CHAPTER 3

Attention

Learning Objectives

- **1.** The Nature and Roles of Attention
  - **1.1.** Failures of Selection
    - **1.1.1.** Failures of Selection in Space
    - **1.1.2.** Failures of Selection in Time
    - **1.1.3.** Sources of Limitation
- **DEBATE:** Cars and Conversation
    - **1.1.4.** Problems in Interpretation
    - **1.1.5.** When the Brain Fails
  - **1.2.** Successes of Selection
    - **1.2.1.** Endogenous and Exogenous Effects in Space
    - **1.2.2.** Components of Attention
    - **1.2.3.** Cross-Modal Links
    - **1.2.4.** Object-Based Attention
- **2.** Explaining Attention: Information Processing Theories
  - **2.1.** Early versus Late Attentional Selection
  - **2.2.** Spotlight Theory
  - **2.3.** Feature Integration Theory and Guided Search
- **3.** Looking to the Brain
  - **3.1.** Electrophysiology and Human Attention
  - **3.2.** Functional Neuroimaging and TMS
- **4.** Competition: A Single Explanatory Framework for Attention?
- **A CLOSER LOOK:** Competition and Selection
- *Revisit and Reflect*

At a very large and very noisy party, you're looking for a friend you've lost in the crowd. You look for her green dress amid the sea of colors. You try to catch the sound of her voice in the general roar. There she is! Somehow, above the loud music and noisy conversation, now you hear her calling your name. But before you can move very far, you are stopped in your tracks by the sound of shattering glass—you turn your head sharply and see that a pitcher has fallen off a nearby table. While others tend to the glass, you set off across the crowded room toward your friend.

The processes by which you were able to spot your friend, hear your name despite the noise of the party, and then quickly turn toward and then away from the sound of the breaking glass involved attention. In the context of human information processing, **attention** is the process that, at a given moment, enhances some information and inhibits other information. The enhancement enables us to select some information for further processing, and the inhibition enables us to set some information aside.

Throughout life—indeed, throughout the day and throughout the minute—we are bombarded by an overwhelming amount of perceptual information; the party is simply a highly dramatic example of what's going on all the time. Our information-processing capacity cannot make sense of the constant input from many sources all at once. How do we cope? How do we manage to keep from being overloaded and thus rendered incapable of action? How do we, moment to moment, choose the information that is meaningful and avoid distraction by irrelevant material? One solution is to focus on some particular piece of information (such as the sound of your own name or a color of interest) and to *select* it for processing in preference to other bits of available information because of its immediate importance in a given situation. Is attention, then, something that we summon up by will that enables concentration on some piece of the incoming stimuli? The answer, briefly, is yes; but this is not the whole story. Even if our intentions and goals are clear and we know exactly what information we are interested in, other aspects of the input, if sufficiently salient, can capture our attention and distract us, as the sudden noise of breaking glass interrupted your search for your friend.

A host of questions immediately arises: while we are attending to one thing, do we actively inhibit and suppress distractions, or do we simply ignore them and let them hover in the background? What happens to the information to which we do not attend? What brain systems and mechanisms underlie these attentional abilities, and what disorders arise when these systems and mechanisms are damaged?

This chapter explores attention as a cognitive ability. We specifically address four issues:

1. What is attention, and how does it operate during cognition?
2. What information-processing models have been devised to understand attention?
3. How have new techniques for studying the brain enhanced our understanding of attention?
4. Attention, according to one contemporary theory, is a competition among different sources of information, all jockeying for further processing. Can such a theory explain both the behavioral and brain perspectives on attention?

1. THE NATURE AND ROLES OF ATTENTION

Although we have an intuitive understanding of what it means to “pay attention” to an object or event, the study of attention has a long and checkered history in cognitive psychology, filled with debate and disagreement. Some have suggested that “everyone knows what attention is;” others have countered that “no one knows what attention is” (Pashler, 1998). For example, Moray (1970) proposed six different meanings of the term *attention*, whereas Posner and Boies (1971) suggested that attention has three components: orienting to sensory events, detecting signals for focused processing, and maintaining a vigilant or alert state. Still others have used terms such as *arousal*, *effort*, *capacity*, *perceptual set*, *control*, and *consciousness* as synonymous with the process of attention. Adding to the difficulty is the problem of designing and carrying out careful and systematic studies of attention, for the very reason that attentional selection seems to occur so naturally and effortlessly that it is difficult to pin down experimentally.

Nonetheless, there is broad agreement that attention involves selecting some information for further processing and inhibiting other information from receiving further processing. One possible way to understand how this might work is to explore what happens when attention fails. After this, we will examine what happens when attention succeeds. Outlining the failures and successes will allow us to develop a clearer idea of what attention is. Thereafter, we will present some theories of attention and some experiments that look at how attention operates in the brain.

1.1. Failures of Selection

When we fail to attend to information, what kind of information do we miss? One sort of failure occurs when there is a lot of information simultaneously present in front of you, as at a party, and you are simply not capable of noticing all of it at once. These failures are referred to as *failures of selection in space*. Failure can also occur with information that unfolds in time. When new information (even if only a small amount) arrives in a rapid stream, spending time processing it will cause you to miss some other incoming information, resulting in what are called *failures of selection in time*. These failures to attend to information in space or in time are a by-product of a system that prevents us from becoming overloaded with irrelevant information—that is, of a system of selective attention. As such, these failures are an important part of effective cognitive processing and highlight the function of attention. Later, when we come to theories of attention, it will be important to remember that understanding attention is as much about information that is not selected as well as information that is selected. In the following subsections, we provide illustrative examples of failures and successes of attentional selection.

1.1.1. Failures of Selection in Space

Failures of selection in space can be of surprising magnitude. You'd notice, wouldn't you, if someone who stopped you on the street to ask for directions suddenly changed into a different person in the middle of the conversation. Actually, you might not. Demonstrations of the failure to detect changes between flashes of the same scene have now been replicated many times. Perhaps the most dramatic of these was a demonstration by Simons and Levin (1998) in which an experimenter stopped pedestrians on a college campus to ask for directions. During each conversation, two people carrying a door walked between the experimenter and the pedestrian. As they did, the experimenter switched places with a second experimenter who had been concealed behind the door as it was being carried. This second experimenter then continued the conversation with the pedestrian. Only half the pedestrians reported noticing the change of speaker—even when they were explicitly asked, "Did you notice that I am not the same person who first approached you to ask for directions?" This failure to detect changes in the physical aspects of a scene has been referred to as **change blindness** (Simons & Rensink, 2005). This phenomenon often goes to the movies: errors in continuity, such as the switch from the breakfast croissant to the pancake in *Pretty Woman*, go unnoticed by many in the audience. We can also be insensitive to changes in modalities other than vision. It has been shown that we miss changes between voices in an auditory scene, a phenomenon referred to as *change deafness* (Vitevitch, 2003).

That we miss some perceptual information is interesting. Even more interesting from a cognitive perspective is the implication that this does not occur by chance: we are *selecting* only partial information from the world around us and are not very attentive to the rest. Change blindness indicates that not all available information is attended and subsequently represented. Fortunately for our evolutionary survival, those aspects of the input that are more relevant and meaningful

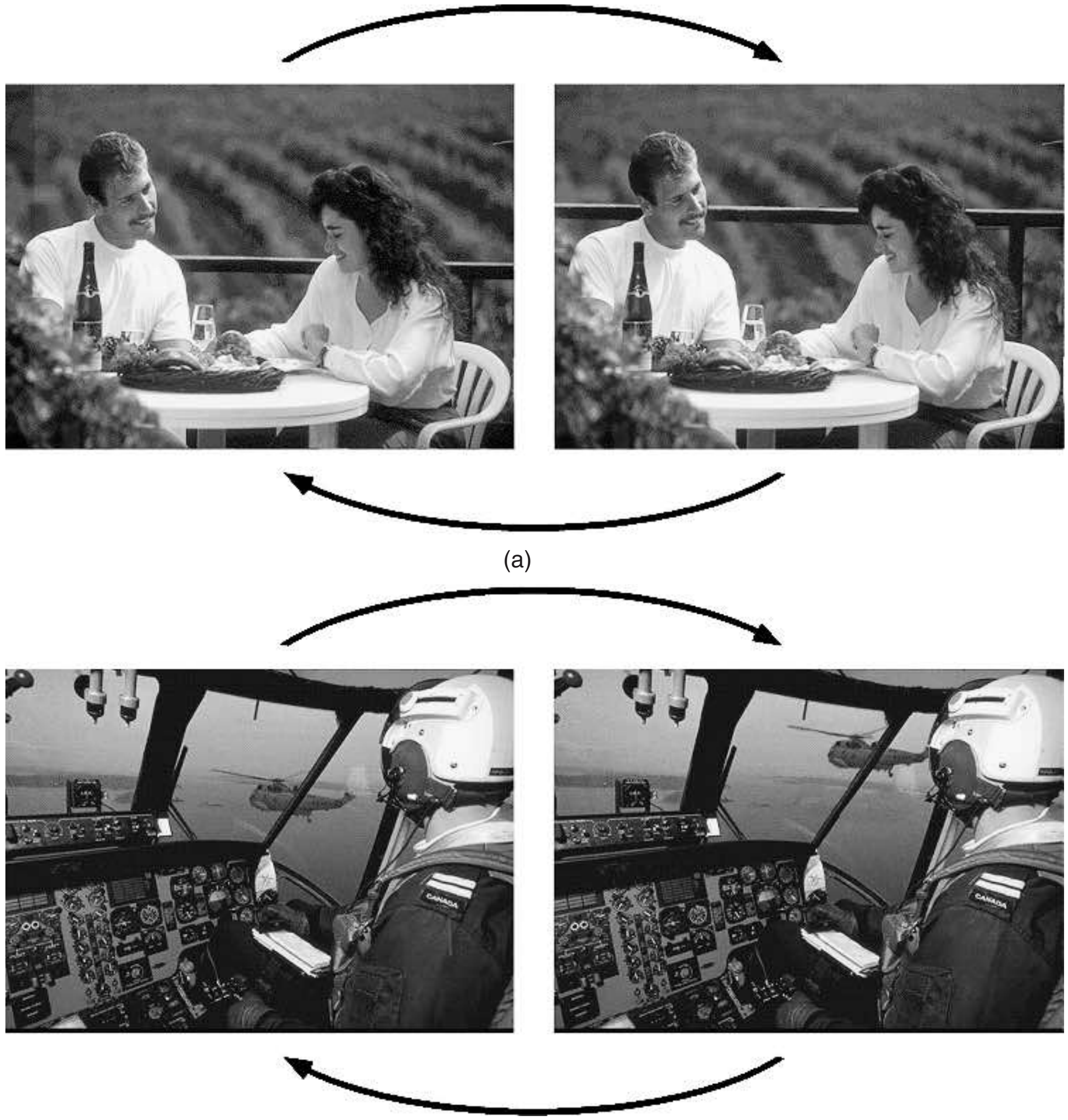


FIGURE 3–1 Changes in the scene

(a) A change of marginal interest (the height of the railing) and (b) a change of central interest (the position of the distant helicopter). Participants required more alternations between the two frames and more time overall to detect a change of marginal interest (average 16.2 alternations, 10.4 seconds) than one of central interest (average 4 alternations, 2.6 seconds).

(From "To See or Not to See: The Need for Attention to Perceive Changes in Scenes," by R. Rensink, K. O'Regan, and J. J. Clark, 1997, *Psychological Science, 8*[5], pp. 368–373. Reprinted with permission.)

may be attended and well noticed, although much other information is not. Rensink and colleagues (1997) showed that changes of "central interest," those relating to the thematic content of a scene, were detected much more quickly than changes of "marginal interest" (Figure 3–1). This finding suggests that although we do extract the most important elements of the visual world, many of the supporting features may be lost.

A further implication is that our attention is driven and controlled via top-down processing, which can change in a flexible and dynamic manner; what is important at one moment may no longer be so at the next, and our goals shift accordingly. If you're hungry, you may notice a basket of luscious-looking fruit on a nearby table—but if you've just eaten, your attention may glide right over it with barely a pause. Knowledge, beliefs, goals and expectations can alter the speed and accuracy of the processes that select meaningful or desired information; that's what is happening when you rapidly scan a book to find a particular reference and are able to skip over large, irrelevant passages of material. The ability to use top-down processing to affect selection and attention is highly adaptive and such processing is an efficient way of extracting critical information from a flood of input.

However, because of the wealth of competing stimuli, top-down attentional selection does not always lead immediately to your goal. For example, in the opening scene at the party, the moment you recognized that the flash of green was your friend's dress was probably not the first moment that patch of green had appeared in your visual field, and the first time you heard your name was probably not the first time your friend had called to you. Further, you were actively diverted from looking for your friend by the sound of the crashing glass—your top-down processing was overridden by a sensory event, that is, by bottom-up attentional processing. The result? Failure in space: attention was captured away from the current goal of seeking your friend.

Failures to select information in space can also occur when far fewer stimuli are present. For instance, if you are presented with only two sources of information simultaneously (say, a drama on television and a story in the newspaper) and are required to process both, you will not be able to do them both full justice. The ability to attend to two sources is impaired compared to the ability to process information from one source alone: there is a cost associated with doing both tasks together. When you try to do both things at once, there are two possibilities: either you will follow the television plot perfectly and lose the news story altogether (or the other way round), or you will lose parts of both show and story.

Concentration on one source of input to the exclusion of any other is known as **focused attention**; in cases of **divided attention**, in which more than one source is attended, the information selected is imperfect (as in the example of following *part* of the newspaper story and *part* of the television story). One explanation for the loss of information when attention is divided is that the two sources of information vie for limited attentional resources, which are sometimes described as "mental effort." An oversimplified but helpful image is that we each have a pool of attentional effort into which each task taps. The harder the tasks, and the more of them there are at any one time, the more of such "mental effort" is drawn from the pool. When the available capacity is less than that required for completion of a task, failures are more frequent. When tasks are easier or fewer, there is less demand on this limited resource.

