven to be particularly well suited to examining costs and benefits. The task is simplicity itself: Subjects must press a button as soon as they detect a brief flash of light. The light is presented either to the left or to the right of a central fixation point, as shown in Figure 11.2.9.

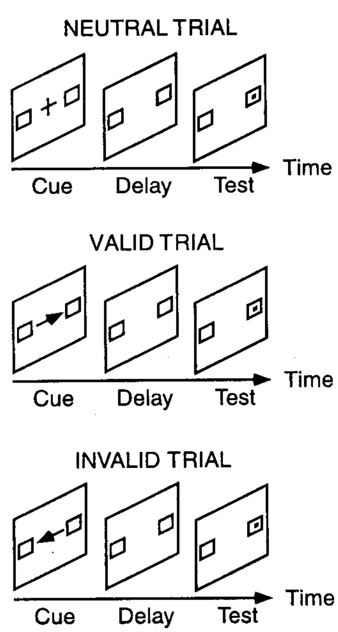


Figure 11.2.9 The attentional cuing paradigm. An arrow pointing toward the left or right box is used to cue subjects to attend there for a target stimulus. On valid trials, the target occurs in the cued location. On invalid trials, it occurs in the uncued location. On neutral trials, the cue is a plus sign.

The crucial manipulation that allowed the costs and benefits of selective attention to be studied was a cue presented in the center of the visual field before the test flash. This cue gave subjects information about where the test flash was likely to appear. A left-pointing arrow (←) indicated that the flash would occur to the left of fixation on 80% of the left-arrow trials. A right-pointing arrow (→) indicated that the flash would occur to the right of fixation on 80% of the right-arrow trials. A plus (+) indicated that the flash was equally likely to occur on either side of the fixation point. The cue was shown 1 second before the test flash appeared, and the subject's reaction time (RT) to respond to the flash was measured. Subjects were instructed to keep their eyes fixated on the center of the screen. To be sure that eye movements did not contaminate the results, however, all trials on which subjects moved their eyes were discarded.

The relation between the central cues (←, →, or +) and the position at which the test flash appeared (on the left or right side of the screen) defines three attentional conditions of interest:

1. *Neutral trials.* When the + cue was presented, subjects got no prior information about the position of the test flash, so they presumably attended equally to both locations. RTs in this divided attention condition constitute a baseline against which RTs in the other two conditions can be compared to evaluate costs and benefits of focused attention.

2. *Valid trials.* On 80% of the arrow-cued trials, the flash appeared on the side to which the arrow pointed. On these "valid" trials, subjects are presumed to move their attention to the location cued by the arrow. If there are measurable benefits of selectively attending to the cued location, detection of the test flash should be faster on these valid trials than on the neutral trials.

3. *Invalid trials.* On 20% of the arrow-cued trials, the flash appeared on the side opposite the arrow, where subjects were *not* expecting it. If there are measurable costs of selectively attending to the cued location, detection of the test flash should be slower on these invalid trials than on both the neutral and the valid trials.

The results of this study are shown in Figure 11.2.10. Posner, Nissen, and Ogden (1978) found that RTs to the valid cues were about 30 ms faster than RTs to the neutral cues and that RTs to the invalid cues were about 30

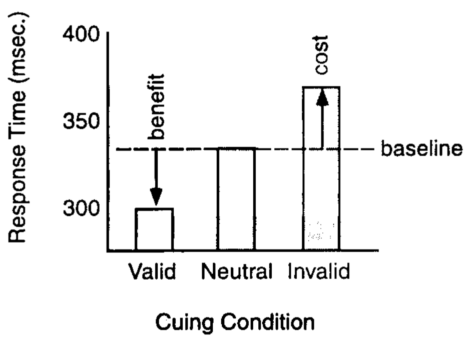


Figure 11.2.10 Costs and benefits of attention. The results in the cuing paradigm show that valid trials are faster than neutral trials (the benefit of correctly directing attention), whereas invalid trials are slower than neutral trials (the cost of misdirecting attention). (Data from Posner, Nissen, & Ogden, 1978.)

