

---

## Selective Attention

We are almost always subjected to a barrage of different sources of sensory information at any instant. You can verify this by pausing and trying to survey the various sensations impinging upon you at this very moment. You will probably find some auditory stimuli (an airconditioner?) and kinesthetic stimuli (the back of the chair?), as well as many visual inputs originating in this text, but also more peripheral visual stimuli. Probably you would agree that you were not aware of more than a few of these stimuli at the moment you began trying to enumerate them. As described in chapter 1, a great deal of dispute surrounds what happens to stimuli that are rejected. The process of selecting from among the many potentially available stimuli is the clearest manifestation of selective attention.

Two questions about selective attention have been the subject of most of the research reviewed here. The first concerns what factors make selective attention more or less easy and efficient. The second is, what are the consequences of selective attention for processing both attended and rejected stimuli? This question leads directly to some of the key issues over which early and late selection theories disagree. The first question, regarding ease and efficiency, is also relevant to that dispute, but less directly so.

Probably the best-known real life example of selective attention is listening to a single voice in a room full of people talking at the same time. For over a decade, most contemporary studies of auditory selective attention examined a laboratory analogue of this everyday phenomenon, beginning with the work of Cherry (1953). Examples of visual selective attention also abound, but the phenomenon of visual selectivity is more

**elusive.** The reason it is elusive is because the casual observer tends to think of visual selective attention as consisting of nothing more than fixations of the eye, and assumes that changes in attention are equivalent to eye movements. As James (1890, p. 437), Helmholtz (1924, p. 455), and presumably many others before them noted, this identification is not valid: we can choose which visual stimulus to attend to without moving our eyes. Consider the case of reading, activity that depends on taking in information from one or at most a few words at a time. For purely sensory reasons, words become less legible as they move toward the periphery, but nonetheless one can easily choose to read the word above or below the fixated word. One can even move one's eyes to scan one line of text while reading the line immediately above or below this line, although doing so feels unnatural and effortful. It is hardly surprising that the distinction between visual selection and eye movements is so easily overlooked, however, because there is rarely any reason to try to fixate one object and attend to another. The relation of eye movements and selective attention has further intricacies that are discussed toward the end of this chapter. The point of this example is to demonstrate that even in the most routine of our activities, visual selection is not exclusively achieved by movements of the eye, but depends on internal selective mechanisms as well.

This chapter describes many laboratory phenomena that illustrate the voluntary control of visual selection without eye movements. For example, an observer may be shown a display of colored letters (too brief to allow an eye movement) and asked to report the identity of the red letter. In other cases, an observer may view a display of several words and try to name which one an arrow points toward. In each case, a well-defined selection criterion determines which stimuli are to be selected and which are to be rejected. The criterion of selection is a simple physical attribute. Logically speaking, whether a stimulus fits this criterion can be determined without identifying it (e.g., as a letter, word, or picture). These tasks satisfy the defining attributes of filtering tasks, as described in chapter 1.

The goal of this chapter is to explore some of the most important empirical results from studies of filtering tasks, and the implications of these results for the mechanisms of selective attention. Limiting the

discussion to filtering tasks does not by any means imply that selective attention mechanisms are involved only when a task explicitly requires filtering. In fact, the next two chapters (on divided attention and set) describe some quite different situations in which selective attention mechanisms may also be involved. For example, in chapter 4 we consider what happens when a single visual stimulus is briefly presented for report, and the observer may be cued in advance where the stimulus will be. Logically, this task too might involve some of the same processes as selection from a crowded display. However, it is useful to begin with filtering tasks, in which the requirement for selection is obvious and inescapable.

### Auditory Selective Attention

#### Cherry's Studies: Reports about the Unattended

The researches of E. Colin Cherry (1953) provided a powerful impetus for contemporary work on selective attention, and they continue to be widely cited. Cherry performed ingenious experiments in which two auditory messages were presented simultaneously, one to each ear (*dichotic* presentation). Dichotic presentation generally gives a listener the experience of two streams of sound, each localized roughly at the ear of input. Although it produces strong localization, it is an unnatural stimulus, because localization is normally achieved on the basis of multiple cues, of which intensity differences are only one (Scharf and Houtsma, 1986).

Cherry instructed his subjects to shadow one of the messages (i.e., repeat it back without delay). He observed several things. First, subjects found it fairly easy to carry out the task, which Cherry thought remarkable in itself (although he acknowledged this was presumably well known to people administering hearing tests). His second observation was that when subjects were asked, after they had finished shadowing, to describe the contents of the rejected message, they could say almost nothing about it, except that sounds had been present. When the message on the rejected ear started out in English and then switched to German, subjects rarely noticed this. The same was true when English speech was played backward on the rejected channel. On the other hand, when the speaker

switched gender (and pitch), or when speech was replaced with a 400-Hz tone, listeners almost always noticed and remarked on it.

Cherry concluded that "certain statistical properties [of the rejected message are] identified, but that detailed aspects, such as the language, individual words, or semantic content are unnoticed" (1953, p. 978). In a well-known follow-up to his work, Neville Moray (1959) examined this striking lack of memory for the rejected message in more detail. His subjects shadowed the message presented to one ear while a message recorded in the same voice was played to the other ear. In some cases, the same word list was repeated thirty-five times in the rejected message. A recognition test on the words from the list disclosed that subjects were no more likely to recognize these words as having been presented than they were to recognize never-presented words. It is not surprising that words on the attended channel were recognized at a reasonably high level.

These results are not peculiar to processing of linguistic input. Diana Deutsch (1986, pp. 32–20) reported informally obtaining the same kinds of effects with melodies played on the piano. Listeners sang along with the melody presented to one ear. When questioned about the melody presented to the other ear, they were able to say very little about it. As Deutsch pointed out, the phenomenon is even more remarkable with musical stimuli than with speech, since music (but not speech) often involves several perceptual streams proceeding in parallel. It is somewhat hard to see how a person who never processed more than one melodic line at a time could have a very rich appreciation of, say, the music of Bach. This reinforces a point made on purely logical grounds in chapter 1: the need to distinguish between the processing people *can* accomplish when they try to divide their attention, on the one hand, and the processing that is accomplished when they try to focus on a particular stimulus, on the other.

### Factors Affecting the Difficulty of Selection

In Cherry's original studies, the ear of arrival (and presumably the subjective localization that resulted) provided the basis for segregating the two messages. He also noted (perhaps no surprise) that the task was much harder when an irrelevant message spoken in the same voice was

played in the same ear as the relevant message. This effect is not restricted to irrelevant spoken messages: Egan, Carterette, and Thwing (1954) observed a similar advantage when they played noise to the same or different ear to which speech was played.