One clear example of what happens when attention is divided comes from a study conducted by Neisser and Becklen (1975). Participants were shown two

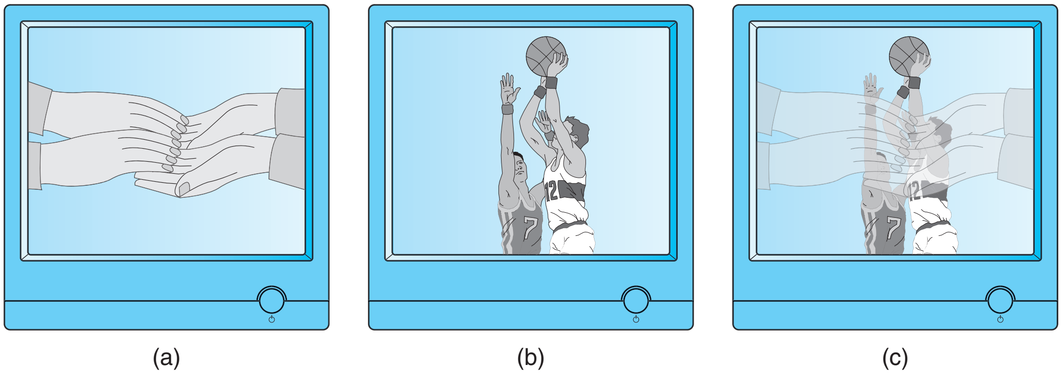


FIGURE 3–2 Divided attention

Drawings of (a) a frame from the video sequence of a hand-slapping game and (b) one from the basketball game. One is superimposed on the other in (c). Participants were shown (c) and asked to track only one of the games. They were successful, although less so than when following either (a) or (b) alone. Tracking both games in (c) proved almost impossible.

(Russell, J. A., and Barrett, L. F. (1999). Core affect, prototypical emotional episodes and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology, 76,* pp. 805–819. Reprinted with permission.)

superimposed video sequences. In one, two people were playing a hand game in which one player tries to slap the opponent's hand; in the other, three men were throwing a basketball and moving about (Figure 3–2). When participants were instructed to track one of the two games, they were successful; but keeping track of both games at once was almost impossible.

A divided-attention task like this one seems, at first blush, artificial. Yet this is exactly what, after much practice with simulations, highly trained air traffic controllers do, with many simultaneous stimuli to monitor (Figure 3–3). Fortunately, those who do this kind of work have sufficient expertise and skill that failures are extremely rare.

1.1.2. Failures of Selection in Time

Just as there are limitations on the quantity of information that can be processed simultaneously in space, there are limitations on the speed with which information can be processed in temporal sequence. These limitations, of differing degree and quality, apply to everyone.

Perhaps the simplest way to determine how fast information can be processed is to ask participants to report the presence of stimuli shown in a rapid sequence. Researchers interested in the question of the temporal constraints of attention have developed experiments that push the attentional processing system to its limit. In such studies (e.g., Shapiro et al., 1984), participants are shown a stream of letters, one of them (denoted by the researchers as the *first target*, or T1, letter) white and the rest

FIGURE 3–3 **Highly divided attention**

Air traffic control personnel are required to track the movements of many planes simultaneously.

of them black (Figure 3–4a). In some of the trials, a second target "X" (denoted as T2 and referred to as a *probe*) was included in the stream of letters at various intervals (either immediately afterward or after a number of intervening letters) following the appearance of the white letter. Each letter was on the screen very briefly, for only 15 milliseconds; the interval between letters was 90 milliseconds. The first part of the experiment was a single task: participants were instructed to ignore T1 (the white letter) and simply indicate whether or not T2 (the probe "X") was present in the sequence of letters. The percentage of correct detection of T2 was recorded as a function of how long after T1 it appeared. Next, in a dual-task condition (one in which two tasks must be performed simultaneously), participants were shown the

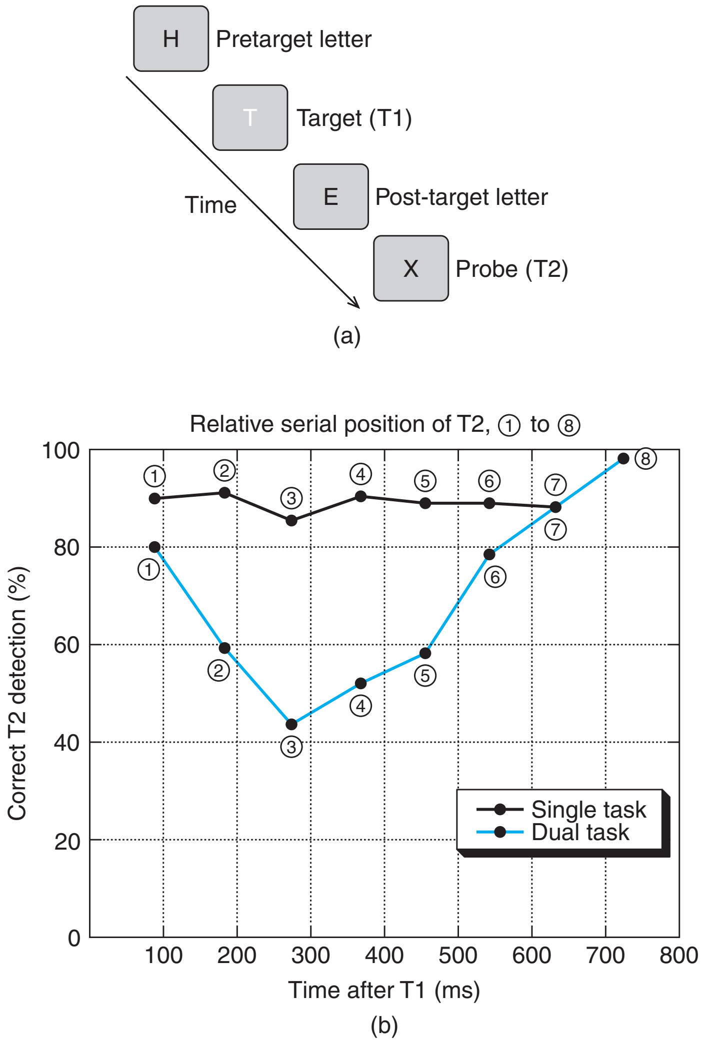


FIGURE 3–4 **Investigating attentional blink**

(a) Target 1 (T1) is white and is embedded in a stream of letters. The probe, the letter "X" (target 2, or T2), is presented at variable serial positions in the stream after the target. (b) The participants were more accurate in the single task (reporting the presence of the X without having to report the white T1 letter) than in the dual task (reporting the presence of an X after correctly reporting the identity of the white T1 letter). The attentional blink occurs after 100 ms, and is present even after a lag of about a half a second—but it is most severe when T2 is presented about 300 ms after T1.

identical stream of letters but this time they were asked to report the presence of T2, as in the single-task case, *and* to identify T1 whenever it appeared.

The results in both conditions are graphed in Figure 3–4b. In the single-task condition, participants consistently did well at detecting T2 regardless of how long after T1 it appeared; this result is not surprising because, following instructions, they ignored T1. The interesting finding is that, in the dual-report condition, participants failed to report T2 on the occasions when it appeared between about 100 and 500 milliseconds after an appearance of T1 (remember, other letters intervene). After greater delays between an appearance of T1 and an appearance of T2, however, participants were again able to spot the T2. The decrease in performance in reporting a T2 if it appears within a certain temporal window following the appearance of a T1 is an instance of attentional blink. As the term suggests, **attentional blink** is a short period during which incoming information is not registered, similar in effect to the physical blanking out of visual information during the blink of an eye. The phenomenon of attentional blink also occurs for two objects (not just letters) that are presented in rapid succession (Raymond, 2003). The hallmark of attentional blink is the missed detection of a stimulus that is presented within a particular time frame after an earlier stimulus is presented. When stimuli are presented so quickly, attention to the first seems to preclude attention to the second—showing the failure to select items in time.

A similar effect involves the failure to detect objects presented in a rapid sequence when some of these stimuli are identical, even when the stimuli are shown for a long enough time to avoid the attentional blink. For example, Kanwisher and colleagues (1997b) showed participants a sequence of nine serially presented displays with two or three consecutive pictures sandwiched in between visually "noisy" patterns called *masks* (Figure 3–5). A large masking field was also shown at the beginning and end of each trial. Each image was shown for 100 milliseconds. The important finding is that when the first and third picture in the series were identical, participants were markedly less likely to report seeing the third picture (the repeat). This was also true when the first and third pictures depicted the same object even if the objects were of

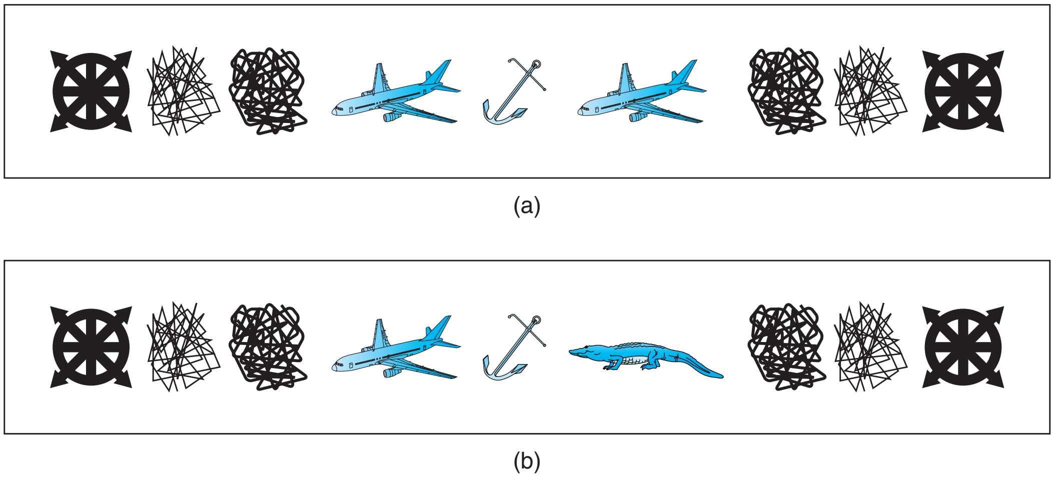


FIGURE 3–5 A demonstration of repetition blindness

Masks and *(center)* representational pictures were shown to participants. (a) When the first and third pictures were identical, participants failed to report the repetition, even if the objects were different in size or shown from different viewpoints. (b) When the first and third picture differed, participants had no difficulty reporting their identities.

different sizes or were shown from different viewpoints. When the two pictures were of different objects, however, participants had no problem identifying the third picture. The failure to detect the later appearance of a stimulus when the stimuli are presented in a rapid sequence has been termed **repetition blindness** (Kanwisher, 1987).

Repetition blindness can occur for words as well as for objects. For example, when the sentence "It was work time so work had to get done" was presented rapidly, participants failed to detect the second occurrence of "work" and recalled the sequence as "It was work time so had to get done" (Kanwisher, 1991). Blindness to a repetition can also be observed if several words come between the two instances of the repeated word or even if they are written in different styles ("WORK" and "work"). It is believed that the failure to encode the second stimulus occurs because it is not individuated or selected as a distinct event when it rapidly follows the first. Instead, the second occurrence is assimilated to the first, and only one event is registered. The phenomenon of repetition blindness suggests that when we do not have much time, we do not form a second, separate representation of an object we have just processed and so are not aware of the repetition.

1.1.3. Sources of Limitation

Why do we fail to select information in space or in time? Some have argued that the failure is on the sensory end; that is, the limitation is, literally, in the eye of the beholder (or, if the stimuli are auditory, in the ear of the beholder). Human peripheral vision is not very precise, and, in many studies, the information that participants miss appears at the edges of the screen. But failures to select all the information present cannot be explained solely by the drop-off in visual acuity for information appearing farther from the center of the visual field. In the Neisser and Becklen study (with superimposed video sequences), for example, all the necessary information appeared in the center of the screen. In the attentional blink and repetition blindness studies, the information is also presented in the center of the field. In these circumstances, then, the failure cannot be one of poor vision. Rather, the limitation appears to have to do with the *quantity* of information. Some models propose the notion of a **bottleneck**, a restriction on the amount of information that can be processed at once; because of the bottleneck, certain critical mental operations have to be carried out sequentially (Pashler & Johnston, 1998).

Divided-attention studies demonstrate that performance is hampered when you have to attend to two separate sources of visual information (like the television screen and the newspaper page) or two separate visual events (the hand-slapping game and the basketball play). There is also an added cost in accuracy or reaction time when you attempt to perform two tasks at once. In all these cases, the decrement in performance is referred to as **dual-task interference**.

One might wonder whether this decrement occurs because there is too much interference when all the information is similar, because it is all visual (or is all uditory) and we simply cannot cope with the quantity of data presented. There is, in fact, greater interference when the sources are both of the same type of information than when they are different (Brooks, 1968). If you are trying to recall a sentence, performing another verbal task such as counting will impair your performance much more than if you are trying to recall a sentence and examine a picture. By the same token, a spatial task such as visualizing a map of the United States and scanning along its borders, noticing which states are wider than they are tall, would impede recall of a picture but not of a sentence.

But the limitation is more general, and the failure to select information can occur even if the two sources of information are of two different types, or even if the information is presented in two different sensory modalities, say, one auditory and one visual, although the interference is not as great as when the types of information are the same. Some investigators believe that the extent of the interference depends on the extent of the "cross-talk" between the various representations and processes drawn upon by the incoming information: if similar representations and processes are activated in the two tasks, they may be confused. For example, both recalling a sentence and counting engage verbal representations; these representations may infiltrate each other and hinder processing. But recalling a sentence and visualizing a map of the United States have less overlap of representations and, hence, produce less interference with each other. A practical issue in dual tasking is discussed in the accompanying *Debate* box.

The bottlenecks in attention we have looked at so far have all been perceptual: too many competing stimuli, or fewer stimuli but of the same type and therefore competing. A bottleneck in attention can also occur when, even with a single sensory input, the outputs required are too great. In such cases, the bottleneck is motor in nature. Consider this: You've just picked up a ringing phone. The only sensory input is the sound of the caller's voice. The call is not for you. Is your roommate around? You have to check. The caller's in a hurry—can you take a message? Sure—where's the pad . . . and while you're looking you've missed the beginning of the message. OK—but now the pencil's dropped . . . gosh, this is taking a long time. . . . One sensory stimulus is requiring a number of responses. Coordinating them will be difficult, and some information will be missed, unless you take all the required responses a little more slowly or in turn. As with multiple sensory inputs, coordinating two output responses is more difficult than simply making a single response. It is not impossible to do two things at once, and, indeed, we can get better at this with practice, but there is usually some associated cost or failure even when one is skilled.