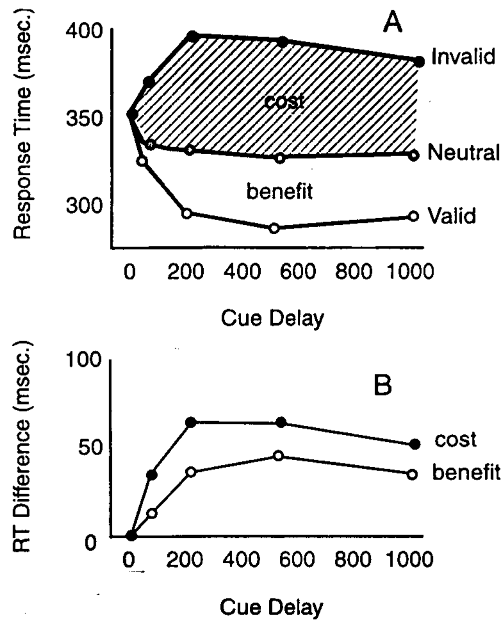


Figure 11.2.11 Temporal development of costs and benefits. Graph A shows the response times for valid, invalid, and neutral conditions as a function of the delay between cue and test displays. Graph B plots the cost (invalid RT minus neutral RT) and benefit (neutral RT minus valid RT) of selective attention. (Part A after Posner, Nissen, & Ogden, 1978.)

ms slower. This indicates that both costs and benefits are present in this particular task and that they are about equal. Given that the 30-ms benefit was obtained on 80% of the cued trials (the valid ones) and the 30-ms cost was obtained on only 20% (the invalid ones), the net benefits outweighed the net costs, at least in this objective sense.

Beyond measuring the basic costs and benefits due to attention, Posner and his associates also wanted to measure how long it takes subjects to shift attention to the cued location. They did this by performing a second experiment in which they varied the time interval between the presentation of the arrow cues and the test flashes (Posner, Nissen, & Ogden, 1978). The shortest interval was 50 ms, and the longest was 1000 ms. The experimenters reasoned that if the test flash were presented too soon after the cue, subjects would not have enough time to shift their attention to the cued location, and neither costs nor benefits would result. As the interval between cue and test increases, however, subjects would be increasingly likely to have completed the shift of attention by the time the test flash appeared. Thus, Posner and his colleagues predicted that both costs and benefits would increase as a function of the cue-to-test interval until some maximum level was reached, indicating the completion of the attentional shift.

The results of this experiment are shown in Figure 11.2.11A. Look first at the neutral trials in the middle. RTs to test flashes after the + cues were not much affected by the cue-test interval, presumably because they provided no information about the location of the test flash. This is the baseline to which performance in the other two conditions should be compared. As in the first experiment, performance on valid trials was faster than that on neutral trials. The difference between these two curves therefore measures the benefit of selective attention (Figure 11.2.11B). Notice that the magnitude of this benefit increases steadily as the cue-to-test interval increases, reaching its highest level at about 400 ms. Performance on invalid trials is again slower than that on neutral trials. The difference between these two curves therefore measures the cost of selective attention (Figure 11.2.11B). Notice that the magnitude of this cost increases as the cue-to-test interval increases, reaching its highest level by 200 ms. Thus, we conclude that attentional shifts from one location to another accrue benefits to the attended location and costs to the unattended location. Moreover, we infer that it takes about 400 ms for people to complete such an attentional shift and that the costs seem to accrue slightly before the benefits.

Voluntary versus Involuntary Shifts of Attention. Attention researchers have extended this experimental paradigm to study the effects of different kinds of attentional cues. You have probably noticed that when there is a sudden change in your field of view, such as the appearance of a new object, it seems to draw your attention automatically. This involuntary summoning of attention appears to be quite different from the voluntary, effortful process of directing attention according to the arrow cues in the experiments just described.

Jonides (1981) extended the cost/benefit paradigm to find out what kind of differences there might be between voluntary and involuntary shifts of attention. He examined voluntary shifts of attention using centrally presented symbolic cues such as the arrows at fixation as described above. These are sometimes called **push cues** because attention must be "pushed" from the symbolic cue to the cued location. He also examined involuntary shifts of attention by presenting peripheral arrows right next to the cued location. Because it was expected that these peripheral cues could effectively summon attention directly to that location, they are sometimes called **pull cues**. Valid and invalid trials for each cue type were constructed by presenting the target object in the cued location or some other location, respectively.