As noted earlier, presenting messages to different ears causes them to be subjectively localized in or near the ears themselves, as a result of the unnaturally extreme difference in the intensity of the signals at the two ears. It turns out that even moderate differences in interaural intensity produce enough localization to improve selective shadowing a great deal compared with zero differences (Treisman, 1964a). Intensity is not the only cue that the auditory system uses to localize sound. Another one is the difference in the time at which a sound arrives at the two ears (and closely related to that, the phase of the signals at the two ears). A sound to the extreme right or left of a listener will arrive at the nearer ear up to about 0.7 msec earlier than it arrives at the farther ear. Despite the very brief intervals involved, the auditory system makes effective use of this cue, especially at low frequencies. A variety of psychophysical studies, principally using measures of detection thresholds, shed light on the relationship to selective attention. They found that people are better able to hear a pure tone superimposed on noise when the interaural timing localizes the tone and the mask at different points in space. This is known as the masking level difference (MLD), and it has been studied extensively since it was first discovered by Langmuir and colleagues (see Scharf and Buus, 1986, for a succinct review).

Suppose a tone is presented to one ear (monaurally) against a background of noise presented to both (binaurally). When the noise at the two ears is correlated and in phase, the threshold for detection is typically about 9 db lower than it is when the noise at the two ears is uncorrelated. Subjectively, the correlated in-phase presentation makes it sound as though a single noise source is located in the center of the head, producing subjective distance from the source of the tone to be detected. Interaural time differences seem to help in selectivity among speech signals as well. Broadbent (1954) delayed one of two messages briefly, and listeners found this helpful in focusing on one of the messages.

The early observation that people can selectively shadow the message played to the left or right ear led to the general (mis-)impression that ears

constitute natural or basic channels that can always be processed separately without difficulty. As noted above, dichotic presentations are quite **unnatural**, and hearing typically relies heavily on elaborate decoding of **the pattern of input to the two ears to construct a representation of sound sources in the three-dimensional world** (see Warren, 1982, for an overview). For this reason, it would be surprising indeed if the two ears constituted fundamental input channels, as these early studies sometimes are taken to imply.

Indeed, later studies provided a more realistic picture. Consider, for example, the findings of Treisman and Riley (1969), who had their subjects shadow the message in one ear while a different message in a **different voice** was played in the other ear. Unlike in the earlier studies, the authors used digital signal processing to synchronize the individual pairs of words on the two channels in a precise way, shrinking or stretching each word to have a duration of 250 msec. Under these conditions, shadowing one message was very difficult, and many intrusions from the rejected channel were noted. The investigators suggested that subjects may find it difficult to determine which ear a sound is coming to unless some of it (perhaps especially the onset) coincides with silence on the other ear. This suggestion is consistent with work in auditory psychophysics that suggests that auditory localization relies heavily on the initial onsets of acoustic stimuli (Haftner and Buell, 1985; Zurek, 1980).

In summary, filtering tasks show excellent selectivity by spatial location at least under conditions in which localization cues are adequate. This, however, should not lead one to think of the two ears as channels per se. Rather, it appears that various cues leading to the perception of sounds having different locations in the world provide the basis for successful filtering; dichotic presentations may or may not provide such conditions, depending on the temporal relation of the two inputs.

People can also filter effectively on the basis of frequency differences between attended and rejected messages. Treisman (1964b) had listeners monitor one of two messages played binaurally (i.e., each of two messages played at equal intensity in both ears). Under these conditions, she found that differences in voice pitch improved selectivity considerably. A woman's voice was shadowed with 74% accuracy when the irrelevant

message was spoken in a man's voice (reading the same type of material), but only 31% accuracy when the irrelevant message was spoken by the same woman. With both messages coming into both ears, such differences may partly reflect frequency-selective masking at peripheral levels in the auditory system, and when the subjects are shadowing there is the complication that they hear their own voice as well. Underwood and Moray (1971) circumvented these problems by having listeners monitor one of two messages to detect digit targets. With both dichotic and monaural presentations, they detected more targets when one voice was male and the other female, compared with when both voices were recorded by speakers of the same gender. Male and female voices overlap in their frequency content and differ in other ways as well. It would be interesting to analyze the critical factors that enhance selection in this situation. It could be studied using synthetic speech, but this does not appear to have been done as yet.

The fact that differences of pitch and location facilitate selective attention was one of the main pieces of evidence cited by Broadbent in support of his filter theory. According to his interpretation, these physical attributes are the only attributes of messages that are extracted, regardless of whether or not the message is attended to. In his view, this preliminary analysis is used to define channels for the purpose of subsequent selection. It is no coincidence that subjects also tend to notice changes occurring in these attributes in rejected messages, as Cherry found also.

The ease of auditory selection also depends on the number of channels on which information is presented. Treisman (1964a) used three different channels to present different messages recorded by the same speaker. Messages on the left and right channels were played to the left and right ears, respectively. A third (middle) channel was created by playing messages at equal volume to the two ears. As noted above, this results in sound that is subjectively localized midway between the ears. When subjects shadowed the message played to the right ear, shadowing was scarcely affected by playing a single additional message on either the middle or left channel, compared with no rejected message at all, confirming Cherry's observations. However, when two irrelevant messages were played, one to the left channel and one to the middle channel, shadowing of the right channel was substantially impaired. This could

potentially have happened either because there were two different messages to be rejected, or because two different spatial channels contained information that had to be rejected. Treisman concluded it was the two channels rather than the two messages that created the problem, because when the irrelevant messages were both played on the middle channel (or on the left channel), the problem was largely eliminated.

Does having to reject two distinct channels impair selection when the channels are defined by frequency instead of location? To answer this question, Treisman placed a male and a female voice on the central channel while subjects shadowed the right channel. Performance was impaired here also, arguing that it is rejecting any two separate channels that poses special difficulty. The problem in monitoring one of three channels did not occur when the irrelevant channels contained speech sounds ("bet bet bet" or "a a a"), rather than speech, and Treisman therefore concluded that the problem could not be attributed to peripheral auditory masking (Scharf and Buus, 1986). The difficulty remained when the messages consisted of passages spoken in Czech. The reason for this difference was not clear. Treisman's observations are intriguing, but they have not apparently been followed up.

What conclusions can be drawn from these studies of selective listening? First, differences in subjective sound location between attended and rejected stimuli clearly facilitate selective attention to a spoken message. Merely having messages presented to different ears does not make it trivial to select a message, however; the messages must be at least partly nonoverlapping in time. Differences in pitch also facilitate selectivity. Finally, the difficulty of auditory selection seems to be increased when the number of rejected channels grows beyond one.

#### **What Is Noticed in the Rejected Message?**

As described, Cherry was struck by how little subjects spontaneously reported about the rejected channel. Broadbent's early selection account drew support from the fact that what subjects tend to notice pertains to just the attributes that serve as effective cues for selection, such as pitch and location. However, under certain conditions subjects sometimes *do* spontaneously notice something of the linguistic content of the unattended channel.

One example was noted by Cherry in his initial report (1953). He tried playing the same message to the two ears, separated in time by a variable delay. He started with a comparatively long interval (he did not make it clear whether the attended or rejected message was leading or lagging) and gradually reduced it. Subjects noticed the repetition when the lag got down to the range between 2 and 6 seconds. Treisman (1964c) examined spontaneous detection of repetition more carefully. Her subjects either shadowed the message in the right ear or simply listened to it, knowing they would have to describe it later. In either case, they were led to believe that listening to the sounds played in the left ear would cause trouble and should be avoided. The task was performed over and over with 40-second passages, and the lag between the repeated messages was shortened until the subjects spontaneously remarked on the repetition. Once this happened, the experiment ended.