Just as divided-attention failures, in which sensory information is not attended, do not result from a limitation in vision, the failures in motor output when you try to do a number of things at once or in quick succession do not result from a limitation in the ability to program your muscles to move. Back to the party for a moment: you've found your friend and are comfortably chatting. Someone offers you a sandwich, but you're holding a glass in your right hand, your dominant one. You hesitate: should you put the glass down (where?) and take the sandwich with your right hand, or take the sandwich with your left hand? The interference, in the form of a slowing down of your actions, that arises when you try to select between two possible responses to even a sole sensory stimulus is referred to as a **response bottleneck**. The additional time needed to move through this kind of bottleneck has been measured experimentally. In one study, for example, participants were instructed to press a button with the left hand when a light appeared on the left of the

Cars and Conversation

DEBATE

The use of cellular phones has skyrocketed worldwide in the last few years. Recent surveys have shown that about 85 percent of cell-phone users in the United States use a cell phone while driving. Redelmeier and Tibshirani (1997) found that cell-phone use was associated with a fourfold increase in the likelihood of an accident. (Note that, in many states, it is illegal to drive and use a handheld phone, which makes sense given the increased risk of accidents when cell phones are used.) What we do not know from this early study is whether the accidents occur more often when people are dialing a number, when they are talking, or when they are reaching for the phone; that is, whether there is more dual-task interference in some conditions than in others.

More recent studies try to sort out these possibilities. For example, Strayer and Johnston (2001) designed an experiment in which participants undertook a simulated driving task requiring them to "follow" a target car. Participants were assigned randomly to one of three conditions: during the driving task, some participants listened to a radio program of their choosing (group 1); others conversed on a cell phone about the presidential election of 2000 or the Olympic games on either a handheld (group 2) or hands-free (group 3) phone. At irregular intervals, a signal flashed red or green; participants were instructed to press a "brake" button when they saw a red signal. The experimenters recorded both the number of times the participants missed the red signal and the time it took for the individual to press the brake button.

The results were straightforward: both cell-phone groups missed the red signal twice as many times as did the radio group; and when they did spot the signal, their response (measured by the time to press the brake button) came later. This lag in response time was more pronounced when the participant was talking rather than listening. The difference in performance between the radio and cell-phone groups could not be explained by different levels of driving skills in the two groups; all participants performed the driving task alone (no cell phone) and there was no difference in performance between the groups in this single-task condition. It was only the addition of the cell-phone task that led to the different results.

In just what way is attention affected? In a follow-up study, Strayer, Drews, and Johnston (2003) showed that the consequence of adding the task of conducting a phone conversation to the task of driving led the participants to withdraw attention from the visual scene. Participant drivers who held cell-phone conversations missed or had imperfect recall of billboard signs along the route. It turns out that these drivers did not really look at the information: the eye movements of drivers who were talking on the phone were not drawn to information along the route, even when it was presented in the center of the visual field and, consequently, they had poorer memory for that information. These failures to process and select information are very similar to those reported in other experiments on failures to select information; in all cases, the participants could not take in all the visual information that is present and, therefore, focused on only small amounts.

computer screen; if a tone sounded, however, they were also to press a pedal with a foot. Preliminary experiments determined that it took about 500 milliseconds for participants to press the foot pedal after the onset of the tone. If the light flashed (requiring the left-hand response) 50 milliseconds before the tone sounded, participants took even longer than the 500 milliseconds to press the pedal. Response selection for the tone–pedal could not begin until the response selection for the light had been completed, accounting for the additional time to press the pedal (Pashler, 1998).

1.1.4. Problems in Interpretation

Although cognitive psychologists have spent a great deal of time and effort examining divided attention and the costs associated with dual tasks, many questions remain. For one thing, researchers can never guarantee that the two sources of input are always being attended simultaneously, or that the two outputs are always being selected simultaneously, or that the two tasks are always performed simultaneously. Even under conditions in which one task apparently demands constant attention (for example, driving), it appears that the same level of attention is not required at every single moment. An effective strategy for multitasking, then, may be simply to switch quickly back and forth between the two tasks (or inputs or responses) rather than try to deal fully with both simultaneously. But how do you pick your moments? Listening to the car radio for a moment, then watching the road, then listening again is not a very practical (or recommended) procedure, no matter how short each alternate period of attention is! We still do not know whether it is possible to perform two tasks at exactly the same time or, if it is, what burden this arrangement places on the cognitive system.

A second problem that has muddied the dual-task waters is that it is not possible to guarantee that when you do two things at the same time, you're doing them in exactly the same way as you would if you were performing each task singly. Several researchers have suggested that the participants learn to restructure the two tasks and combine them into a single task (Schmidtke & Heuer, 1997; Spelke et al., 1976). If this is in fact what happens, and the dual tasks morph into a single task, it is difficult to separate out and quantify performance on each of the tasks. Also, in this morphing, each task has in some way been altered to make the combination possible, and so comparisons of the cost of dual tasks relative to the single task may not be legitimate.

In any event, dual tasking may not be impossible, and one can develop immunity to its adverse effects by becoming more proficient at one or both tasks. Let's look again at the dual task of using a cell phone and driving. Using simulated conditions, researchers asked drivers to drive normally and, while doing so, to perform a secondary task such as changing the radio station or selecting and calling a number on a cell phone (Wikman et al., 1998). To perform the secondary task, novice drivers frequently looked away from the road for more than 3 seconds at a time, a long (and dangerous) amount of time when you are on the highway. Under the same conditions, experienced drivers glanced away from the road only for brief periods. Because they were proficient at driving and this skill had become more automatic for them, the experienced drivers knew how much time they could spend on the secondary task without greatly affecting their driving. With enough practice and experience, a task can become more automatic and, as a result, less interference will be observed when it is performed in conjunction with another task.

In the late 1970s two researchers used just this terminology and described processing as being either *automatic* or *controlled* (Shiffrin & Schneider, 1977). Just as our example of driving experience suggests, they found that people use automatic processing on easy or very familiar tasks but are forced to use controlled processing for difficult or new tasks. However, controlled tasks can also become automatic with practice over time.

1.1.5. When the Brain Fails

The failures of selection described so far are part and parcel of the human experience—we have all experienced them at one time or another. This normal pattern of failures, however, is massively exaggerated in those who suffer from **hemispatial neglect**, a deficit of attention in which one entire half of a visual scene is simply ignored. The cause of hemispatial neglect is often a stroke that has interrupted the flow of blood to the right parietal lobe, a region of the brain that is thought to be critical in attention and selection (Figure 3–6a on color insert). When these patients are asked to copy or even draw from memory a clock or a daisy (Figure 3–6b on color insert), they do not attend to (that is, they fail to select) information on the side of space opposite the lesion and so do not incorporate this information in their pictures. Similarly, when they are asked to put a mark through all the lines that they see on a page set before them, their results often show gross disregard of information on the left side: they cancel lines far to the right, as if the left part of the image were simply not there (Figure 3–6c on color insert). The reason for failure to select the information on the left (opposite the lesion) is not that they are blind on that side and fail to see the information. Rather, they do not seem to orient toward information on the left side of the scene before them and attend to it. If it is pointed out to them that there is information missing on the left side of their drawing, they may then go back and fill in the missing information; but, left to their own devices, they apparently do not select information from the left side. The neglect is not restricted to visual information (further demonstrating that it is not a visual problem per se)—such patients may also ignore sounds or touch delivered to the left side, or even fail to detect smells delivered to the left nostril.

Hemispatial neglect can make daily life unpleasant and sometimes dangerous. Patients may eat food from only the right side of the dish, ignoring the food on the left, and then complain about being hungry. (If the plate is turned around so the remaining food is on the right side, the problem is rectified.) They may shave or apply makeup to only the right half of the face. They may attend to only the right portion of text, reading a newspaper headline

SPECTACULAR SUNSHINE REPLACES FLOODS IN SOUTHWEST

as

FLOODS IN SOUTHWEST

and may even neglect the left of some words that appear on the right side too, reading the title as

FLOODS IN WEST

They may neglect a left sleeve or slipper, and leave hanging the left earpiece of their eyeglasses (Bartolomeo & Chokron, 2001).

People with hemispatial neglect suffer deficits in their mental imagery as well as in perception. Even when there is no sensory input, and the image is solely created by their own memory, the side opposite the site of damage to the brain is a blank. Bisiach and Luzzatti (1978) demonstrated this by asking a group of hospitalized patients with hemispatial neglect, all of them residents of Milan, to describe in detail the Piazza del Duomo, a landmark of their city. Because the piazza was not in sight, this instruction required the patients to generate a mental image of it in order to describe it. Even though the information was being read off entirely from an internal representation and not from sensory input, the patients reported few details on the side of space opposite the damage. This was not the result of memory failure or forgetting: the experimenters then asked the patients to imagine walking around the piazza and viewing it from the side of the piazza opposite their point of view in their first mental image. Once again, patients neglected the details on the side of space opposite their lesion, but now described the buildings and shops they had ignored in their first description. The finding that hemispatial neglect patients neglect the left of their mental images suggests that attention can operate not only to select information from real perceptual input but can also select from a scene that is internally generated.

There are some situations in which left-sided information can capture the attention of the patient with neglect. Very strong and salient information that appears on the neglected side of the input may successfully capture the patient's attention in a bottom-up fashion; a bright light or a sudden sound on the left side may do the trick. Top-down guidance may also be helpful: specifically instructing the patient to attend to the left may reduce the extent of the neglect, but such guidance must be frequently reiterated.

Although the most obvious deficit in these patients is a failure to attend to left-sided information in space, there is also evidence that there may be a failure to select information in time. For example, Cate and Behrmann (2002) presented letters (such as an "A" or an "S") briefly (100-millisecond exposures) on the left and right sides of a computer screen and asked a patient with left-sided neglect to report which letters appeared. When a letter was presented alone on the left side, the patient identified 88 percent of the letters correctly; this percentage was comparable to the 83 percent reported correctly when the letters appeared alone on the intact right side. The interesting finding occurred when both letters were presented together. (In some patients, it is under these dual presentation conditions that the deficit for the left emerges most strongly.) If the letter on the left appeared first and remained on the screen for about 300 milliseconds before the appearance of the right letter, it was reported correctly almost as often as when it appeared alone. When the right letter appeared ahead of the left letter by 300 milliseconds however, the left letter was correctly reported only about 25 percent of the time. If, however, the right letter was presented first but enough time, say, 900 milliseconds, elapsed before the left letter appeared, the attraction associated with the right letter apparently decayed and the patient was free to attend to the left letter. In this case, detection went back up to about 80 percent. These results are reminiscent of those in the attentional blink experiment.

This study makes two important points: (1) When the left letter is on its own and has no competitor to its right, it can survive the neglect; (2) when a competitor is present on the right, the probability of the left letter's being detected following a short temporal interval is reduced because the patient is still attending to the more rightward letter. If enough time passes, however, the patient can attend to the left letter once again (Husain et al., 1997). Because the time also affects the outcome, it suggests that spatial (left–right) and temporal attentional mechanisms interact in determining how much information is neglected.

Are there patients who neglect information that appears on the *right* side after a stroke to the *left* side of the brain? Yes, but not many. The explanation usually offered lies in the greater, and asymmetric, specialization of brain areas in humans. In humans, the areas involved in processing language are generally in the left hemisphere, and attentional and spatial processes may therefore have been shifted into the right hemisphere. In humans, then, damage to the right hemisphere gives rise to neglect more often and with greater severity than does damage to the left hemisphere.

1.2. Successes of Selection

Fortunately, the normal attention system is not as dumb as it may appear. Despite the many failures of selection we are prone to—which may come about because too much information is present at any one location, or because too much information streams rapidly in time, or because our attention is divided—there are many conditions under which we can successfully and efficiently select the necessary information from the input presented to us.

1.2.1. Endogenous and Exogenous Effects in Space

When you were looking for your friend at the party, two kinds of information affected your search. One came from within you: your knowledge of the color of her dress. The other came from outside you: the sound of breaking glass. (In this case, knowing the color of dress helped; the glass crash distracted.) These two types of sources of information have been found to be highly effective determiners of what information is attended.

When you came into the room, you searched very broadly and rapidly for all green things in the array of colors that surrounded you and then, within this subset of green things, specifically for your friend's dress. This sort of attentional process, which has a voluntary aspect, is top down; it originates from *within* (from your own knowledge, in this case) and hence is called **endogenous attention**. But this form of goal-driven or top-down attention, can be overridden: salient and powerful stimuli can capture attention and direct you away from the task at hand. Attention thus captured is described as **exogenous attention**, because it is driven in a bottom-up fashion by stimuli generated *outside* oneself. (Like the sound of the breaking glass, strong color can provide an exogenous cue, and can be very useful: little children on field trips and prisoners in work-release programs wear brightly colored shirts for the same reason.)

The most systematic studies that examine endogenous and exogenous forms of attention are based on the idea of *covert attention,* which was developed by Hermann von Helmholtz (1821–1894), the German physiologist and physicist. Von Helmholtz demonstrated that although the eyes may be directed at a specific spot, visual attention can be directed elsewhere "covertly," that is, without overt motion of the eyes. (Helmholtz achieved his experimental conditions by flashing a spark of light, which, as a highly salient stimulus, prevented physical motion of his eyes toward the region of space he was attending.)