Several differences have been reported between voluntary and involuntary shifts of attention using push and pull cues:

1. *Pull cues produce benefits without costs.* Push cues produced both benefits and costs relative to the neutral condition. In contrast, pull cues produced benefits without corresponding costs.

2. *Pull cues work faster.* When the cue-to-test interval was varied, the results indicated that an equivalent shift of attention took only about 100 ms instead of 200–400 ms.

3. *Pull cues cannot be ignored.* When the validity of push cues was lowered to chance level (50%), subjects were able to ignore them. When the validity of pull cues was comparably reduced, they still produced significant benefits. Indeed, they did so even when subjects were instructed to actively ignore them.

Three Components of Shifting Attention. The results of these experiments on attentional cuing clearly demonstrate that attention has measurable effects on a task as simple as detecting the onset of a visual signal. They also show that it can be moved under either voluntary or involuntary control. Moving attention from one object to another seems intuitively simple enough, but how exactly does it happen?

Posner has suggested that a sequence of three component operations is required to shift attention from one object to another (e.g., Posner & Petersen, 1990; Posner, Walker, Friedrich, & Rafal, 1984):

1. *Disengagement.* Since attention is normally focused on some object, the first thing that must happen is to disengage it from that object.

2. *Movement.* Once it is disengaged, attention is free to move and must be directed toward the new object.

3. *Engagement.* After reaching the target, attention must be reengaged on the new object.

Moving attention from one object to another seems so simple that it is hard to believe that it is composed of these separate processes. However, evidence from neuropsychology suggests not only that they are distinct operations, but that they are controlled by three widely separated brain centers.

Patients with damage to parietal cortex (see Figure 11.2.12) show a pattern of costs and benefits on the cuing task, indicating that they have difficulty *disengaging* their attention from objects. Patients with damage to the superior colliculus in the midbrain show a different pattern of results, suggesting that they have difficulty *moving* their attention. (These patients are also severely impaired in making voluntary eye movements, a fact that suggests an important connection between attention and eye movements to which we will return at the end of this chapter.) Finally, patients with damage to certain centers in the thalamus, including the lateral pulvinar nucleus, appear to have difficulty *engaging* their attention on a new object. Thus, the seemingly simple and unitary operation of shifting attention from one object to another actually requires a coordinated effort among three widely separated regions of the brain. When neural functioning in these areas is impaired, attentional movements fail in predictable ways (see Posner & Raichle, 1994, for a review).

11.2.3 Theories of Spatial Attention

How can we understand visual attention theoretically? As is often the case in cognitive science, the first step in theorizing about a mental process is to find an appropri-

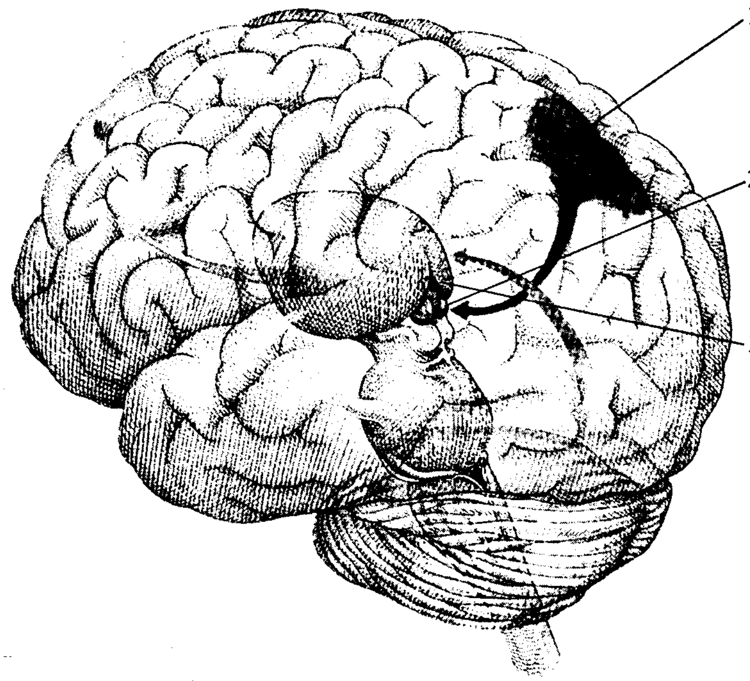


Figure 11.2.12 Three brain centers that are involved in orienting attention. Areas of parietal cortex control the disengagement of attention from objects; circuitry in the superior colliculus in the midbrain controls the movement of attention from one location to another; and certain centers in the thalamus, including the lateral pulvinar nucleus, control the engagement of attention on a new object. (From Posner & Raichle, 1994.)