When both channels contained the same sentences, all subjects noticed the repetition of content and whether the passages were spoken in the same voice or a different one. However, the lag had to be much shorter for the repetition to be noticed when the message in the rejected channel led. Here, repetition was noticed at about 1½ seconds, on average, compared with about 4½ seconds when the message in the attended channel led. The results were similar when the messages consisted of speech played backward, showing that understanding the words and sentences was not critical for detecting repetition.

How does spontaneous noticing of repetition bear on the extent to which the to-be-ignored material is processed? Broadbent (1958) assumed the existence of a sensory memory that holds onto 1 or 2 seconds' worth of rejected material, now commonly referred to as echoic memory. Of course, he supposed that this memory included physical, but not semantic, descriptions of the sounds. To account for repetition detection, one must suppose that when new sensory information arrives, it is matched against this memory. The existence of some matching processes can be argued for on other grounds, since listeners detect periodicity in repeated segments of white noise up to about 1 or 2 seconds in duration (Guttman and Julesz, 1963; Warren and Bashford, 1981). In this case, listeners hear the periodicity as a whooshing or "motorboating," depending on the frequency; obviously, this sort of implicit periodicity detection

might not be related to the conscious detection of repetition studied by Cherry and Treisman. In any case, since people can detect repetition even when the rejected message lags, one would have to suppose that sensory memory for the attended message is compared with incoming material on the rejected channel, as well. Since the matching is acoustic rather than lexical or semantic, this does not conflict with an account like Broadbent's; it should also be kept in mind that noticing a repetition in a long segment might reflect merely intermittent or low-fidelity matching, and does not imply continuous high-fidelity matching.

What poses more of a problem for a theory like Broadbent's, however, is Treisman's observation that repetitions of the same sentences are detected even when they are spoken in different voices. Of course, detecting repetition of the same words by a different speaker might not necessarily involve semantic processing of words, but it surely requires more than a direct comparison of elementary acoustic features. Treisman also tried presenting a messages in English to one ear and the French translation to the other ear to listeners fluent in both languages. Relatively few of them detected the "repetition," so perhaps the degree of semantic processing implied by the different-voice same-message results may not be that great.

All of these results can be reconciled with Broadbent's model simply by supposing that subjects occasionally relax their selectivity and sample the rejected channel. When the selected message leads, any word from the rejected channel that was checked would be one that occurred just a little earlier on the selected channel. It is reasonable to suppose that these selected words would be in a postsensory short-term storage, so it is hardly surprising that the repetition would be readily detected in this situation, even when the lag is longer than the life of auditory sensory memory. Things are a little more complicated when the rejected message leads. Any word sampled on the rejected message would not yet have been spoken on the attended channel, so to explain repetition detection one must suppose that the word that was sampled would itself have to be held onto until its match appeared on the attended channel. It is not hard to see why under these conditions repetition detection would be less effective, with the result that repetition could be detected only at relatively shorter lags. However intriguing, detection of speech

repetition does not rule out even such extreme perceptual selectivity as that suggested by Broadbent's account.

The results can also be accounted for in terms of Treisman's attenuation theory (see chapter 1). She supposed that when the attended message led, processing these words primed the detectors for them, and therefore when the same words appeared on the rejected channel, they were recognized and, at least briefly, *selected* themselves. Her account of the rejected-leads condition was less explicit, but it seemed to rely on sensory memory in the manner suggested in the previous paragraph. There is no doubt that the result could be reconciled with late selection theories as well.

Findings regarding detection of repeated speech are therefore not critical for choosing among very different accounts of the fate of unattended stimuli. However, the detection phenomena are intrinsically interesting and deserve further study for this reason alone.

In a well-known study, Moray (1959) verified what many have noticed outside the laboratory: a listener tends to notice his or her own name spontaneously when it happens to appear in a rejected message. As Loftus (1974) pointed out, many textbooks have implied that people usually or even always notice their own names. In fact, the results were less dramatic: when the rejected message contained the word "You may stop now," subjects noticed and reported this only 6% of the time. On the other hand, the message "John Smith [the name of the subject], you may stop now" was noticed and acted upon 33% of the time. This went up to 80% when subjects were clued in advance to be on the lookout for new instructions (it is not clear whether or not they were led to expect new instructions to appear in the rejected channel). Of interest, Oswald, Taylor, and M. Treisman (1960) observed that sleeping subjects are often awakened by their own name.

The fact that people detect their own name was cited as support by late selection theorists such as Deutsch and Deutsch (1963). Treisman (1960) accounted for it by supposing that detectors for certain stimuli were perpetually primed, with the result that detection could occur even on the basis of attenuated stimulus information. Since the effect occurs only intermittently, it can be reconciled with a model such as Broadbent's simply by supposing that filtering lapses from time to time. This

possibility illustrates a general difficulty in interpreting selective listening results: effects of rejected stimuli may be due to processing that happens only occasionally, and even voluntarily, when subjects become curious about what they are missing on the other channel.

#### Subtler Measures of the Fate of Rejected Sounds

The research described showed that listeners consciously trying to attend to a message have little or no enduring memory for the words or the meaning conveyed on a rejected channel, except for their own name or sounds identical to those on the attended channel. These are fascinating observations, but not theoretically decisive with respect to the fate of rejected stimuli. Subsequently, many researchers claimed to find stronger evidence against selective perception. In these studies, the contents of the unattended channel were neither spontaneously reported nor revealed by memory probes. Instead, more indirect and subtle sorts of evidence proved that their contents had been analyzed to the point at which their meaning was registered. Most of these studies were reported after Deutsch and Deutsch proposed their late selection theory in 1963, and the results were often taken to provide strong support for that theory.

If an individual can say nothing about the contents of a rejected channel, how could one infer from the person's behavior that the person had, at some level, analyzed the meaning of the stimuli? Experimenters have shown a high degree of ingenuity in finding ways to demonstrate this. Treisman (1960) created dichotic messages consisting of two passages, one in each ear. The subject was instructed to shadow one channel and ignore the other. At an arbitrary point the two passages switched, with the message formerly on the left channel continuing in the right channel, and vice versa. When the two passages consisted of ordinary prose recorded in the same voice, about 30% of the time the subjects' shadowing switched channels along with the message.<sup>1</sup> When this occurred, subjects generally shadowed only a few words from the "wrong" channel, and promptly jumped back to the channel they were supposed to be shadowing. One intriguing aspect of the results is that when this happened, the subjects rarely noticed that anything funny had happened. Figure 2.1 shows a typical record, with the words the subject shadowed in capital letters. A less well-known but intriguing observation was that

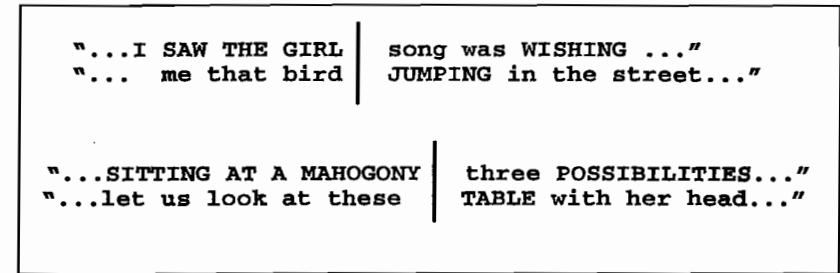


Figure 2.1

Examples of the intrusions observed by Treisman (1960). In each selection, the capitalized words are the ones spoken by the subject, whose instructions were to shadow only one line (the one written on top). Reprinted with permission.

when one prose passage was recorded in a man's voice and the other in a woman's, the subjects often noticed the switch but rarely followed it.