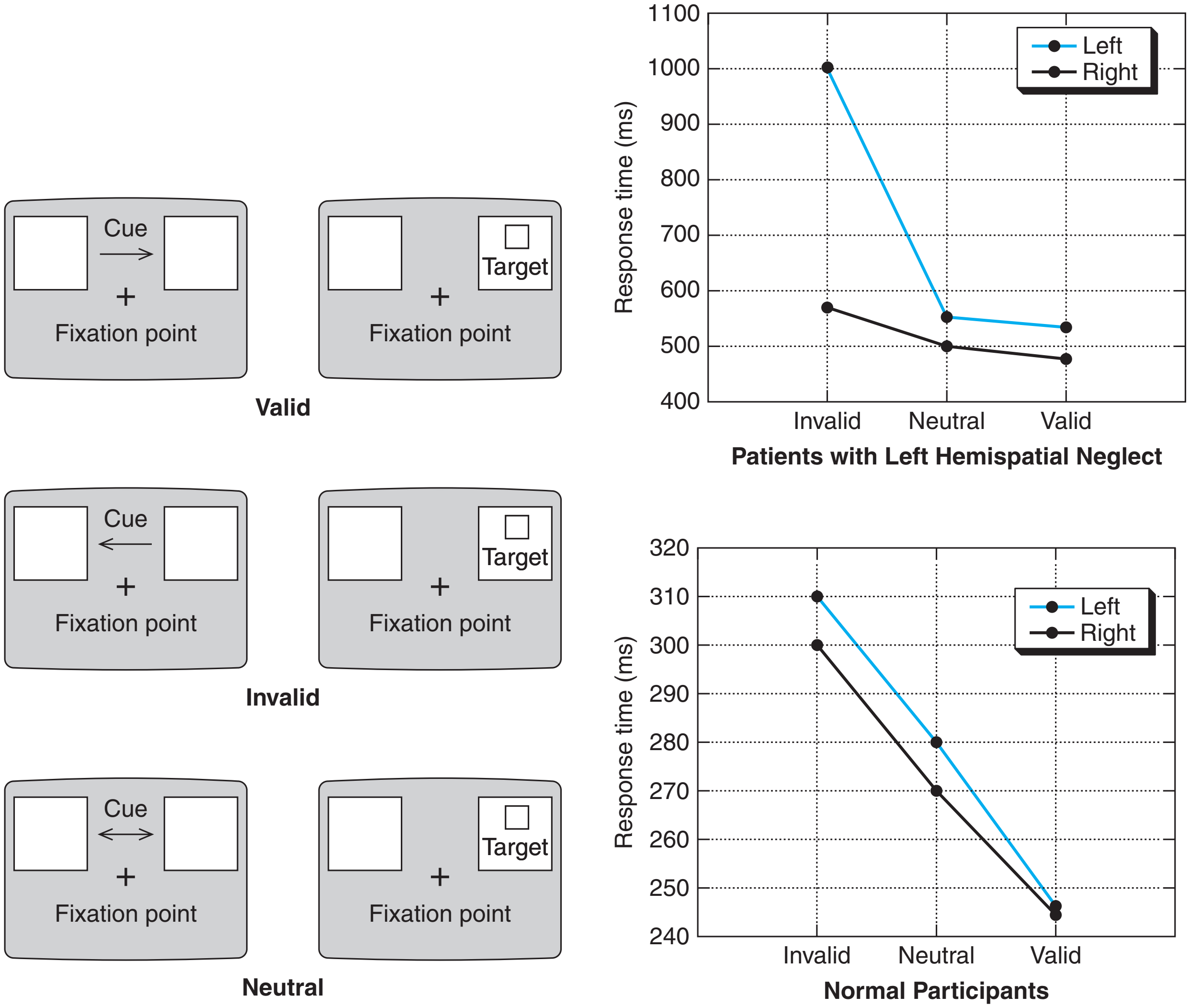
In modern studies, investigators seek to understand how endogenous and exogenous cues influence information processing (Posner et al., 1980, 1982). In one experiment, two boxes appear on a computer screen, one to the right and one to the left of a central fixation point (Figure 3–7a). An endogenous cue such as an arrow leads the participants to focus attention to that location even while their eyes are kept on the fixation point. The arrow is a symbol; only after its meaning is understood does a participant know how to shift attention—and hence attention is controlled by endogenous processes. On a large proportion of trials (usually around 80 percent), designated as "valid trials," a target such as a small box is subsequently presented at the cued location and participants press a response key as soon as they detect the presence of the target. In "invalid trials" an arrow cue appears pointing in the direction opposite the position of the target. Finally, in other, "neutral," trials, the target appears at the same location but the arrow is not informative—it points both leftward and rightward.

This study produced two main results. First, participants detected the target faster (and more accurately) in the valid condition compared to a neutral condition, suggesting that attending to a location can facilitate processing in that location, even in the absence of eye movements. Second, participants detected the target in the invalidly cued location significantly more slowly than in the neutral condition (and, obviously, more slowly than in the valid condition). What was happening? The participants were deceived or misled by the cue in the "wrong" position, which guided their attention in a top-down, endogenous fashion; the subsequent shifting of attention to the other side of the display cost them time. In normal participants, this pattern of costs and benefits is roughly the same whether the target appears on the left or the right of fixation (Figure 3–7b). Data from patients with hemispheric neglect show a very different pattern (see Figure 3–7b).

Attention is also facilitated by a valid cue and inhibited by an invalid cue when the cue is exogenous. An example: this time two boxes are on the screen and one brightens momentarily, presenting a salient bottom-up cue. The target (e.g., a small white box as in Figure 3–7a) then appears in either the cued or the noncued box; in the neutral condition, both boxes brighten. Attention is automatically drawn to the side of the display with the salient bright box. As with endogenous cuing, the participants detect the target faster when it appears in the location indicated by the cue (the valid condition) and detect it more slowly in the invalid condition.

Although the pattern of findings is very similar in the endogenous and exogenous cases, there is one difference. In the exogenous version, attention can be rapidly and automatically drawn toward the powerfully salient brightening cue and no extra processing time is needed. But exploiting the arrow cue in the endogenous version requires that participants process the cue perceptually, understand its content, and then use this information in a top-down fashion. If the arrow appears and the target follows immediately thereafter, participants show neither the benefit nor the cost of the presence of the cue. Given enough time (perhaps 150 milliseconds) to process and implement the information provided by the cue, the facilitation and inhibition emerge clearly.

These studies amply demonstrate that facilitation and inhibition of detection are influenced by the direction of attention alone, without overt eye movement. More recent results suggest that although it is possible under experimental conditions



(a)

(b)

FIGURE 3–7 Endogenous cueing task

(a) In the valid trial, the arrow cue correctly indicates the location of the upcoming target. In the invalid trial, the target appears on the side opposite the cued direction. In the neutral condition, the cue arrow is doubleheaded and therefore not informative about the likely location of the upcoming target. Usually, in such experiments, there are many more valid than invalid trials so participants take advantage of the predictiveness afforded by the cue. (b) Data from normal participants and from patients with hemispatial neglect. Results from normal participants show the advantage—the faster detection time—afforded by the valid cue: target detection in this condition is even better than in the neutral condition. Note that detection time is slowed in the invalid condition, showing the cost when the arrow cue misleads the participant. There is no real difference in detection performance if the target appears on the left or right side. Patients with left hemispatial neglect were substantially slower when the left target was invalidly cued (note the need for a greater ra on the *y* axis). In this case, the arrow cue points to the right and attention is directed to the right. When the target appears on the patient's neglected left side, target detection is very slow.

(such as those described here) to dissociate covert attention and eye movements, under more natural circumstances, the two are tightly linked and may even rely on the same underlying network in the brain (Corbetta & Shulman, 2002). In fact, some researchers have suggested that the coupling of attention and eye movements is particularly advantageous: attention can scout the visual scene first and then the eyes can be moved to regions containing particularly useful or salient information (Hoffman & Subramaniam, 1995).

1.2.2. Components of Attention

What happens when patients with hemispatial neglect perform this attentional cueing task? Patients with lesions to the right parietal lobe, many of whom also suffered hemispatial neglect, detected the valid targets normally on the nonneglected right side and almost normally on the neglected left side (Posner et al., 1984, 1987). That is, they could still take advantage of the cue even though their performance was a little poorer when the target appeared on the neglected side, and this was true whether the cue was endogenous or exogenous. In the invalid trials, when the cue occurred on or pointed to the neglected left side but the target appeared on the nonneglected right side, these patients detected the target more slowly than in the valid trials. The extent of the slowing, however, was in the normal range; neurologically healthy participants, as we have seen, also showed a cost in the reaction time in invalid trials. The important result is that in those invalid trials in which the cue occurred on (or pointed to) the nonneglected right side and the target appeared on the neglected left side, the patients with brain damage needed roughly an additional 500 milliseconds to detect the target.

The findings from patients with brain damage led Posner and colleagues to construct a model for attention that involves three separate mental operations: disengaging of attention from the current location; moving attention to a new location; and engaging attention in a new location to facilitate processing in that location (Posner, 1990; Posner & Cohen, 1984). In the case of the right-hemisphere patients, the model suggests that when the cue directed attention to the nonneglected right side and the target appeared on the neglected left side, the patients had trouble disengaging attention from the good right side, and this deficit produced the dramatically slower target-detection times. No "disengage" problem was apparent for targets on the nonneglected side when the preceding cue indicated the neglected side. In this model then, there are several subcomponents of attention, and results indicate that the parietal lobe (especially on the right) plays a key role in one of them.

Interestingly, Posner and colleagues found other patient groups who appeared to be impaired in either the "move" or "engage" operations posited by their attention model. Patients with damage to the midbrain and suffering from a disorder called *progressive supranuclear palsy* seemed to have no difficulty with "disengage" or "engage" operations (Posner et al., 1985). Rather, they were slow in responding to cued targets in the direction in which they had difficulty orienting, suggesting a problem in *moving attention* to the cued location. On the other hand, patients with lesions to the pulvinar, a part of the thalamus (a subcortical structure), were slow to detect both validly and invalidly cued targets that appeared on the side of space opposite their lesion, but performed well when the targets appeared on the intact side

(Rafal & Posner, 1987). These results led the researchers to suggest that the thalamic patients cannot *engage attention* on the affected side. The different patterns of performance revealed by these three patient groups support the idea that attention can be decomposed into three separate functions (disengage–move–engage) and that each can be selectively affected depending on which brain structures are damaged (for an overview, see Robertson & Rafal, 2000).

1.2.3. Cross-Modal Links

Although many studies have focused on attentional effects in vision, there is also evidence that facilitatory and inhibitory effects can be found within and across different sensory modalities. Once you spotted your friend at the party visually, you suddenly heard her calling your name, but it was likely she'd been calling you for some time. Why did seeing her make her voice more audible?

A series of experiments has demonstrated cross-modal priming under both exogenous and endogenous conditions (Driver & Spence, 1998; Kennett et al., 2002). (As noted in Chapter 2, *priming* occurs when a stimulus or task facilitates processing of a subsequent stimulus or task.) In one experimental design, participants held tactile stimulators that could vibrate at either the thumb or index finger in each hand. Four light-emitting diodes were placed in corresponding locations in space. When participants were asked to report the location of the tactile stimulus, the presence of a noninformative visual flash on the same side of space speeded responses. The reverse condition was also true: a random tactile stimulus primed responses to visual targets on the same side. When participants crossed their hands, the priming was found to be aligned with external space (that is, the left hand crossed to the right side of the body primed detection of a visual stimulus on the right). These effects have also been found between audition and touch, and between audition and vision.

Similar cueing effects have been found when participants expect a stimulus in a location in one modality and an unexpected stimulus appears on the same side of space in a different modality. For example, when expecting a visual stimulus on the right, participants are quicker to detect a random tactile event on that side of space than one on the left. This finding suggests that directing attention to one side of space in one modality automatically results in attention to that location for other modalities as well.

1.2.4. Object-Based Attention

In life, we are surrounded by objects, animate and inanimate, of all sorts, and attention is directed toward them as well as to locations in space and positions in temporal sequence. Recent studies of object-based attention show that when attention is directed toward an object, all the parts of that object are simultaneously selected for processing (e.g., Jaomasz et al., 2005). An immediate example: you think of your friend in green as a single object (and no doubt, because she's a friend, as greater than the sum of her parts). As you focus on her, you would be more likely to notice the watch on her wrist, because it is a part of her, than to notice the watch worn by the person standing next to her, even if that person is standing just as close to you as she is—that watch is a part of a different object.

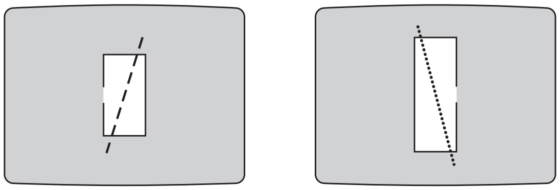


FIGURE 3–8 A behavioral demonstration of object-based attention

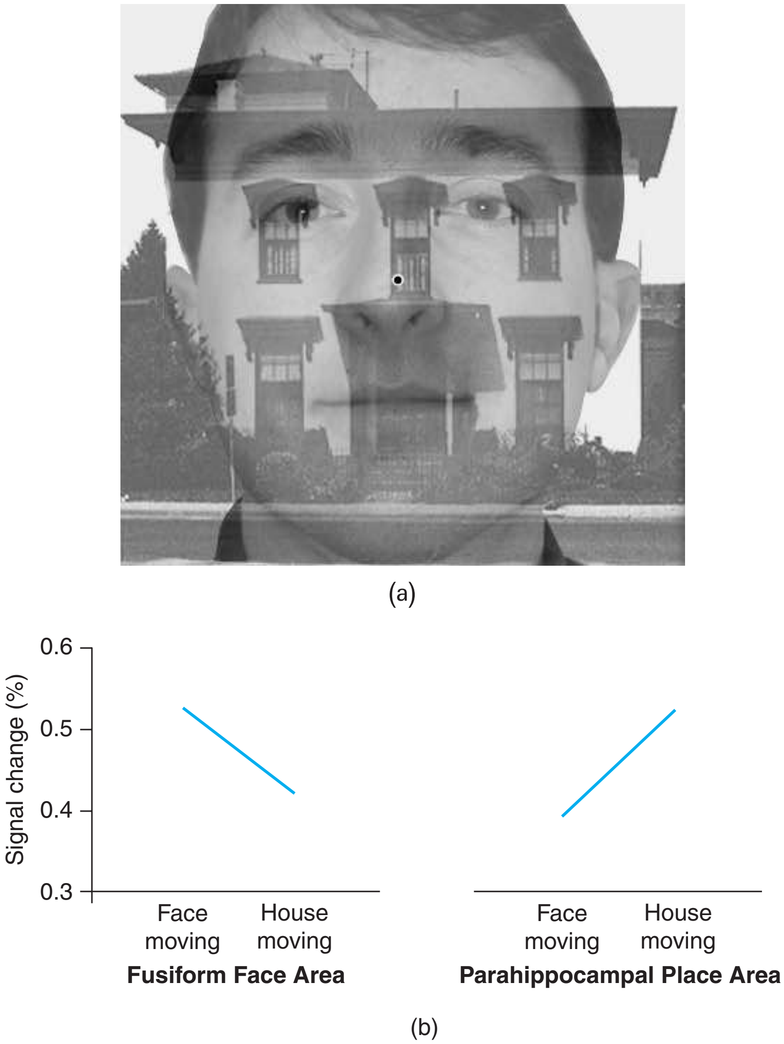
These stimuli were used to test whether two features belonging to the same object are processed better than two features from different objects (details in text). Participants performed better when the two features belonged to the same object, evidence for object-based attentional processing, in which selection of one feature of an object automatically results in the selection of the object's other features.