ate metaphor. Because attention is a rather mysterious, unobservable entity, theorists have tried to understand it in terms of physical systems they can observe directly and therefore understand better. We have already encountered one such metaphor: Attention is like an internal eye. The internal eye metaphor captures some important facts about attention. Two of these facts are that it appears to move from object to object and that it has a fovealike center where processing is concentrated.

The internal eye metaphor is of limited theoretical utility, however, because there is a sense in which it is the operation of the eyes (i.e., vision) that we are trying to explain in the first place. The eye metaphor therefore has the potential problem of infinite regress: If attention is like an internal eye and if the real eye has an internal eye of attention, then does the internal attentional eye also have its own internal eye? And what about the internal eye of that internal eye? Even if the answers to these questions are negative—there may be just one internal eye of attention—many theorists find it preferable to liken attention to something simpler and better understood than an eye.

The Spotlight Metaphor. Among the most crucial aspects of the internal eye metaphor are the selection of one region on which to concentrate processing and the ability to move from one region to another over time. A simpler metaphor that captures both of these characteristics is that attention is like a spotlight (e.g., Posner, 1978). According to this **spotlight theory**, the object at the location where the spotlight of attention is focused is "illuminated" so that it stands out and can be processed more effectively than the less illuminated objects in other regions. Once that object has been processed, the attentional spotlight can be shifted to a different location by moving along a path from the present object to the next one, presumably through the disengage/move/engage sequence of elementary attentional operations mentioned earlier.

The spotlight metaphor of attention has proved both popular and productive. Many experimental results can be understood in terms of it, and new experiments have been devised to test some of its predictions. Consider how Posner's cuing results might be explained, for example. On a cued trial, subjects use the cue to move the attentional spotlight from the central cue to the appropriate location. If the test flash occurs there (a valid trial), it is already in the spotlight of attention, so it can be processed quickly without requiring any subsequent attentional shift. However, if the test flash occurs on the unexpected side (an invalid trial), the spotlight is in the wrong location and must be moved to the correct one before the response can be made. If there is no directional cue (a neutral trial), the spotlight stays in the center and would be moved only half as far as it would on an invalid trial. The spotlight metaphor can thus account for the basic pattern of results shown in Figure 11.2.10. It can also account for the results obtained when the cue-to-test interval is varied, because it will take some amount of time for the attentional spotlight to reach the cued side from the center.

Further experiments have tested a number of predictions derived from the spotlight metaphor. Some have received strong support, but others are controversial. Among the most interesting predictions are the following:

1. *Rate of motion.* The amount of time it takes to shift attention to a target object should increase systematically with the distance over which it must be moved, as though a spotlight were scanning from one place to another. Tsal (1983) has obtained evidence supporting this prediction and has estimated the rate of motion experimentally at about 8 ms per degree of visual angle.

2. *Trajectory.* When a spotlight is moved from one object to another, it illuminates the objects along the path between them. Some evidence suggests that the same is true when attention is moved (e.g., Shulman, Remington, & McLean, 1979).

3. *Size.* Spotlights are generally fixed in size. Eriksen and Eriksen (1974) reported evidence suggesting that the attentional spotlight is about 1 degree of visual angle in size. (As we will see, however, there is also evidence that it can vary in size.)

4. *Unitariness.* A spotlight can be moved from place to place, but it cannot be divided into two or more separate regions. Eriksen and Yeh (1985) reported evidence suggesting that the same is true of attention. (But again there is also evidence for the opposite conclusion.)