Treisman accounted for the channel-jumping phenomenon within the framework of her filter-attenuation theory. She proposed that extraction of the meaning of the *attended* message resulted in a lowering of the threshold for recognizing words most likely to follow. She was less clear about why subjects actually switch over to shadowing them, and why they switch back a short while later. Note, however, that at the point of a switch, the material on the *attended* channel does not make sense. It may be, therefore, that subjects sample the other channel even if they would not otherwise have done so. Thus, one cannot necessarily infer from these results that subjects *routinely* analyze the rejected message to the point of determining its meaning.

Other auditory studies employed even more subtle measures. One example is the autonomic response to words that were previously paired with shock. Corteen and Wood (1972) performed the first such experiment, which involved a conditioning phase and a testing phase. In the conditioning phase, subjects listened to a series of words through headphones. The words were drawn from a list of three different city names, and each was repeatedly paired with a mild but unpleasant electric shock. In the testing phase, the subjects shadowed a prose passage presented through the right ear while trying to ignore the contents of the left ear. Shock-conditioned city names (as well as new city names and unrelated

nouns) were included in the unattended left-ear message. Conditioned galvanic skin responses (GSR) were detected for both shock-associated city names 38% of the time and for the other city names 23%; comparable GSR responses occurred only for about 10% of unrelated nouns. Shadowing errors did not increase when critical words occurred on the rejected channel, and on questioning, the subjects denied having noticed any shock-associated words in that channel. Corteen and Wood concluded that the contents of the unattended channel were processed to the point of semantic analysis.

The result of this test is widely cited in introductory textbooks as evidence that the meaning of unattended messages is fully analyzed. However, subsequent research weakened the case quite a bit. To start with, the observations sometimes proved difficult to replicate (Wardlaw and Kroll, 1976). Others raised questions about how the results should be interpreted. For example, von Wright, Anderson, and Stenman (1975) measured GSR responses to rejected words. They also compared the magnitude of GSR response to the critical words (conditioned words and associates of them) in the monitored and unmonitored channels. The mean GSR response to a critical word in the monitored channel was more than twice the response to such words in the unmonitored channel. These results involve averages over trials, and cannot tell us whether GSR responses to stimuli in the rejected channel occurred less frequently than responses to attended items, or whether they occurred equally often but were smaller in magnitude. What the results do show, however, is that the Corteen and Wood effect does not prove full processing.

Dawson and Schell (1982) analyzed the Corteen and Wood effect with unusual thoroughness. They replicated the original findings quite satisfactorily, and also employed additional measures to eliminate trials on which subjects lapsed in their selective monitoring of the relevant channel. These measures included shadowing errors, subjects' reports of having noticed items on the irrelevant channel, and, in one condition, overt detection responses (subjects were instructed to press a key whenever they heard one of the words).<sup>2</sup>

Dawson and Schell's main result was that when analysis was restricted to trials in which *none* of their measures indicated possible lapses, the GSR responses to the contents of the rejected channel were almost, but

not completely, abolished. The data also showed an interaction with laterality: when the to-be-ignored material was played to the right ear, there were more signs of processing in the GSR than when it was present in the left ear. In light of these findings, it seems plausible that the Corteen and Wood effect simply shows that a subject's rejection of the distractor message occasionally lapses, voluntarily or involuntarily. To put it differently, the results do not force one to conclude that rejected as well as attended messages are identified fully in every trial. A filter that functioned as Broadbent suggested, but that was prone to lapses and switches, could certainly account for the results.

Another finding commonly taken to confirm semantic processing of rejected messages was reported by Lewis (1970). Lewis examined the speed (rather than accuracy) of individual shadowing responses to words in the attended channel as a function of the identity of the word occurring at the same time on the rejected channel. Shadowing response times (RTs) were slowed by about 30 msec when the rejected channel contained a synonym of the to-be-shadowed word; they were speeded up by about the same amount when the rejected channel contained an associate. This result does not seem to be terribly robust. Treisman, Squire, and Green (1974) confirmed the effect, but found that it occurred only early in the list. By the seventh word, no increase whatever could be detected. These authors suggested that lapses in monitoring occurred before subjects had a chance to occupy themselves fully with monitoring the relevant channel (p. 645).

Further evidence against completely unselective processing comes from an interesting experiment performed by W. A. Johnston and Dark (1982). The subjects' primary task was to monitor a dichotically presented word list for any target words (the names of states) and to repeat any words they heard from the target category. In the selective-attention condition, all the state names were in the message played to one ear, and subjects attempted to monitor only this ear. In the divided-attention condition, state names were distributed evenly in both ears, and the subjects attempted to monitor both. Occasionally, the subjects were interrupted with a visually presented probe word (e.g., bark) to which they were supposed to respond with an immediate free association (e.g., dog). Semantic processing of spoken words was revealed in subjects'

word associate choice. Each probe word was homophonous with two entirely different meanings. Subjects' associations were scored according to which of these two meanings they were related to (e.g., dog and birch are examples of associations related to the two different meanings of the probe bark). Sometimes two prime words tending to favor one or other of the two meanings of the probe word (e.g., growl and noise) were presented in either the attended or the unattended ear.

Consider first the selective attention condition. When the two prime words were in the attended channel, subjects' associations followed the primes 69% of the time compared with a baseline of 50%. When the prime words were in the unattended channel, the subjects produced prime-related associations only 52% of the time, which was not significantly different from 50%. The corresponding measure in the divided-attention condition was at an intermediate level (59%). As Johnston and Dark point out, this priming measure is sensitive to processing of nontargets in the auditory monitoring task, but it does not support the idea that nontargets in attended and rejected spatial channels are processed to the same degree (see also Johnston and Wilson, 1980).

A study by MacKay (1973) provides a final illustration of the tendency for breakthrough of the unattended effects in audition to become less convincing as the effects are investigated more carefully. Subjects shadowed ambiguous sentences in one channel, while the contents of the other channel contained potentially disambiguating information. It appeared that the unattended material biased subjects toward the related reading of the sentence they were shadowing. For example, subjects tended to interpret the sentence "They threw stones toward the bank yesterday" as being about a financial institution when the word "money" was in the unattended channel, and as being about riverbanks when the unattended word was "river." Newstead and Dennis (1979), however, found that MacKay's effect occurred only when just a single word was in the unattended channel. When the unattended channel contained a series of words, the effect disappeared. They suggested, quite plausibly, that an unexpected isolated sound in the unattended channel may disrupt the ordinary state of channel selection.