(Duncan (1984). Image taken from Palmer, S. E. (1999). *Vision Science:* Photons to Phenomenology. Cambridge, MA: The MIT Press. Reprinted with permission.)

Many studies of object-based attention demonstrate that an object and its associated parts and features are selected together. In one of the best-known studies (Duncan, 1984), participants saw in the center of a computer screen a rectangular box with a gap in one side and a line through the box (Figure 3–8). When they were instructed to respond with two judgments about a single object—whether the box was big or small in relation to the frame of the screen and whether the gap was on the left or right side—accuracy of report was high. In fact, participants were almost as good at reporting on the two features as when they were required to make only a single judgment, on box size *or* gap side. Similar results were obtained when participants were asked to make the two judgments about the line itself—was it slanted or upright, was it dashed or dotted? In a further condition, the two judgments the participants were asked to make concerned the box *and* the line, for example, the size of the box and the texture of the line. This time, although again no more than two judgments were required, accuracy fell significantly.

The important aspect of this study is that both objects—box and line—were superimposed one on the other in the center of the screen, thus occupying the same spatial location. The results with one object (box *or* line) and two objects (box *and* line) cannot be explained by preferential attention to a particular location in space. Instead, the result is compatible with the idea of object-based attention. Apparently, our perceptual system can handle two judgments quite well when attention is focused on a single object. When attention must be divided across two different objects, however, making two judgments becomes very difficult and performance suffers badly (this is similar to the cost that occurs under dual-task conditions). These findings support the idea that attention can be directed to a single object and all features of the object attended.

Neuroimaging has confirmed these behavioral results: when we attend to one aspect of an object, we perforce select the entire object and all its features (O'Craven et al., 1999). Participants were shown pictures of a semitransparent face superimposed on a semitransparent house and instructed to fixate on the dot in the center of the superimposed images (Figure 3–9). On each trial, the position of either



(a)

(b)

FIGURE 3–9 fMRI support for object-based attention

(a) Example of overlapping stimuli used in an fMRI experiment investigating the effect of attention to faces and to houses on activation in brain regions thought to specialize in processing of one or the other of those objects. (b) Changing levels of activation in the fusiform face area and the parahippocampal place area when the participants attended to motion and the moving object was a face and when it was a house.

(O'Craven, K., Downing, P. E. and Kanwisher, N. (1999). fMRI evidence for objects as the units of attentional selection. *Nature, 401,* pp. 548–587. Reprinted with permission.)

the house or the face shifted; participants were instructed to attend only to the house, only to the face, or to the shift in position, and their brain activation was measured by functional magnetic resonance imaging (fMRI). This study exploited the fact that different regions of the brain respond more to houses or buildings (the parahippocampal place area), to faces (the fusiform face area), and to motion (area

MT). If attention were directed toward a spatial location, then we might have expected to see activation in all three brain regions, corresponding to the three stimulus types, given that all three occurred in the same region of space. If, however, selectively attending to one attribute of an object also enhances processing of other attributes of the same object, then we would expect greater activation in the brain area representing co-occurring attributes of the selected object. And, indeed, that was the case. When motion was selectively attended, not only was MT activated, as expected, but the area representing the attended object (face or house) was also activated. Thus, for example, the fusiform gyrus face area was activated when the face was moving compared to when it was static, even though the face itself was not preferentially attended. The same was true for the parahippocampal place area when the house moved. This suggests that in object attention, more than one feature is simultaneously selected (say, house-and-moving or face-and-moving) and the corresponding neural areas reflect this coactivation.

Evidence from brain damage also supports the notion of object-based attention. Although hemispatial neglect is predominantly thought of as a deficit in processing in which attention to the left side of space after a right-hemisphere lesion is demonstrably poor, it has been shown that the left side of individual objects is also neglected. In one study with neglect patients (Behrmann & Tipper, 1994; Tipper & Behrmann, 1996), participants were shown a static schematic image of a barbell, with each end a different color (Figure 3–10). Participants were instructed to press a key when they detected a small white flashing light in either the left or right end of the barbell. As expected of patients with left-sided neglect, detection on the left was poorer than on the right. But was this because the target appeared on the left of *space* (space-based neglect) or because the target appeared on the left of the *object* (object-based neglect)?

In a second condition, while participants watched, the barbell rotated so that the original left end (as identified by its color) appeared on the right of space and the right end of the barbell appeared on the left of space. Surprisingly, when the target—the flashing light—appeared on participants' "good" side, the right of space (but in the end of the barbell that had previously appeared on the left), participants took longer to detect its presence. The poorer detection in the rotating case occurred presumably because the target fell on what had been the left end of the barbell. Similarly, when the target appeared on the left of space, performance was better than in the static condition, because it now fell on the right end of the object. A further, and important, finding was that when the two circles that depicted the ends of the barbell were not joined by a connecting bar, participants' detection of targets on the right of space was always good and detection of targets on the left of space was always poor. In other words, no object-based neglect occurred, presumably because the two circles were not perceived as two ends of a single object; therefore, only space- and not object-based attention operated.

An even more extreme case of object-based selection—or rather its deficit—can be seen in patients with Bálint's syndrome (for more information about this syndrome, see Rafal, 2001). This neurological disorder follows after bilateral (that is, on both sides of the brain) damage to the parietal-occipital region; it is sometimes referred to

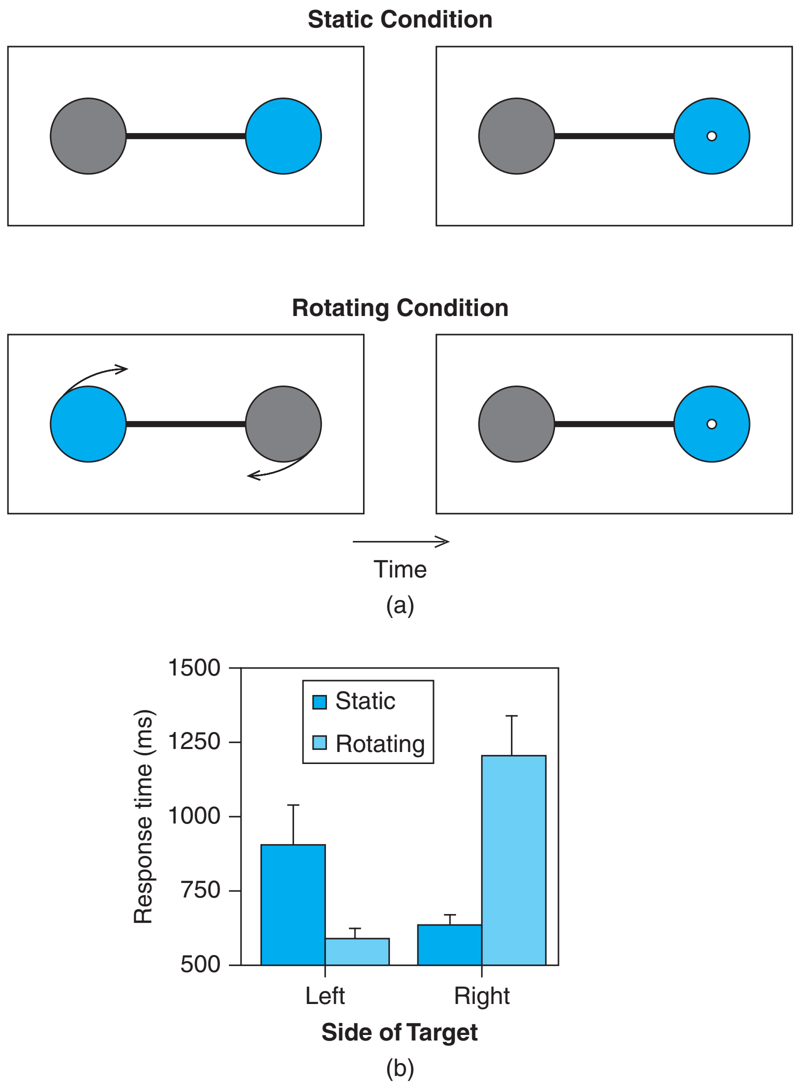


FIGURE 3–10 A demonstration of object-based hemispatial neglect

(a) The barbell display used to assess neglect for the relative left of a single object in people with right-hemisphere lesions and left-sided neglect (see text for details). (b) The crucial result: in comparison with the static condition, performance in the rotating condition is poorer (that is, the time to detect the probe is longer) when the probe appears on the right of space and better when it appears on the left of space. The decrement in performance on the right of space and improvement on the left is attributed to the fact that the probe is appearing on the left and right, respectively, of the *object.*

(Modified from Behrmann, M. and Tipper S. (1994). Object-based attention mechanisms: Evidence from patients with unilateral neglect. In: Imilta, C. & Moscovitch, M. (eds), *Attention and Performance XV: Conscious and nonconscious processing and cognitive functions.* Cambridge, MA: The MIT Press. Reprinted with permission.)

as simultanagnosia. (*Agnosia* is a defect in recognition; *simultanagnosia* is the inability to recognize two things at the same time.) Bálint's patients neglect entire objects, not just one side of an object as in hemispatial neglect. The disorder affects the selection of entire objects irrespective of where the objects appear in the display, and a whole object (a line drawing) may be neglected even if it occupies the same spatial position as another drawn object. These patients are able to perceive only one object at a time; it is as if one object captures attention and precludes processing of any others.

However, the failure to select more than one object can be reduced if the objects are grouped perceptually. In one such study (Figure 3–11), Humphreys and Riddoch (1993) had two patients with simultanagnosia view a display containing colored circles (each itself an object). On some trials, the circles were all the same color and on other trials half the circles were one color, half another. The patients were to report whether the circles in a display were all the same color or were of two different colors. In some displays, the circles were unconnected (the random condition); in others, circles of the same color were connected by lines, and in still others (the single condition), circles of different colors were connected (the mixed condition). In displays of unconnected circles, the patients found it difficult to judge color, especially when there were two circles of different color—there were just too many different objects. If, however, the two differently colored circles formed a single object by virtue of a connecting line, some of the difficulty of attending to each circle separately was offset, and the patients did better in the mixed condition than in the other two conditions. The cost of dividing attention between differently colored circles was reduced when these circles were joined to make a single object, and this improvement was more dramatic than in either the random or single condition.

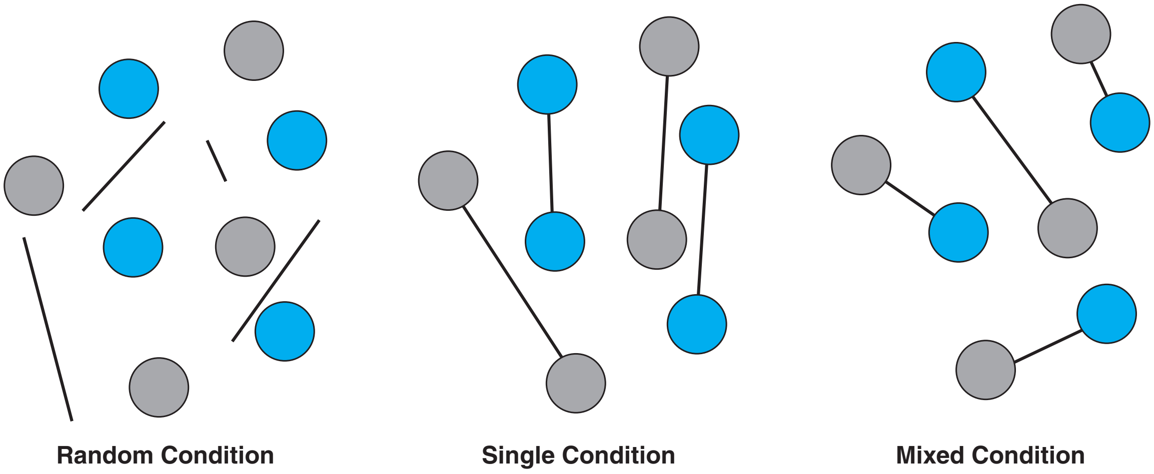


FIGURE 3–11 Simultanagnosia

Patients with simultanagnosia can attend to only a single object at a time; these patients can be helped when information in the visual scene is bound together to make a single object. Patients were shown displays with circles of two different colors, some disconnected (the random condition), some linked according to color (the single condition), and some in which differently colored circles were connected (the mixed condition). Joining the circles (thus making them into single objects) made it easier for these participants to perform the task of judging whether one or two colors were present; joining circles of two different colors (the mixed condition) led to the most improved performance.

✓ Comprehension Check:

1. Selection of information can fail when too much information is present at any one time or when information comes in faster than we can process it. What are some examples of failures in space and in time?
2. What are the distinctions between exogenous and endogenous forms of attention?

2. EXPLAINING ATTENTION: INFORMATION-PROCESSING THEORIES

"Paying attention" encompasses dynamic processes that involve the enhancement or selection of particular information and the inhibition of other information. Attention may be thought of as a mechanism that controls processing so that we are not overwhelmed by too much information. Endogenous factors such as one's knowledge and goals and exogenous factors such as the salience of external information can influence the selection. But how does attention work? A number of different information-processing theories have attempted to capture the dynamics of attentional effects. Although none of these theories can explain all the attentional phenomena we have described so far, the theories offer important insights into the principles that underlie attentional effects.

2.1. Early versus Late Attentional Selection

Almost all the experiments described here show that we can attend only to some of the sensory information that surrounds us and not to all of it. In the language of information processing, this selective aspect of attention is often a consequence of inadequate channel capacity or a fundamental limitation in the flow of information. One question is, when does selection occur, early or late in processing? Where is the bottleneck? How much and what kind of information is processed before it, how much and what kind after it? That essentially is the problem of "early versus late attentional selection."