Despite its successes, there are a number of problems with the simple spotlight metaphor that have led theorists to consider alternatives. One difficulty is that, despite Eriksen and Eriksen's (1974) conclusion that attention covers only about 1 degree of visual angle in their particular experiment, it seems that under normal viewing conditions attention can cover a much wider area of the visual field, such as when you look globally at a large object or even a whole scene. It also seems that attention can be narrowed to a tiny region of the visual field, as when you scrutinize a small detail. These considerations have led to an alternative metaphor.

The Zoom Lens Metaphor. The **zoom lens theory** likens attention to the operation of a zoom lens on a camera that has variable spatial scope (Eriksen & St. James, 1986). The analogy is not exact, however, for the idea is that attention can cover a variable area of the visual field is usually coupled with the further assumption that varying the size of the attended region changes the amount of visual detail available within it. With a relatively wide attentional scope, only coarse spatial resolution is thought to be possible, whereas with relatively narrow scope, fine resolution is possible.

Shulman and Wilson (1987) tested this idea experimentally. They showed subjects large letters made up of small letters, like the stimuli Navon (1977) used to study global and local processing (see Section 7.6.3), as illustrated in Figure 7.6.9. On some trials, subjects had to identify the large letters, and on other trials the small ones. Shortly after each such trial, they had to respond to a sinusoidal grating that was either low in spatial frequency (wide fuzzy stripes) or high in spatial frequency (thin fuzzy stripes) (see Section 4.2.1). Shulman and Wilson found that responses to low-spatial-frequency gratings were enhanced after subjects had attended to the large global letter and that responses to high-spatial-frequency gratings were enhanced after subjects had attended to the small local letters. This is precisely what would be expected if attention worked like a zoom lens that took time to be adjusted to different sizes and spatial resolutions, large sizes being associated with coarse resolution (low spatial frequencies) and small sizes with fine resolution (high spatial frequencies). These findings are therefore widely cited as supporting the zoom lens metaphor.

Notice that the spotlight metaphor is actually compatible with the zoom lens metaphor in the sense that they can be usefully combined. One can easily conceive of a spotlight that is variable in size as well as position. If the total power of the spotlight is fixed, then a wide beam will illuminate a large region dimly, and a narrow beam will illuminate a small region intensely. This connection between beam width and brightness is not exactly the same as the presumed relation between attentional scope and resolution, but it provides a relatively simple metaphor for thinking about how attention might be distributed over space in a way that includes position, scope, and effectiveness.

Space-Based versus Object-Based Approaches. The metaphors for attention that we have considered thus far—an internal eye, a spotlight, and a zoom lens—all have one important thing in common: They assume that attention selects a region of space. A spotlight, for example, illuminates whatever lies within its beam, whether it is an object, part of an object, parts of two or more nearby objects, or nothing at all. An important alternative to these **space-based theories** is the possibility that attention actually selects a perceptual object (or group of objects) rather than a region of space (e.g., Duncan, 1984). Notice that these **object-based theories** of attention allow a good deal of leeway in how attention might be deployed, because of differences in what constitutes a perceptual object. It could be directed at a single complete object, part of an object, or even an aggregation of objects, as discussed in Chapter 6 when we considered the hierarchical structure of perceptual organization.

Identifying perceptual objects as the domain of selective attention might make object-based accounts seem too ill-defined to be useful, but it does impose significant constraints on the distribution of attention. Unlike space-based theories, for example, object-based theories of attention cannot account for selection of arbitrary portions of two or more different objects, even if they are located within a spatially circumscribed region such as might be illuminated by a spotlight. Also unlike space-based theories, object-based theories can, under certain conditions, account for attentional selection of several discontinuous regions of space. These conditions require that the "object" of attention be a perceptual grouping of several objects whose members are typically defined by some common property (such as color or motion) with other objects interspersed between them. If attention can be allocated just to the set of objects within such a group, it need not occupy a connected region of space. The spotlight and zoom lens metaphors require that a unified region of space be selected.