What can be concluded from these studies about the extent of processing of unattended speech? Certainly, the indirect effects of the content

of the rejected channels rule out any model claiming people can *always* and *completely* exclude speech sounds from anything more than simple physical processing. The meaning of rejected stimuli can affect listeners' behavior in quite diverse ways. What is lacking, however, is persuasive evidence that rejected message is analyzed to the *same* degree as attended message. The results are consistent with an alternative interpretation according to which semantic analysis occurs when selection lapses and/or is restricted to recognition of stimuli that are particularly salient (or primed, in Treisman's terminology). The unmistakable tendency has been for more careful studies to find less rather than more evidence for semantic analysis of unattended stimuli.

We now turn to visual selective attention, asking the corresponding questions about ease of selection and extent of processing of unattended stimuli.

### Visual Selective Attention

#### Ease of Selection

What makes it easier or more difficult for a person to attend selectively to a particular visual stimulus and exclude others? As noted, in daily life we are typically aware of spatial selection only when we make an overt movement of our eyes. Reading, for example, obviously requires deliberately taking in information from different positions on the page at different times, and eye movements are an obvious manifestations of this selection process. They do not, however, *constitute* selection.

The most straightforward evidence that people can select visual stimuli without eye movement is obtained with visual filtering tasks. The instruction "tell me the name of the single red letter in the display of ten letters" would be an example. Here color (red) is the selection criterion and letter identity is the reported attribute. How can we compare the effectiveness of selection by different criteria? One way would be to determine how well an observer can perform visual filtering tasks as a function of which attribute serves as selection criterion, where the criterion is varied between blocks of trials while the displays and the reported attribute remain constant. (This sort of comparison was roughly accomplished by some of the early selective shadowing studies described in the preceding section.)

No one seems to have made explicit comparisons of this sort using either RT or accuracy as the dependent measure. Many studies examined RTs for selective report, but with spatial location serving as the selection criterion throughout (e.g., Eriksen and Hoffman, 1972a, 1972b). Other studies used accuracy measures to compare different selection criteria, but the goal was to examine visual sensory (iconic) memory, rather than the efficiency of selection per se. For that reason, cues were presented at variable intervals after the offset of the display (e.g., Averbach and Coriell, 1961; von Wright, 1970). For example, when color<sup>3</sup> was the selection attribute and letter identity was the reported attribute, the subject might see a brief display, knowing that either red or green items would have to be reported, but not which. Shortly after the display offset, a tone would sound, and its pitch would indicate whether red or green items would be reported. Accuracy was reasonably high so long as the cue was not delayed more than a few hundred milliseconds after display offset, reflecting sensory persistence from the display (usually termed iconic memory; see chapter 3). Selection from iconic memory probably involves the same process as selecting from a display that is still physically present (evidence for this will emerge throughout succeeding chapters). Assuming this is the case, iconic memory studies provide at least some comparative information about the effectiveness of different selection criteria.

The first comparison was carried out by Sperling (1960) in his classic partial-report studies. His subjects were shown a brief display of characters. After a variable interval, a tone indicated which subset of a display of letters and digits they should report. The selection criterion was either spatial location or category (letter vs digit). Accuracy was much better for location cueing than for category cueing. Von Wright (1970) performed a more systematic comparison in which a display of eight characters was presented for 100 msec. A high or low tone sounded as the display disappeared, cueing the subject to report a subset of the display, composed always of four items. Performance was excellent when the selection criterion was location, color, size, or brightness, but much worse when it consisted of an attribute that might plausibly depend on identification (e.g., orientation, letter vs digit, vowel vs consonant).

These observations were sometimes taken to show that information in iconic memory is precategorical, that is, that objects were not identified.

This interpretation of the results would be consistent with early selection but not late selection theories. As several authors pointed out (e.g., Allport, 1989; Duncan, 1981), however, the results do not require this interpretation. Even if the items were all identified, identity information might not be a good selection cue. Consider a filing system in an office, for example. One can easily retrieve file folders based on whatever is written on each file's tab, but this obviously does not mean that no other information is contained in the file itself. Indeed, later studies found that selection by category (letter vs digit) can yield a partial-report superiority effect (Merikle, 1980). Based on findings of this kind, some authors suggested that visual persistence *must* include categorical information; this inference is also dubious.

What, then, can be concluded about the question we started with: the relative ease of selection using different selection criteria? Selection by simple, physical attributes such as location, color, size, and brightness is clearly effective in filtering tasks involving brief visual displays. This closely mirrors findings from auditory selection discussed earlier in the chapter. Although this has been taken to argue that unattended stimuli are not processed to any deeper level (Broadbent, 1958; Kahneman, 1973; Kahneman and Treisman, 1984), the argument is weak. Attended and unattended stimuli could even be processed to exactly the same extent, with only precategorical attributes facilitating selection.

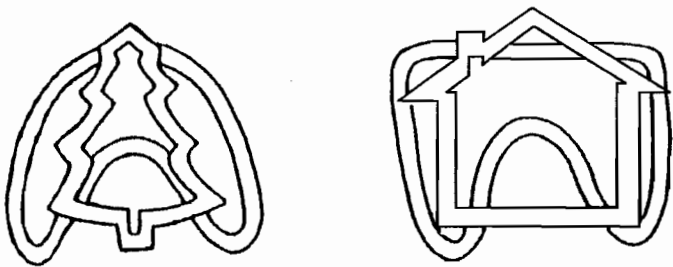
### Fate of Rejected Stimuli

**Memory and Spontaneous Noticing** Moray's classic observations with word lists repeated over and over in the rejected ear in a shadowing task showed that people typically have little enduring memory for the contents of a rejected auditory channel. The same is true of rejected visual stimuli. Neisser (1976) had subjects read text in which alternating lines consisted of different prose passages, each written in a different color. Subjects had no difficulty reading only the red lines, and showed little memory of the unattended message.

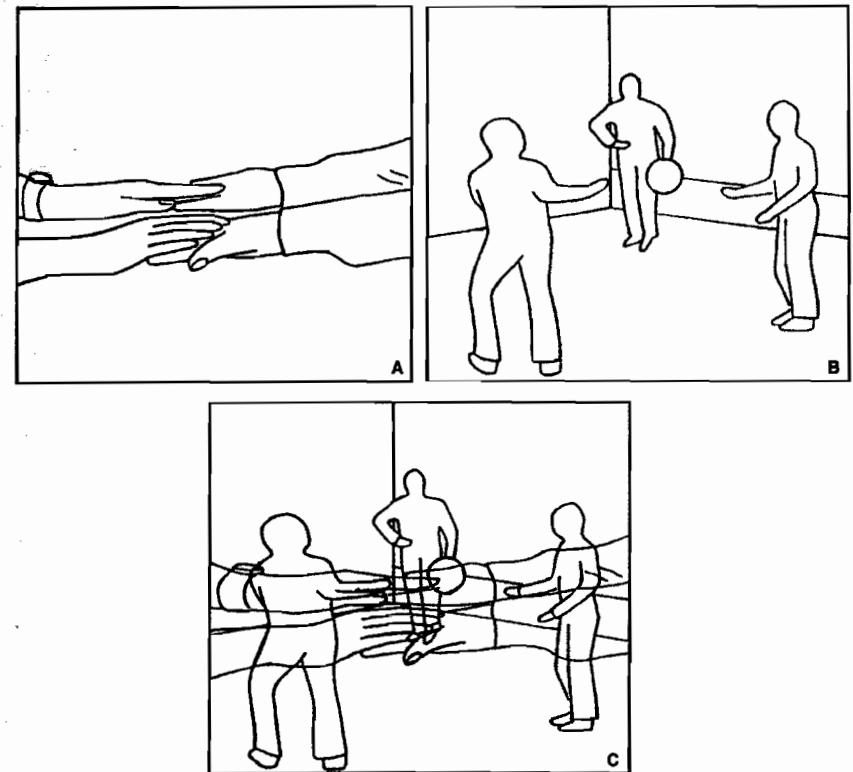
There is little doubt, however, that a reader generally fixates directly on each attended line, giving the excluded lines a double disadvantage: greater eccentricity (hence poorer acuity) as well as attentional exclusion. Some studies used spatially overlapping forms to minimize this problem.