The British psychologist Donald Broadbent (1926–1993) favored the view that selection is made at an early stage of processing. He proposed a model of the attentional system as containing a limited-capacity channel through which only a certain amount of information could pass (Broadbent, 1958). The many sensory inputs capable of entering later phases of processing, he believed, had to be screened to let only the most important information through. In his view, at an early stage of processing, information comes into a very brief sensory store in which physical characteristics of the input are analyzed: in the visual modality, these characteristics are motion, color, shape, spatial location; in the auditory modality, pitch, loudness, and, again, spatial location. Broadbent argued that the bottleneck is immediately after the sensory store, and only a small amount of information, selected on the basis of physical characteristics, passes through for further, semantic processing.

Broadbent's ideas were well received at the time; they successfully accounted for a range of empirical evidence. Some of this evidence had been presented by E. Colin Cherry (1953), another British psychologist, who recruited volunteers to participate in an auditory experiment. Using a technique called *dichotic listening* (the literal meaning is listening with "two ears"), he played competing speech inputs through headphones to the two ears of his participants. For example, the right ear might receive "the steamboat chugged into the harbor" while the left ear simultaneously received "the schoolyard was filled with children." Cherry instructed participants to "shadow," that is, to follow and repeat as rapidly as possible one stream of speech input and to ignore the other. Cherry found that participants had no memory of what was played in the unattended ear; in fact, they did not even notice if the unattended message switched to another language or if the message was played backward. They did, however, notice whether the sex of the speaker was different or whether the speech became a pure tone.

Cherry's results (1953) are consistent with the early-selection bottleneck theory: unattended inputs are filtered out and attended signals are admitted through on the basis of their physical characteristics. Changes in the physical aspects of a stimulus were attended, but if there were no such changes, the stimulus would either be attended or filtered out. Consistent with the claim that unattended stimuli are filtered, when the same word list was played to the unattended ear of participants 35 times (Moray, 1959), the participants never noticed. The failure to detect the repeated word lists indicates that the unattended signals were not processed deeply and the participants did not have a representation of the words or their meaning.

But one important piece of evidence suggests that a theory of early selection cannot be the whole story. Only a theory of late selection—which holds that, before the bottleneck, *all* information is processed perceptually to determine both physical characteristics *and* semantic content—could account for the finding that some information *could* be detected in the unattended channel even when there was no change in its physical features. This was especially true if the information was salient and important to the participant. Hearing your friend call your name above the din of the party is a good example of how unattended but high-priority information can still be detected. Hearing your name at a loud party is such a good example of this phenomenon that it's known as the **cocktail party effect**. By early-selection views, the cocktail party effect should not be possible; but there it is. Because it now seemed that unattended inputs were able to intrude and capture attention, Broadbent's ideas had to be modified.

Additional evidence to support late selection came from a number of studies using dichotic listening. In one (Treisman, 1960), a different message was played into each ear of participants. The logical content of each, however, was confused: the left ear heard "If you are creaming butter and *piccolos, clarinets, and tubas seldom play solos*"; the right ear heard "Many orchestral instruments, such as *sugar, it's a good idea to use a low mixer speed*." Participants were told to shadow the right ear, but some must have switched channels to follow meaning: they reported a reasonable sentence about an orchestra, and believed they had shadowed the correct ear all along.

The late-selection idea was also tested by presenting a participant's own name on the unattended channel, a controlled equivalent of the cocktail party effect (Wood & Cowan, 1995). About one-third of the participants reported hearing their own name (and none reported hearing a different name). This finding is difficult to accommodate within the early-selection view; it is also difficult to accommodate it entirely within the late-selection view, because only one-third of the participants detected their names on the unattended channel. One possible explanation is that this one-third occasionally switched attention to the unattended channel. This may indeed be what happened: when Wood and Cowan instructed participants ahead of time to be ready for new instructions during the task, 80 percent of the participants now heard their name on the unattended channel. This finding undermines late selection and suggests that participants may, for one reason or another, switch attention to the other channel despite instructions.

How can we reconcile these various results? One suggestion is that some kind of analysis must be done before the bottleneck, or filter, so that one's own name or other salient information can pass through (Moray, 1970). Arguing against this, another suggestion holds that the early-selection view requires only a slight modification (Treisman, 1969): that in each person's "dictionary," or lexical store, some words have lower thresholds of activation than others. Thus, information is still filtered out early, but words that are well known by the listener are more easily detected and require less analysis—and hence the information that does get through the filter is sufficient. Thus, one's own name or a shouted "Fire!" would appear to pass through the bottleneck and capture the listener's attention. Also, words that are highly probable given the semantic context (such as *piccolo* following shortly after *instruments*) may also pass through to our awareness.

2.2. Spotlight Theory

Like a spotlight that highlights information within its beam, in this view, spatial attention selectively brings information within a circumscribed region of space to awareness, and information outside that region is more likely to be ignored. This metaphor works—up to a point.

Consistent with the idea that spatial locations can be enhanced when they fall in and around the spotlight, participants who correctly named letters appearing at multiple locations in the visual field were more likely to succeed at an orientation discrimination task when the forms to be discriminated appeared near the letters. These data (Hoffman & Nelson, 1981) suggest that information is enhanced when it appears near the current position of the spotlight.

But the spotlight metaphor breaks down. For one thing, a number of experiments, discussed earlier, have shown that attention can be directed to a single object, even if superimposed on another object, demolishing the idea that a "spotlight" of attention highlights information in a particular spatial region. If that were true, all objects would be selected together, but we know that one object can be preferentially selected. Another difficulty is the assumption of the spotlight model that the beam of attention sweeps through space. If that's what happens, one would expect that if an obstacle intervened in the course of the sweep, attention would be captured or hampered by this obstacle. But it isn't. In a study that investigated this expectation (Sperling & Weichselgartner, 1995), participants monitored a stream of *digits* appearing at a fixation point. At the same time, they were to attend to a rapidly changing stream of *letters* appearing to the left of the fixation point and to report any appearance of the letter "C." Occasionally another character appeared between the fixation stream and the letters on the left—that is, it interfered with any "spotlight beam." But this "interference" made no difference: whether or not intervening information appeared, the time taken to detect a letter "C" was constant. These results suggest that attention is not influenced by the presence of spatially intervening information, as would be expected from a spotlight model.

Rather than thinking about attention as a spotlight where information outside the selected region is simply ignored, more recent studies have begun to characterize attention as a dynamic process in which information selection is automatically accompanied by active inhibition of other information. Thus, rather than as a spotlight, attention can be understood as a competitive system in which tuning into one thing results in the inhibition of competing information. Attending to green things will result in inhibition of things of other colors (helping you find your friend on our party example), attending to a single object such as a person will result in inhibition of other people or objects, attending closely to music may inhibit to some degree unrelated visual information. Attention, then, is a dynamic push–pull process that involves both increasing and decreasing the likelihood that certain locations or objects will be processed in detail. These ideas are elaborated in greater detail in the final section of this chapter.

2.3. Feature Integration Theory and Guided Search

This theory, which has a very different emphasis from ideas of bottlenecks, filters, and spotlights, is mostly concerned with the role attention plays in selecting and binding complex information. This question has been particularly well investigated in experiments using a visual search task. In this design, a display is presented on a computer screen; participants are instructed to find a target piece of information and press a response key when they do. For example, participants may be instructed to search for the circle in a display such as that shown in Figure 3–12a. In

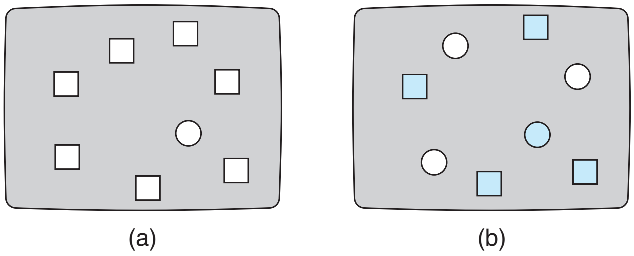


FIGURE 3–12 Selecting and binding complex information

Schematic depictions of (a) a disjunctive visual search display and (b) a conjunctive one. In which is it easier to spot the "odd-man-out"?

a separate block of trials, participants may search for the colored circle in a display such as that shown in Figure 3–12b. Try it: the difference between these two types of trials will be immediately apparent. It is absurdly easy to locate the target on the left, but the display on the right poses more difficulty. Displays like that on the left are referred to as **disjunctive** (or **feature**) **search** trials. In these trials the target differs from the other characters or symbols—the distractors—by a single feature, such as shape (circle among squares). A **distractor** is a nonrelevant stimulus that is supposed to be ignored. Displays such as the one in Figure 3–12b are **conjunctive search** trials, in which the target is defined by a conjunction of features—here, (blue versus white) and shape (circle versus square).

Your experience with Figure 3–12 no doubt suggested that disjunctive search is generally easier; you're right. Even increasing the number of elements in the display in disjunctive search does not slow down target detection—the search can be done effortlessly and rapidly. The target seems to pop out at you; this kind of search is described as *preattentive;* that is, it takes place before attention is engaged. Because the target pops out regardless of the number of elements in the display, it is likely that the search is conducted in parallel across the display; that is, all elements are evaluated at the same time. In conjunctive search, however, each element must be attended and evaluated individually to determine whether or not it is the target. Adding more elements to conjunctive search slows you down substantially and, in fact, an additional increment of time is required to detect the target for each additional item included in the display. Because you must examine each item serially to see whether it has the requisite conjunction of attributes, the time to find the target increases dramatically as the number of distractors increases.

The cognitive difference between disjunctive and conjunctive search is well captured by *feature integration theory* (FIT) (Treisman & Gelade, 1980). According to FIT, the perceptual system is divided into separate maps, each of which registers the presence of a different visual feature: color, edges, shapes. Each map contains information about the location of the features it represents. Thus, the shape map of the display in Figure 3–12a would contain information about something of a particular shape at the right of the screen. If you know that you are looking for a target defined by that shape, you need only refer to the shape map, which contains all the shapes present in the display. The shape you are looking for will pop out of this shape map and target detection proceeds apace, irrespective of the number of distractors of another shape. Looking for a conjunctive target, however, requires the joint consultation of two maps, the shape map and the color map. FIT suggests that attention is required to compare the content of the two maps and serves as a kind of glue to bind the unlinked features of, say, "COLORness" [blue] and "circleness" to yield a [blue] circle.

Feature integration theory has illuminated other aspects of the way attention operates in visual search. One important finding is that you can search faster for the *presence* of a feature than for its *absence;* participants are able to find the "Q" (essentially a circle *with* a tail) among the "O"s in Figure 3–13a much faster than the "O" (essentially a circle *without* a tail) among the "Q"s in Figure 3–13b. In fact,

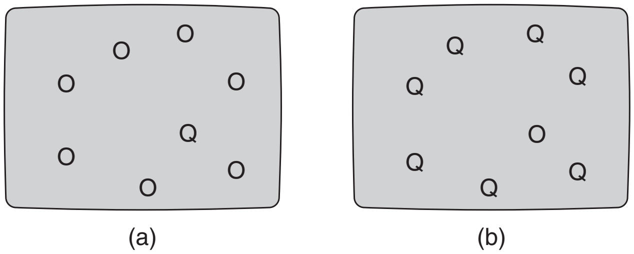


FIGURE 3–13 **Looking for an absence**

Schematic depictions of a visual search display in which the target has the critical feature (a) present or (b) absent. Participants found that spotting the "Q" among the "O"s—the element *with* a feature (a tail)—took less time than seeing the "O" among the "Q"s—the element *without* a feature.

search time for the target "O" increased dramatically as the number of "Q" distractors increased, but it did *not* increase when more "O" distractors were added to the display surrounding the "Q" (Treisman & Souther, 1985).

FIT is also supported by some of the sorts of errors that occur when attention is overloaded or selection fails. For example, participants sometimes make **illusory conjunctions**, that is, incorrect combinations of features. For instance, if participants report on the elements present in Figure 3–12b, if the display is presented very briefly, they may report the presence of a white square. This response, incorrectly combining features (COLORness and squareness) present in the display, suggests that these features are registered separately but that they are not properly bound together. When attention is overloaded or the features are not selected together, the isolated features remain unbound and may be incorrectly attached to other features (Treisman & Schmidt, 1982).

Do brain studies support the distinction between disjunctive and conjunctive processes? Some neuroimaging studies have indicated that different types of features really are being registered by partially distinct neural mechanisms, as assumed by feature integration theory. But the evidence is not incontrovertible, and some findings from patients with hemispatial neglect present a challenge to FIT. Patients with hemispatial neglect fail to take into account information on the side of space opposite the lesion, and the deficit has been assumed to be one of failing to attend to that side. According to FIT, disjunctive search is preattentive and does not engage attention, whereas conjunctive search does involve attention. If this distinction is correct, then one might predict that neglect patients would be able to perform disjunctive search well, even when the target appears on the neglected side. The findings suggest that this is not true. Behrmann and colleagues (2003) tested the visual search performance of a large group of neglect patients on the "Q"s and "O"s search task, manipulating the number of elements in the display from 1 to 16 items. As expected, these patients took a very long time, compared with control participants, to detect the presence of the "O" among the "Q"s on the right side (remember, this is the more difficult search, for the *absence* of a feature). They were also impaired in their search for the “Q” among the “O”s when the target “Q” was located on the left of the visual display. The “Q” did not pop out for these patients: either they failed to locate the target or took a very long time to do so, suggesting that even disjunctive search may require attention and the preattentive–attentive distinction between these forms of search may not hold.

Additionally, even behavioral studies with neurologically unimpaired participants have found that some conjunctions are easier to detect than a purely serial search model predicts (Nakayama & Silverman, 1986). Consequently, a new theory, *guided search,* was proposed (Wolfe, 2003; Wolfe et al., 1989). As indicated by the name, output from a first stage of information processing “guides” later serial search mechanisms. Although the first stage is similar to that of FIT in that it is constructed of different feature maps, it differs in that items that cannot possibly be the target are eliminated in parallel in the feature maps. In the example of Figure 3–12, processing within the color feature map would label all the white items as distractors and all the color items as potential targets. The same sort of labeling would occur for the square versus circle stimuli within a shape feature map. Thus, by the time information reaches the second attentive stage, the number of candidate targets is already much reduced compared to the total number of items possessing one feature of the target. Guided searching accounts for a relatively efficient search of conjunction targets by allowing information from the preattentive feature stage to reduce the number of items over which attentionally demanding serial searches occur.