Some of the strongest evidence for an object-based view of attention comes from a neurological condition known as Balint's syndrome, in which patients are unable to perceive more than one object at any time. We will discuss this syndrome later (in Section 11.2.7) when we consider the physiology of attention, but there is also good experimental evidence for object-based attention with normal perceivers. One of the most widely cited studies is an experiment by Duncan (1984). He reasoned that if attention is allocated to objects rather than to regions of space, it should be easier for subjects to detect two different properties of the same object than two properties of different objects that lie within the same region of space. He showed subjects displays like the one illustrated in Figure 11.2.13. Each stimulus consisted of two objects: a box with a gap in one side and a line running through the box. Each object had two relevant attributes. The box was either short or long and had the gap slightly to the left or right of center. The line was either dotted or dashed and tilted clockwise or counterclockwise from horizontal. After a brief presentation, subjects had to report either one or two attributes. When two attributes were tested, they could belong to the same object or different objects. Duncan found that if the two attributes belonged to different objects, subjects were worse at detecting the second property than

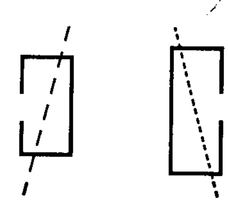


Figure 11.2.13 Stimuli for Duncan's experiment on object-based attention. Subjects had to report two features of a stimulus display that varied on four dimensions: line slant (left or right), line type (dashed or dotted), box length (long or short), and gap placement (left or right). Performance was better when the two features belonged to the same object than to different objects.

the first. But if they belonged to the same object, no such difference was obtained.

These results can easily be explained within an object-based view of attention. When the two properties come from the same object, no shift of attention is required because attention is defined by the single object. When they come from different objects, an attentional shift is required to detect the second property, taking additional time and therefore reducing accuracy. This pattern of results is more difficult to square with space-based theories, however, because the two objects occupy essentially the same region of space. A roughly circular spotlight that illuminates either the box or the line, for instance, will necessarily illuminate the other object. It is therefore not clear how such a difference in detecting properties would arise unless objects were somehow implicated in the allocation of attention. Notice that if a space-based theory allows the attentional spotlight to be shaped tightly around specific objects (e.g., LaBerge & Brown, 1989), they take on a significantly object-based flavor.

There is also recent evidence from a Posner-type cuing experiment suggesting that attention operates at an object-based level. Egly, Driver, and Rafal (1994) showed subjects displays containing two rectangles oriented either horizontally or vertically, as shown in Figure 11.2.14. After the initial presentation, the edges of one end of one rectangle brightened briefly, providing a pull cue to attend to that location. Subjects were then to make a response as quickly as possible when a dark square appeared anywhere in the display. When the cue was valid, the target square appeared briefly in the cued end of the cued object. There were two different types of invalid cues trials, however. In the *same object condition*, the target square appeared in the opposite end of the cued rectangle. In the *different object condition*, it appeared in the uncued object but at the same end as the cue.

As usual in the cuing paradigm, responses were faster when the target appeared at the cued location than when it appeared at either of the uncued locations. The results for the uncued locations showed object-based attentional effects, however, in that the same object condition was faster than the different object condition,

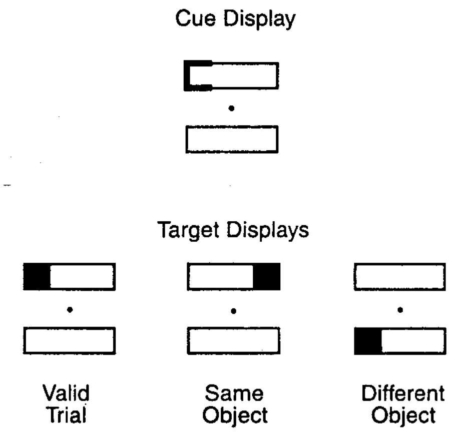


Figure 11.2.14 Stimuli for an experiment showing both space- and object-based attention. One end of two rectangular boxes was cued before presentation of the target in one of three locations. Same-object responses were faster than different-object responses, indicating an object-based attentional effect.

even though their distances from the cued location were equal. This finding suggests that switching from one object to another incurs an additional cost that cannot be attributed to distance.[3]

A third finding that lends credence to the object-based view concerns an observer's ability to keep track of several moving objects at once. Pylyshyn and Storm (1988) showed subjects a random array of static dots and designated several of them as target objects to be attended by flashing them on and off. All the dots then began to move in quasi-random (but continuous) trajectories, and subjects were instructed to try to keep track of the ones that were initially designated as targets. After several seconds of such motion, one of the elements flashed, and subjects were asked whether or not that particular dot was one of the initially designated targets. Pylyshyn and Storm found that subjects could track as many as five dots at once.