Rock and Guttman (1981) showed subjects series of pairs of overlapping forms (figure 2.2). Each pair was viewed for 1 second, and subjects reported their aesthetic judgment about the red figure. After viewing ten pairs, a surprise recognition memory test was given. Subjects were able to recognize the attended items almost as well as when no rejected item was present. The rejected items, however, were recognized at a level little better than chance. This was true even when the ignored line figure was the outline of a familiar object (e.g., house), or when subjects were asked about the contents of a display just 1 second after it had disappeared.

Following up on observations of Kolers (1972), Neisser and Becklen (1975) presented video images from two cathode-ray tube (CRT) displays to observers in such a way that the images appeared in the same part of their visual fields and at the same distance. This was achieved by placing a half-silvered mirror between the observer and one of the CRTs; the mirror reflected the image from the other CRT at equal contrast. The effect is described as similar to what one sees looking out a window from a well-lit room at dusk: the scene out the window and the reflection from the room are superimposed. Neisser and Becklen showed subjects video images of two types of events, one consisting of human hands interacting in a game, the other showing complete figures of people tossing around a ball (figure 2.3). Subjects were given the task of detecting particular events in the videos, such as a ball being thrown. They were able to monitor whichever video image they chose with only minimal interference from the other. However, they were extremely poor at monitoring



**Figure 2.2**  
Examples of the overlapping forms used by Rock and Guttman (1981). Within each pair, one form was colored red. Reprinted with permission.



**Figure 2.3**  
Outline tracings of typical video images of (a) a hand game alone, (b) a ballgame alone, and (c) hand game and ballgame superimposed. Reprinted from Neisser and Becklen (1975, figure 1) with permission.

both scenes. The results mirrored Cherry's findings with dichotic messages very closely.

Other findings mirrored Cherry's results as well. Just as Cherry tried changing the sounds played in the excluded ear, Neisser and Becklen placed various peculiar events in the videos (e.g., disappearance of the ball, substitutions of players). Only rarely did the observers who were monitoring the other video channel notice and report any of these events.

**Indirect Measures of Identification** We saw earlier that investigators have used a variety of indirect measures to assess the extent to which

rejected auditory stimuli are perceptually analyzed, such as switching ears in shadowing and conditioned GSR changes. Even more diverse measures were employed to provide indirect evidence of processing of unattended visual stimuli. Most studies involved filtering tasks in which location was the selection criterion. Subjects made a deliberate response to a target stimulus at a fixed position. Excluded stimuli were presented at the same time or nearly so, usually in adjacent positions, but sometimes overlapping, as in Rock and Guttman's study. The critical evidence for processing of rejected stimuli consists of some sort of change in the speed or accuracy of responses to the target as a function of the relationship between the rejected stimulus and the target.

The best-known demonstration of this type is the *Stroop effect* and the large family of (apparently) related effects (Stroop, 1935; MacLeod, 1991). In the classic Stroop effect, the subject reads aloud the color of the ink in which a word is printed. When the word spells out the name of an incompatible color (e.g., GREEN printed in red ink), one is substantially slower to respond "red" than in a neutral condition (e.g., CHAIR printed in red ink). The difficulty caused by incompatibility is so large that it is quite noticeable when one reads through a long list of incompatible color/word stimuli. With long lists, overt errors as well as hesitations tend to occur. Most recent Stroop studies, however, have used individual computer-controlled trials, with a mixture of matching and mismatching color/word stimuli in different trials. Under these conditions, subjects can usually maintain high accuracy levels, and the effect of mismatches shows up as a slowing of response times on incompatible trials. A wide variety of effects at least somewhat analogous to the Stroop effect have been reported in the literature over the years. For example, if people report the number of characters on the screen, they are slower when these characters are numerals incompatible with the correct response, such as a display of four 3s (Flowers, Warner, and Polansky, 1979). Another Stroop-like task calls for the subject to say aloud the direction in which an arrow points; a left-pointing arrow is responded to more slowly when the word RIGHT is embedded in it, rather than the word LEFT (Shor, 1971).

The classic Stroop effect produces a substantial slowing on incompatible trials compared with neutral trials, often well over 100 msec difference. A speed-up in compatible trials is sometimes observed, although it

is usually more modest. The slowing of RTs due to color/word mismatch can also be induced by a color word that is separated from the target color patch, although as Kahneman and Henik (1981) noted, the Stroop effect is larger in the standard case in which the word and the color belong to the same object. Gatti and Egeth (1978) presented a color word at eccentricities of 1 to 5 degrees from a central patch; the subjects named the color of the patch. Increases in spatial separation reduced the size of the Stroop effect, but even at 5 degrees, considerable interference remained (35 msec).

The classic Stroop effect certainly demonstrates that one cannot simply and completely turn off one's word-recognition machinery. Gatti and Egeth's results show that one cannot even completely prevent a word in a position known to be irrelevant from being read. As will emerge below, more ambitious inferences drawn from the effect are more problematic.

Analogous to the Stroop effect, but less colorful, are the so-called *flanker effects* observed in choice reaction-time tasks. Eriksen and Hoffman (1973) had subjects make a speeded response to a central letter, pressing the left key for an A or a U and the right key for an H or an M, for example. Next to this letter were other letters that the subject attempted to ignore. When these flankers were associated with the opposite response from the correct one on that trial (e.g., an H surrounding a target A), RTs to the target letter (here, A) were slowed, compared with the case when the distractors were associated with the correct response. Flanker effects are generally smaller than Stroop effects, which is not surprising given the effect of spatial separation noted by Gatti and Egeth. The size of the effect is reduced when the flankers are presented more than 1 degree from the central items, but this may be due to reduced acuity rather than complete selectivity at the greater distances (Egeth, 1977). Closely analogous flanker effects have also been observed in tasks involving words. Shaffer and LaBerge (1979) presented two identical words above and below a central target word. The subject's response depended on the semantic category of the central word. When the flanking words were drawn from a different category than the target, RTs were slowed by about 40 msec.

Another indirect measure of perceptual processing which has been explored in many studies is *semantic priming*. Word recognition occurs more quickly when someone has just read a semantically associated

word. Priming is commonly observed by presenting a *prime* word to which no response is made, followed by a target letter string to which subjects make a speeded *lexical decision*, indicating whether the letter string is a word or a nonword. The response to a target word is faster when the prime is semantically related to it (e.g., doctor-NURSE) rather than unrelated (e.g., chair-NURSE).