✓ Comprehension Check:

1. What are the differences between a spotlight view of attention and a feature integration view of attention?
2. Distinguish between “attention operating early” and “attention operating late” and provide examples of studies that support each of these two views.

3. LOOKING TO THE BRAIN

The study of attention has become a very hot topic in the twenty-first century because of a number of very successful studies of the neural basis of attention. These in turn have furthered our knowledge of the mechanisms that give rise to attention. For example, it is now a well established behavioral finding that when attention is directed to a location, perception of that information is enhanced. Until fairly recently, however, we did not know whether this was because the target at the cued location was more efficiently processed in the visual areas of the brain or because the motor system was biased to produce faster responses. Both these explanations are reasonable accounts of the findings of faster target detection. To explore further, researchers have conducted attention studies with animals and humans using various biological methods.

3.1. Electrophysiology and Human Attention

In the late 1960s, technology was developed that allowed researchers to measure with considerable precision the variation in electrical activity generated by the brain. Although it was known that faint waves are emitted by the brain in response to a stimulus, it had not previously been possible to average these tiny signals and relate them specifically to the processing of that stimulus. More sensitive electrodes placed on the scalp and more powerful computers to make the calculations could do the job. With technological advances, it became possible to distinguish between an event-related potential (ERP)—the change in electrical activity in response to a stimulus—and the sorts of brain activity that go on all the time. Investigators were now able to explore the neural mechanisms associated with various cognitive processes, including the phenomenon of selective attention.

The major result from some of these ERP studies was that directing attention toward a stimulus results in an increase in the amplitude of the waveform as early as 70 to 90 milliseconds after the onset of the stimulus. These changes are recorded in the first positive, or P1, wave over lateral occipital regions of the scalp (in the visual system) and suggest that attention enhances the early processing of visual stimuli in the brain, which leads to better perceptual detection of the attended target stimulus. For example, a study that recorded ERPs during tasks involving covert attentional cueing (such as that depicted in Figure 3–7) found a difference in the P1 waveform (and also in the first negative wave, N1) between cued and uncued trials for targets in both the left and right visual fields. Larger sensory ERPs were recorded in early stages of visual processing when the targets were in cued locations (Mangun & Hillyard, 1991). Attending to a location apparently increased the amount of visual processing, giving rise to a larger ERP signal. A similar increase in the sensory ERP occurs when attention is exogenously drawn to a location if the target appears within 300 milliseconds of the exogenous cue. Again, an early occipital wave is enhanced, consistent with enhanced visual processing of the target.

Taken together, these results suggest that exogenous, automatic attention and endogenous, voluntary attention (in other words, bottom-up and top-down forms of attention) have at least some underlying processes in common, an implication consistent with the behavioral findings discussed earlier. Furthermore, the enlargement of early waveforms at the occipital cortex is consistent with the idea that selection occurs early in the processing stream and that incoming sensory signals may be enhanced early as a result of attention. But, as we will see, some attentional processing may also occur later.

Interestingly, just as cross-modal links confer benefits in behavioral tasks, interactions between different primary sensory areas revealed by ERP studies have similar effects. For example, paying attention to either an auditory or a tactile stimulus appearing on one side of space resulted in enhanced ERPs within the first 200 milliseconds at electrodes in primary visual areas. Thus, attending to one side of space in the tactile *or* the auditory domain automatically resulted in enhanced attention to visual information on that side. This result indicates that when a salient event takes place at a given location in one modality, spatial attention is directed to that location in other modalities as well (Eimer & Driver, 2001; Eimer et al., 2002). This seems to be a very efficient way of wiring up the attentional system, and the result at a crowded party is that when you spot your friend visually, you are better able to hear and localize the sound of your own name.

3.2. Functional Neuroimaging and Transcranial Magnetic Stimulation

Data from ERP studies have been very useful in demonstrating that some attentional modulation occurs during the first phases of cortical processing, and in showing similarities across endogenous and exogenous cueing and across visual, auditory, and tactile domains. These data have been able to do this because of the temporal precision of the ERP technique, allowing measurement of changes in brain waves over time even down to a matter of milliseconds. But ERP methods are not that good at indicating exactly which region of the brain is responsible for generating the brain waves. Because the electrodes are placed on the head and the potentials are recorded at the surface of the scalp, we can never be perfectly sure of the location of the brain region producing the potentials. Functional neuroimaging serves as a good complementary approach: its temporal precision is not as good as that achieved by ERP methods, but its spatial precision is far better. The two main functional neuroimaging methods, positron emission tomography (PET) and fMRI, measure blood flow or metabolism in very particular regions of the brain. Their use in attention studies can demonstrate the regional consequences of attending to a stimulus. (For a fuller discussion of methodological issues, see Chapter 1, Section 4.)

In one of the early PET studies, participants were required to shift attention between locations—in this case, boxes aligned horizontally across the visual field—and to press a button when a target appeared in one or the other of the boxes (Corbetta, et al., 1993). The experimenters found that the superior (i.e., upper) parietal lobe in the right hemisphere was consistently activated during attentional shifts compared to periods when fixation was maintained in the middle of the screen (Corbetta et al., 1993; Vandenberghe et al., 2001). The involvement of the superior parietal lobe was also evident in another visual search study, especially when the target contained a conjunction of features. Although other regions of the brain, including the basal ganglia, thalamus, insular cortex, frontal cortex, and the anterior cingulate, also show enhanced activation during attentional switching in visual search tasks, the parietal lobe seems to play the primary role.

Another experimental design used PET to monitor brain activation as participants inspected an image for a change in either its color, motion, or shape (Corbetta et al., 1990). In addition to the parietal cortex, brain areas associated with processing motion were activated when participants attended to motion. Similarly, when participants attended to color, brain areas associated with color were activated. The importance of this correspondence is the demonstration that while the parietal cortex plays an important role, it is intimately linked with other brain areas that reflect the attentional modulation of the relevant features of a display.

The studies of attention have become increasingly sophisticated, and there have been attempts to differentiate the neural processes associated with different forms of attention. For example, Corbetta and Shulman (2002) showed that different neural systems are used when attention is directed to a location before the appearance of a stimulus (an endogenous condition) and when an unexpected salient stimulus appears (an exogenous condition) and redirects a participant's attention. They found that in cases of endogenous attentional orientation, a network of frontal and dorsal parietal areas (including the intraparietal sulcus, IPS; superior parietal lobule, SPL; and the frontal eye fields, FEF) was involved. Searching for your friend in a crowded room involved voluntary reorientation of attention that would activate this frontoparietal network of brain areas. On the other hand, effects involving exogenous mechanisms of attention, such as those arising from the harsh sound of shattering glass, were found to activate a more ventral system that included the temporal parietal junction (TPJ) and ventral frontal cortex. The authors hypothesized that this latter system is involved in the detection of unexpected salient or novel information. The two systems were described as functionally independent but interactive (Figure 3–14). Information from

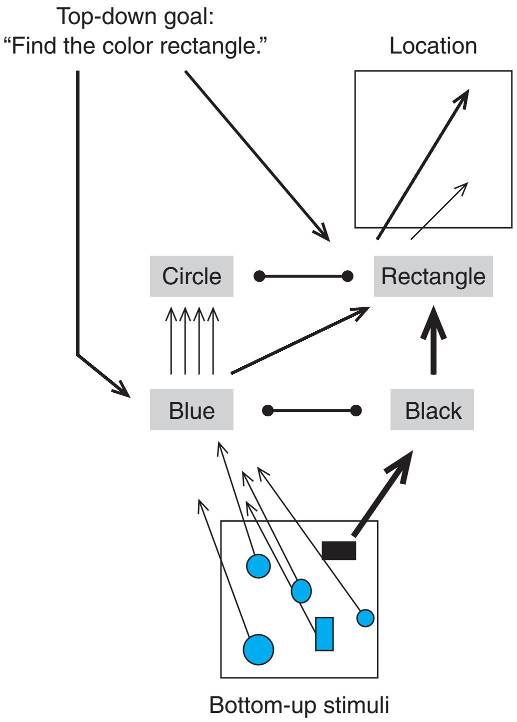


FIGURE 3–14 The interaction of attentional systems

Where is the [blue] rectangle? Finding it requires the interaction of top-down (endogenous) and bottom-up (exogenous) attentional systems. Arrows indicate activation; lines with dots on their ends indicate inhibition.

the ventral system may interrupt processing in the voluntary frontoparietal system and reorient attention toward the bottom-up, salient stimulus (as you were drawn to the sound of the crashing glass). Conversely, information about the importance of the stimulus from the voluntary system may modulate the sensitivity of the ventral system, providing a top-down way for your intentions or goals to influence how strongly your attentional system will be distracted by exogenous information.

The notion that attention operates over time as well as over space—for which there is behavioral evidence from studies on attentional blink and repetition blindness—is supported by evidence from neuroimaging. You've been looking for your friend for some time without success. Have you missed her? Has she left? Some people are beginning to leave the party. As more and more people start looking for their coats, you become ever more vigilant in looking for a patch of green. Such expectations for the appearance of a stimulus in time have also been reflected in activation of the frontoparietal brain areas, suggesting that the voluntary attentional neural network operates temporally as well as spatially (Coull et al., 2000; Wojciulik & Kanwisher, 1999). It is probably not surprising that after damage to this region, patients not only are impaired in attending to information on the side of space opposite the lesion (as in hemispatial neglect), they are also impaired in attending to information that is presented in a rapid temporal stream.

Evidence from a completely different technique, transcranial magnetic imaging (TMS), has also confirmed the critical role of parietal cortex in attention. This technique, which was also discussed in Chapter 1, allows a magnetic field to be passed through the neurons in a particular region of the brain, putting those neurons into a state in which they are inhibited from being activated by incoming stimuli. This technique, which is used with nonneurological participants, in effect induces a temporary "lesion" in healthy brains lasting a few seconds or minutes (with no demonstrable consequences after the study is completed). TMS studies with normal participants can therefore be thought of as analogous to studies of patients with damage to specific regions of the brain (such as patients with hemispatial neglect). TMS applied to the right parietal cortex of normal participants led them to require more time for conjunctive search but not for searches for simple features (Ashbridge et al., 1997). (Recall from earlier in the chapter that conjunctive search, but not simple-feature search, requires attention.) Interestingly, this increased time to search after TMS can be eliminated by training participants on the conjunctive search task (Walsh et al., 1998), perhaps thus making it more automatic and less demanding of attention. Another TMS study, which disrupted the IPS in one or the other brain hemisphere, resulted in impaired detection of stimuli on the side of space opposite the site of TMS; but this was true only when two stimuli were presented, one on the left and one on the right (Hilgetag et al., 2001). Together these studies appear to support the idea that the superior and posterior parietal lobes are involved in attentional shifts, and that damage to these areas results in a bias for attention to be directed to the side of space that is registered by the "intact" hemisphere, as is true in hemispatial neglect.

The ERP studies provide much support for early selection and enhancement of visual cortex processing during periods of attention; the PET and fMRI studies indicate that many other areas of cortex are activated as well in attention. These include parietal cortex and frontal cortex; imaging studies have also shown activation in occipital cortex, confirming the ERP studies (Gandhi, et al., 1999). The most important message from these studies is that attention does not simply engage one area of the brain. Rather, attention is implemented in a wide, distributed circuit to which different brain regions contribute. "Attention" involves selection that can occur early, late, or at both times, and be driven by our will or by the strength of stimuli in the environment and on the basis of space or objects or time. And so the most fruitful way of thinking about attention is as a dynamic system that flexibly allows for selection in many different ways. These studies of neural systems in the human brain complement the behavioral studies described in previous sections.

Comprehension Check: ✓

1. Describe two different methods that have been used to study the brain basis of attention.
2. Do the ERP studies support an early or a late view of attentional selection? In principle, could ERP results provide support for both views? If so, how? If not, why not?

4. COMPETITION: A SINGLE EXPLANATORY FRAMEWORK FOR ATTENTION?

Conceptualizations of selective attention have undergone many transformations over the history of attention studies. Early theories drew analogies between attention and a filtering mechanism or bottleneck that operated in accordance with a set of early perceptual or later semantic criteria. Later theories recast attention as the selective distribution of a limited amount of cognitive resources. Now attention was seen not as a discrete gateway or bottleneck but as a modulatory influence that could increase or decrease the efficiency with which demanding processing is performed. Considered in this way, attention is a far more flexible mechanism—capable of facilitating or inhibiting processing of the input—than a simple spotlight or filter. Data from many differently designed studies have suggested ways in which attention may be implemented in the brain; many different brain areas, from the posterior occipital lobe, to the anterior frontal lobe, appear to be involved in attentional processing.

Is there a general theory of attention that embraces the findings from neural studies and observed behavior? The answer is yes. This is the theory of *biased competition,* or *integrated competition,* developed by Desimone and Duncan (1995) and Duncan and colleagues (1997).

From the perspective of this theory, attention is seen as a form of competition among different inputs that can take place between different representations at all stages of processing. In a simple competition model, the input receiving the greatest proportion of resources (say, because it possesses salient bottom-up attributes) would be the one that is most completely analyzed. A very strong bottom-up signal (such as a shattering glass at a party) would be rapidly and efficiently processed, even above the ambient noise of the party. Competition (and selection, which is the result of the competition) between noise-of-breaking-glass and general-noise-of-party would occur in auditory cortex. The same sort of competition would occur with inputs in other sensory modalities.

The competition between inputs can be biased by the influence of other cognitive systems. Focusing on visual processing, Desimone and Duncan (1995) argue that attention is "an emergent property of many neural mechanisms working to resolve competition for visual processing and control of behavior" (p. 194). Instead of characterizing attention as a spotlight that highlights particular regions of space for processing, or as a bottleneck or filter, these investigators characterize attention as an integral part of the perceptual or cognitive process itself. Competition occurs because it is impossible to process everything at once; attention acts as a bias that helps resolve competition between inputs. So, for example, if the input contains a gray circle, competition would occur between different color representations (or the neurons that constitute the representations), and gray would win; the neurons of that representation would fire, and gray would be considered the winning input. The source of the bias can come either from the features in the external stimulus (exogenous) or from the relevance of a stimulus to one's goals of the moment (endogenous).