If one believes that this tracking task requires attention (and Pylyshyn and Storm do not), these results pose significant problems for space-based theories of attention. First, most space-based theories assume that the region to which attention can be allocated is a unitary,

[3] More recent experiments have shown that the "objects" in question are perceptually completed objects rather than retinally defined objects. Moore, Yantis, and Vaughan (1998) concluded this after finding results similar to those of Egly, Driver, and Rafal when the different ends of the same-object stimuli had been retinally separated by an occluding object. They also found this pattern of results when the two ends of the same object were defined only by illusory contours.

convex, connected area. However, the tracking task seems to require observers to attend to a number of disconnected spatial regions at the same time. Equally important, the trajectories of regions through space are defined only by virtue of the objects that traverse them. How else could attention be allocated to the proper regions of space at the proper times?

These results have another feature that is at least somewhat troubling from the object-based perspective, however: They seem to indicate that attention can be split among multiple objects. One could, of course, extend the object-based view specifically to allow for attention to be divided among some small number of objects. Indeed, this is how Pylyshyn and Storm (1988) interpreted their results. But there is another possibility that does not require giving up the unitary nature of attention. Perhaps observers keep track of the multiple dots by grouping them initially into a single superordinate object and attending to that group as a unitary entity. The designated dots could then be perceived, for example, as the corners of a virtual polygon whose shape changes over time as the dots move. Yantis (1992) has tested predictions from this hypothesis and found support for them in several experiments.

The current debate between object-based and space-based theories of attention often implies that they are mutually exclusive—that one or the other is correct but not both. This would presumably be the case if attention operates at just one level in the visual system. But what if attention operates at multiple levels? At an early image-processing level, such as Marr's primal sketch, a space-based definition of attention is the only thing that makes much sense, because in this low-level representation, coherent "perceptual objects" have not yet been designated. But at a higher level, after organizational processes have identified figures against grounds, objects could certainly be the basis for allocating attention. Both hypotheses may therefore be correct, just at different levels of the visual system.

11.2.4 Selective Attention to Properties

The theories of selective attention we have just discussed—including spotlights, zoom lenses, and even object-based theories—are designed to account for spatial selection. They cannot be the whole story of visual attention if its capabilities extend to selection of different properties, however. When you inspect a prospective purchase at a clothing store, for example, you seem to be able to focus selectively on its color, style (shape), texture, and size as you consider the garment. This ability appears to imply that attention must have important nonspatial components that select for other sorts of properties. Such evidence is anecdotal at best, however. Can people really attend to different properties of the same object independently or does attending to one necessarily result in perceiving them all? In this section we will consider experimental evidence that bears on this question.

The Stroop Effect. Early experiments seemed to indicate that if an object is attended, certain properties are processed automatically, even if the observer is trying to ignore them. This implies that selection by properties is either nonexistent or incomplete.

The best-known evidence for this conclusion comes from the **Stroop effect**, named for the psychologist, J. Ridley Stroop, who discovered it in 1935. The Stroop effect refers to the fact that when subjects are required to name the color of ink in which color words are printed, they show massive interference when the color word itself conflicts with the ink color to be named (Stroop, 1935). Examples of stimuli that produce this effect are shown in Color Plate 11.1.

You can demonstrate the Stroop effect for yourself by timing how long it takes you to name the ink colors in the column of X's on the left (the control condition) versus the conflicting color-word condition in the center versus the compatible color-word condition on the right. Even without timing yourself with a stopwatch, you will find it much more difficult to get through the middle column than the other two. This fact indicates that shape information is being processed automatically whenever the color of the word is attended and that the response to the identity of the word interferes with naming the ink color. This finding seems to imply that, unlike our intuitions based on everyday experience, selective attention to color, independent of shape, may not be possible after all.

One might wonder whether this interference is specific to color. Further research has shown that it is not. Stroop interference occurs, for instance, if the subject's task is to name an object that is drawn in outline around the name of a different object, as illustrated in Figure 11.2.15A. It also occurs when subjects have to name the