As with flanker and Stroop effects, semantic priming has sometimes been found when the prime is presented at the same time as a target word to which the subject responds. In one study of this kind, Dallas and Merikle (1976) had subjects name the word in a precued position, ignoring another word above or below it. Subjects were faster by about 20 msec when the two words were related rather than unrelated. It has been suggested that this sort of priming reflects automatic (unselective) processing of the prime word. If the primes are helpful to performance, however, one cannot assume that subjects really have an incentive to ignore them. In that case, such observations cannot test the idea of unselective processing as this concept was defined above. Other observations sometimes taken to support automatic perceptual analysis involve priming effects that occur when the target is not related to the prime in most trials. Since these primes convey no information about targets, some maintain, any processing of them must be obligatory rather than intentional or strategic (Neely, 1977; Posner and Snyder, 1975b; Tweedy and Lapinski, 1981). This argument is not compelling, either. In many of these priming experiments the prime is the only word present on the screen in a position that will be relevant to the subject's task; this would hardly seem to be an optimal condition for exclusion. In addition, the effects observed from primes that do not convey information is often restricted to beneficial effects on performance in related trials, with no decrement in unrelated trials. This means that observers have no obvious incentive to try to ignore them. These various priming results provide little or no basis for concluding that word identification is unselective. Stronger evidence would have to come from the Stroop and flanker effects when the task explicitly demands filtering.

**Indirect Measures: A Closer Look** The diverse effects described in the preceding section seem to suggest that when people try to ignore familiar

visual stimuli, these stimuli are nonetheless sometimes processed to the point of recognition. Such effects are often cited as evidence for unselective perception. That is, they occur because all stimuli undergo full processing to the point of recognition regardless of the observer's desires and regardless of what other stimuli are being identified at the same time.

As seen above, when indirect effects of rejected auditory stimuli were closely scrutinized, the evidence they provided for wholly unselective processing became less convincing. The same is true for effects of rejected visual stimuli. Consider first the obvious possibility that all such effects might simply reflect occasional lapses in selection, rather than completely unselective processing occurring whenever extraneous stimuli are presented. This possibility was raised in connection with several auditory findings, and it can be raised here as well. How can one be sure that flanker and Stroop effects, for example, reflect analysis of rejected stimuli that occurs in every trial or even in most trials? Oddly, this question has only rarely been tackled despite the large literature that deals with processing of rejected stimuli.

Eriksen, Eriksen, and Hoffman (1986) presented some relevant data. Their subjects performed a Sternberg memory-scanning task, deciding whether a centrally presented probe letter belonged to a memory set that was presented earlier. This probe item was flanked by other letters that sometimes did, and sometimes did not, belong to the memory set. Substantial flanker effects were observed (the exact nature of the effects is not presently relevant). Eriksen et al examined not just the mean RTs, but also the distribution of individual RTs in the compatible and incompatible conditions. Suppose that incompatible flankers produced a constant slowing of, say, 30 msec in every trial in that condition. In that case, one would expect that the cumulative distribution function for incompatible trials would be just like the distribution for compatible trials, but shifted to the right by 30 msec. By contrast, if the effect stemmed from, say, a 300-msec slowing that occurred only in one-tenth of the trials, the effect on mean RTs would be the same, but one would expect a different pattern in the RT distributions. The incompatible condition would include more very slow responses, and these would replace trials over most of the range of the compatible trials' distribution. Thus, the compatible and incompatible trials would differ little among

the fastest RTs, and would diverge progressively more at the slower RTs. The data of Eriksen, Eriksen, and Hoffman appeared to fit the first prediction, not the second, suggesting that the incompatible flankers produce modest, reliable slowing, rather than large, occasional slowing. Nevertheless, further analyses are in order,<sup>4</sup> and the generality of the conclusion has to be verified. As the literature stands, many discussions of processing of rejected stimuli simply assume that effects of unattended stimuli reflect events occurring in every trial rather than just occasionally, but this has not been adequately demonstrated.

A second fact that weakens the interpretation of the indirect effects is that rejected stimuli almost invariably come from the same set as the relevant stimuli or at least from a set related to those stimuli. Therefore, it is quite reasonable to suspect that the subject may be primed for these stimuli, as suggested by Treisman. According to Treisman's proposal, unprimed stimuli would not produce the same effects, and thus the indirect effects simply do not reflect true stimulus identification. Some evidence on this point was reported by J. O. Miller (1987). His subjects made a speeded classification judgment to a central character that was flanked by other characters. Rather than employing flankers that belonged to the relevant set, as in the flanker experiments discussed earlier, Miller used letters that were not assigned to any response in the instructions given to the subjects. However, certain flanker letters were more often presented when one particular response was appropriate rather than another. This correlation meant that the response could be reasonably well, but not perfectly, predicted by knowing the identity of the flankers. Miller found that the speed of responses to the target was affected by the flankers, with faster responses occurring when the response was more likely given the flankers, and slower responses occurring with the less likely flanker-response combinations. Remarkably, subjects appeared to have little awareness of the correlation, although such assessments of awareness are notoriously treacherous (see, e.g., Brewer, 1974).

Miller stated that his results demonstrated semantic processing of unattended stimuli that were not "primed by task relevance" (p. 419). This conclusion may be questioned, however. The very fact that the flankers are correlated with the response means that subjects have no

obvious incentive to ignore them. After all, they do convey information about what the correct response would be, albeit not as much as the central target. Thus, it might well be adaptive for subjects to base their response partly on whatever information they might have about the identity of the flankers, as well as information about the target itself, especially if the former happens to become available sooner than the latter in a given trial. This criticism was raised earlier regarding the idea that primes that convey no information about targets could only have effects because they are automatically processed. The general problem is this: to the degree one finds effects of processing distractor items that are mostly beneficial to performance, one cannot assume subjects are trying to exclude them. Of course, effects of correlated stimuli may have interesting implications for skill learning and other psychological issues.

Other evidence relevant to the selectivity of perception comes from several recent studies. Broadbent and Gathercole (1990) investigated the word-flanker effect first documented by Shaffer and LaBerge (1979). They successfully replicated the finding that nearby words produced flanker effects. However, when entirely novel words were used on each trial, the effects were eliminated. Lambert, Beard, and Thompson (1988) reached similar conclusions, although they were not able to eliminate the effect entirely when the two words were side by side. Broadbent and Gathercole suggested that once subjects had experienced a particular word, they may be able to identify it based on only partial cues, rather than full-blown word recognition. In line with this hypothesis (related to Treisman's idea of priming), it is interesting that the Stroop effect can be produced when only a few letters of an irrelevant color word are presented, such as OR or XXANGE<sup>5</sup> (Singer, Lappin, and Moore, 1975). These studies certainly weaken the argument for unselective processing based on indirect effects of rejected visual stimuli.