The competition that takes place between possible inputs occurs in multiple different brain regions. For example, competition in earlier areas of the visual system will tend to be influenced by exogenous factors such as color or motion. This competition will, in turn, affect more anterior brain regions to which these lower-level areas send information for further processing. There, however, endogenous factors, such as relevance or goals, will tend to bias the competition in regions of the brain involved in formulating plans for how to pursue specific goals. This later competition can also send information back to the regions of lower-level processing and modulate the influence of exogenous factors at that level. The theory holds that many different brain regions are involved in such competition, and because they are connected, the competition is integrated across them. The ultimate winner of the competition—the item that is ultimately attended—is determined by a consensus among all the different regions working together. Given this perspective, findings that so many different brain areas contribute to attentional selection are not surprising.

One of the original motivations for the idea of competition came from a single-cell recording study in which monkeys were trained to perform a visual search task (Moran & Desimone, 1985). The important result was that when two targets were in the same receptive field, they competed for the cell's responses. However, when one of the objects was the target and the other was a distractor, the neuron responded primarily to the target stimulus and blocked processing of the distractor stimulus. If you imagine these competitive processes occurring all along the processing stream, one way to understand attention is to think of it as a gating mechanism that biases processing according to a combination of external salience and internal goals. The outcome of the competition is a winner that is selected for further, preferential processing.

Several ERP and fMRI studies have shown how the ERP waveform or the activation of various brain regions is enhanced or magnified under conditions of competition. These increases occur when participants perform more difficult discriminations (Lavie, 1995), when distractors compete with targets, and when the demands of a task are increased. When task demands are increased as in dual-task conditions, less activation is observed in areas associated with a simultaneous secondary task, reflecting the decrease in processing of nonselected information.

Competitive effects are also seen when stimuli appear simultaneously rather than sequentially, probably reflecting mutual suppression by simultaneously competing stimuli. When four visual stimuli were present, there was less activation in visual cortical area V4 than when only one was present (Kastner et al., 1998). However, when participants were asked to attend to just one of the four stimuli that were presented simultaneously, activation went up to a level similar to that observed with a single stimulus presented alone (see accompanying, *A Closer Look* box). In the framework of integration competition theory, attending to a single object effectively reduced the amount of competition from other stimuli and biased processing toward that stimulus.

Other demonstrations of attentional modulation, obtained with both fMRI and ERP techniques, have been found in V1 and other early visual cortical areas (Brefczynski & DeYoe, 1999; Gandhi et al., 1999; Luck & Hillyard, 2000; Noesselt et al., 2002; Somers et al., 1999). some researchers have even found very early attentional effects in the lateral geniculate nucleus of the thalamus, a key informational relay station between the retina and visual cortex in the back of the brain (O'Connor et al., 2002). These results indicate that, as the theory holds, information can be sent back to areas that accomplish earlier processing and thereby bias subsequent processing. Moreover, there is some suggestion that feedback connections to early visual cortex are involved in determining our conscious experience of visual information (Pascual-Leone & Walsh, 2001).

Competition is also evident when a participant is required to divide attention across two perceptual features. Compared with a simple condition in which no stimulus appears on the screen and the participant stares at a fixation point, attending to the color or the shape of a stimulus that appears on the screen leads to greater activation of many visual areas. Moreover, if the participant has to switch between attending to the color and attending to the shape, additional areas of the brain become engaged and—no surprise—regions of parietal cortex are also active (Le et al., 1998; Liu et al., 2003).

Many of the findings that demonstrate failures of selection in space or in time can be explained by the idea of competition between stimuli. For example, in covert attentional cueing (see Figure 3–7) the invalid trials can be thought of as cases in which there is competition between the location indicated by the invalid cue and the location where the target appears; in the valid condition, the location

A CLOSER LOOK

Competition and Selection

We examine here an investigation that explored mechanisms by which competition by stimuli might be enhanced or decreased. The work, by Sabine Kastner, Peter De Weerd, Robert Desimone, and Leslie Ungerleider, was reported in 1998 in "Mechanisms of Directed Attention in the Human Extrastriate Cortex as Revealed by Functional MRI," *Science, 282,* 108–111.

Introduction

The investigators used fMRI to test ideas about competition and attentional selection. The idea being explored is that the visual system is limited in its capacity to process multiple stimuli at any given time. The hypothesis: in order for objects to be selected, competition must occur between available objects to yield a "winner." The suppression of the eventual winner by the "losing" stimuli would give rise to a reduction in the signal measured by fMRI. In addition, the investigators argue that this suppression can be overcome even when multiple objects are present: if attention is directed to one of the objects specifically, the response would be enhanced and lead to stronger fMRI signals.

Method

Eight participants viewed images appearing on a screen while lying in the fMRI scanner. Two conditions were used in the first experiment. In the *sequential condition,* four complex images were shown in random locations on the screen but only one object was ever present at a single time. In contrast, in the *simultaneous condition,* the same four complex images were present, but in this condition all four were present at the same time. Because competition (and suppression) can take place when all four are simultaneously present, the expectation was that the fMRI signal from visual cortex in the simultaneous condition would be lower than the sum of the four fMRI signals in the sequential condition. The second experiment was the same as the first but now, in some sets of trials, participants were instructed to attend to a particular location in which one stimulus was presented and to count the occurrences of a particular target stimulus at that location.

Results

Markedly weaker signals were observed in many visual areas of the brain in the simultaneous than in the sequential condition, supporting the idea that the stimuli presented together competed with one another and, in so doing, led to the suppression of some of them. When stimuli are presented one at a time, each can activate the brain fully and so the sequential condition leads to stronger fMRI signals. It is interesting and important that in the second experiment, when the participants attended to the location of one of the stimuli, there was an increase in the strength of the fMRI signal, and this signal was even larger than in the sequential condition in some visual areas of the brain.

Discussion

The hypothesis was that competition among multiple stimuli would lead to suppression and that this would be reflected by a reduction in the fMRI signal. This was indeed the case, and this finding provides support for the idea that attention is a dynamic process in which stimuli compete for selection. Competition, then, may be the means by which unwanted stimuli are filtered out (they are suppressed and have little or no activation). The second experiment shows that one can enhance some stimuli by attending to them even if there are many stimuli present; this finding shows that along with suppression, enhancement occurs when one selects a subset of items for further processing.

of the valid cue and the location of the target are one and the same, and hence cooperation rather than competition prevails. Also, the effects of divided attention can be interpreted as the result of competition between different inputs or different tasks, as opposed to the noncompetitive case, in which the focus is exclusively on a single input or a single task. The improvement, in the form of automaticity, that comes with greater practice with dual tasks may be thought of as a reduction in the competition between the two tasks. Moreover, the performance of patients with hemispatial neglect can also be understood within this framework of competition. If the damage to the right side of the brain allows the intact hemisphere to produce a bias away from the left side and toward the right side, that bias increases the competitive strength of right-sided stimuli and reduces that of left-sided stimuli.

Failures of selection in time lend themselves to a similar explanation. The failure to report T2 in the attentional blink task (see Figure 3–4) might arise from competition between T1 and T2. Reporting T2 when it is not preceded by T1 is not problematic—there is no competition. However, the requirement to report T2 when it is preceded by T1 and is very similar to T1 in appearance (say, the letters are "A" and "H") establishes a highly competitive environment and reduces the chances of detection of T2. Competition may also explain the failures of selection in time observed with patients with hemispatial neglect. When visual stimuli are presented on both sides, reporting of the stimulus on the neglected side improves depending on the timing of presentation of the two stimuli and their grouping. One might think of these two factors, time and grouping, as biases that can influence the outcome of the competition between stimuli on the right and on the left.

It seems, then, that almost all the behavioral experiments we have discussed can be interpreted in terms of a competition between "stronger" and "weaker" stimuli, with strength defined by a combination of bottom-up and top-down influences. Although not all the details of biased competition have been worked out yet, this framework allows us to explain a wide range of findings; its direction is promising. The advantage of this theory is that it underscores the idea that attention is a bias in processing, and that processing occurs through cooperative and competitive interactions among brain areas. Because the different brain areas are connected, they will all contribute to the selection of the target. By combining the behavioral results with the inferred involvement of particular brain areas, it is possible to begin to develop an understanding of how attentional effects are manifested in the neural system and how those changes affect cognition and behavior.

Comprehension Check: ✓

1. How might an attentional blink result from competition between two stimuli?
2. Give an example of how a salient bottom-up signal might compete with and win over another stimulus.

Revisit and Reflect

1. *What is attention, and how does it operate during cognition?*

Attention is the process whereby we can select from among the many competing stimuli present in our environment, facilitating processing of some while inhibiting processing of others. This selection can be driven endogenously by our goals (e.g., to find a particular friend, to follow an instruction, to use an arrow to direct attention), or exogenously by a salient or novel stimulus that captures attention away from the task at hand (e.g., bright light, loud noise). Because there is too much information at any given moment for us to cope with, attention is the mechanism by which the most important information is selected for further processing. The type of information that we miss and the conditions under which we miss it are, therefore, the flip side of the cognitive processes involved in attentional selection. Being unaware of the posters on the wall at a party is a failure of selection that is a property of selectively searching for features of a friend. Although we are capable of processing only a limited quantity of information in both space and time, fortunately selection does not occur randomly. Both our goals and the salience of information around us determine where and to what we attend. This balance between endogenous and exogenous factors not only allows us to accomplish our goals effectively, such as finding an individual in a crowd, but also to be sensitive to important external information, such as a fire alarm or crashing glass.

Think Critically

- Describe the differences between endogenous and exogenous attentional processing in space and in time.
- What would it be like if you were equally aware of all the visual and auditory details of your environment at once? Would this be an advantage or a disadvantage?
- Does studying in a noisy environment such as a coffee shop help you focus or does it distract you? Do the level of noise and the difficulty of the subject matter or its type (verbal, pictorial) affect the suitability of a study location? How?
- How do cross-modal processes (e.g., visual-to-auditory) facilitate attentional selection of goal-relevant information such as looking for a friend in a crowd? How can cross-modal processes hinder it?

2. *What information-processing models have been devised to understand attention?*

Different models of attention have each been successful in capturing particular aspects of attentional processing. The debate over whether attention operates at an early or late stage highlighted two aspects of attention. First, attention can have an effect on the very earliest levels of perceptual processing by reducing the amount of information that enters into our cognitive system. Second, some unattended information reaches very late stages of processing, which shows that not all unattended information is entirely filtered out. Information that is contextually consistent with our goals or likely to be of extreme importance, such as our name, penetrates the attentional filter. The spotlight metaphor for attention reflected the reality that space is a powerful coordinate system for our perceptual systems, and that attention operates on these sensory systems directly. For example, turning toward the sound of crashing glass at a party might result in the incidental selection of other things in that spatial location such as a piece of furniture, which you would have otherwise failed to notice. Later theories, such as feature integration theory and guided search, proposed more complex models of attention that involved early preattentive and later attentive stages of processing. These theories provided a mechanism for how attention integrates information. As theories change over time, they build on ideas from previous theories and increase in the detail of explanation. In this way, our understanding of attention builds over time.

Think Critically

- According to research findings, would it be more effective to search for your friend at a crowded party along a particular dimension (e.g., color of her dress, her height) or in terms of a combination of dimensions? Which and why?
- From knowledge of the different theories of attention, what advice would you give to advertising agencies for creating advertisements that are likely to be noticed and read? What advice would you give to Web masters who want to control the distractability of Web advertisements on their pages?
- In what way is the spotlight an appropriate metaphor for attention, and in what way is it not?
- According to feature integration theory, what is the difference between preattentive and attentive processing?

3. *How have new techniques for studying the brain enhanced our understanding of attention?*

Together, ERP, TMS, PET, and fMRI studies have corroborated and extended information-processing concepts of attention. They have shown that attention modulates processing in early sensory areas such as the primary visual cortex but that the attentional signal may be generated from processing in the parietal and frontal lobes. The parietal and frontal areas associated with attention are separated into two neural systems that are interconnected. The more dorsal system is involved in endogenous attention and tightly connected with the motor systems that produce eye and other body movements. This system underlies the voluntary selection of relevant information and the transformation of it into discrete actions, such as moving one's eyes toward a person in green. The more ventral system is sensitive to the appearance of new exogenous stimuli, such as the sound of breaking glass, and this system can modulate, and be modulated by signals from the dorsal system. The results suggest that the attentional system in the brain involves highly interconnected areas that interact to produce effective selection of relevant information.

Think Critically

- Damage to which areas of the brain would impair the endogenous and exogenous attentional systems, respectively?
- What deficits in searching for a friend in a crowded room would you expect to occur if you had brain damage in each of the two attentional systems?
- What properties of the neural systems involved in attention have TMS, ERP, PET, and fMRI helped us to understand?
- What areas of the brain have been found to be involved in attentional processing, and how are they involved?

4. *Attention, according to one contemporary theory, is a competition among different sources of information, all vying for further processing. Can such a theory explain both the behavioral and brain perspectives on attention?*

   The competition framework characterizes attention as a signal that biases processing toward the most relevant or salient features, which are then processed further. Attention as a biasing signal acts within as well as between perceptual and cognitive systems. The outcome of the bias present in one phase of processing is passed on to other phases and acts as a bias there. The effects of competition are dynamic, just as experiments have shown attentional effects to be. According to the competition framework, the reason why it is so difficult to find your friend in a crowded room is because there are too many competing objects that are either too similar, such as other people, or too salient, such as crashing glass and loud voices. Your friend, as an object, does not immediately win the competition for processing. If the party were less crowded or noisy, it would be easier for properties of your friend to be selected and other properties to be inhibited. This example also points out the continuous nature of competition: rather than selection's being binary, biased competition suggests that the selection process is continuous and graded.

Think Critically

- How does biased competition differ from the other theories of attention?
- How can the idea that the information we are aware of is essentially the "winner" of competing information be used to inform laws about using cell phones while driving?
- How can we use the principle of biased processing to produce more effective road signs?