Further evidence along the same lines comes from experiments showing that these effects can be modulated either by adding additional stimuli or by changing task demands. Suppose the indirect measures of processing discussed here reflect wholly unselective and capacity-unlimited analysis undergone by *all* stimuli, limited only by acuity and other nonattentional factors. In that case, their magnitude should not be affected

when the perceptual demands involved in the subject's voluntary task are altered; nor should they be changed when the number of unattended items present in the display is changed.

A number of careful studies carried out in the 1980s tested these predictions, and the results paint a fairly consistent picture. Kahneman and Chajczyk (1983) had subjects name the color of a patch of ink in the center of the screen. Above or below this patch were other stimuli. When a single incompatible color name was presented (e.g., RED above a green patch), the expected Stroop slowing was observed. In another condition, two words were presented, one above and one below the patch. Consider the case where one was the incompatible color name and the other was a completely irrelevant word (MOST). According to the unselective identification account, one would expect the Stroop effect to be unchanged, since the color name is identified regardless of the presence of the irrelevant word. However, the addition of the irrelevant word reduced the color/word incompatibility effect from 72 to 36 msec. In short, merely adding to the number of *unattended* channels reduced the impact of the contents of one of those channels.

Investigations of the letter-flanker effect yielded similar findings. Yantis and Johnston (1990) had their subjects decide whether a particular target letter (which varied from trial to trial) was present in a circular display of eight letters. The subjects only had to check one location, indicated by a spatial precue, because the target was always present in that location (if it was anywhere). Thus, each display contained seven rejected letters. Was each of these seven letters identified despite being rejected, as flankers apparently were in Eriksen and Hoffman's task? If so, when one of these seven letters was a copy of the target, one might expect to find a speed-up of the positive responses. A series of experiments with excellent statistical power showed, however, that redundant targets in noncued positions either had no effect or almost no effect (< 10 msec). There was no hint that the identity of distractors beyond those that immediately flanked the target made any difference whatever, despite the fact that with the circular arrays, acuity was constant for the different positions in the display. When positions were not cued, the redundant targets produced the expected speed-up. These results again show that the evidence for supposedly unselective processing can be made to virtually

disappear when the number of rejected stimuli is increased and conditions for selection are optimized.

It should not be assumed, however, that the flanker effect is so fragile that it will disappear with any minor variation in the task. J. O. Miller (1991) tried five manipulations that he thought might have a chance of making the flanker effect disappear. He used flankers that did not have abrupt (possibly attention-grabbing) onsets, or he presented sequential displays to let the subject lock onto the center position, or he increased the number of *relevant* stimuli. None of these variables eliminated the effect. However, Miller did not examine the effect of increasing the load of *rejected* stimuli (as did Yantis and Johnston). Furthermore, some of the manipulations he examined did seem to reduce the flanker effect, so it is possible that if all of them had been used together, the effect would indeed have been eliminated.

Further evidence against the unselective processing hypothesis comes from recent investigations of semantic priming. W. A. Johnston and Dark (1985) presented four words for 67, 200, or 500 msec. Subjects attempted to report only the words in two prespecified positions among the four. Recognition memory for the words in the irrelevant locations, assessed at the end of the session, was essentially nil, in line with conclusions described earlier. In certain trials, however, an indirect measure of semantic processing was employed: the sequence of brief word presentations was interrupted for a probe test in which the subject tried to report a test word that slowly emerged from visual noise. When the test words were identical or related to the words that were present in the relevant positions in the report task, they were more easily identified (a form of priming). However, there was no priming for words that were presented in irrelevant locations in the report task (except for some marginally significant priming when the exposure durations were 500 msec—clearly so long that some relaxation of selectivity could well be expected).

Similarly, Hoffman and MacMillan (1985) presented a prime word very briefly, surrounded by several letters, and the subject had to search among these letters for a digit. Immediately thereafter, a target string for lexical decision was presented. The letter-search task virtually abolished the priming effect. Priming occurred, however, when the subject was instructed to ignore the letters and silently read the prime.

Thus, when only one or two letters are present in a display and subjects respond only to the one at a prespecified location, there is little doubt that the nontarget item(s) are processed at least to some degree, probably in most trials. Whether they would still be processed if (a) they were unrelated to the material the subject was attempting to process, and (b) the subject had no incentive to process them, seems to be an open question. However, the hypothesis of completely unselective processing fails when the number of rejected items is increased (in flanker and Stroop tasks), or the extent of concurrent visual processing is increased in studies involving words (in priming situations).

What conclusions can be drawn? There is good reason to doubt the claim that rejected visual stimuli are subject to full analysis. The measures used to support the hypothesis do not behave as one would expect once subjected to closer examination. Logically speaking, one could still contend that rejected stimuli *are* completely analyzed in every trial, but these measures just happen not to reveal it. It is indeed perfectly possible that such analysis occurs, but it is behaviorally silent. Although this view cannot be refuted, it deprives the unselective perception hypothesis of the support that was used previously to argue in its favor. In any case, the attractiveness of this dodge will be reduced further when we turn to the results of divided attention studies in chapter 3.

Do the data help in evaluating Treisman's (1960) suggestion that rejected stimuli are, in some not clearly defined sense, attenuated, and that once this attenuation has occurred, the only stimuli to be recognized are those that are primed in one way or another? Certainly, this hypothesis is broadly consistent with a number of facts reviewed above, especially those from more rigorous studies of recent years. However, the hypothesis cannot be said to have been decisively confirmed, either. The very nature of the indirect measures used in these studies makes the question difficult to answer conclusively. Ideally, one would like to know whether the brain has resolved the identities of stimuli that are unrelated to the task at hand. However, the indirect measures (e.g., Stroop, priming) consist of average differences between responses to related and unrelated stimuli—these provide no evidence about whether different possible unrelated stimuli were distinguished.

**Negative Priming** Dalrymple-Alford and Budayr (1966) made an interesting observation about the color-naming task commonly used to elicit the Stroop effect. When a subject responds to the color of the ink in which a word is printed (say, red), the response is slowed if the word is, for example, GREEN (the basic Stroop effect). Dalrymple-Alford and Budayr found that when, in a given trial, the ink color is the same as the distractor on the previous trial, subjects are slower to respond "green" compared with an appropriate control. This effect—slower responses to a stimulus on trial *n* that was rejected in trial *n*-1—is termed *negative priming*, to distinguish it from the more common facilitatory priming effects.

Since that time, other examples of negative priming have been reported. Tipper and Driver (1988) showed subjects overlapping red and green figures, to which the subjects responded by identifying the red figures and ignoring the green ones. When the rejected (green) distractor in trial *n* was identical to the (red) target in trial *n*-1, responses in trial *n* were slowed. It was harder to select what had previously been rejected. Tipper and Driver also found the same inhibitory effect for responses to items that were semantically related, rather than identical, to the previous distractor, and the inhibition occurred even when a verbal response in trial *n*+1 followed a manual response in trial *n* (Tipper, MacQueen, and Brehaut, 1989).

Various researchers (e.g., Tipper and Driver, 1988; Allport, 1989) suggested the following view of negative priming. In each trial, perception is wholly unselective, with all elements semantically analyzed. The selection process in trial *n* works by inhibiting those activated codes that are associated with the distractor item. This inhibition persists until trial *n*+1, at which time it results in increased RTs when the inhibited code must now control the response. This theory, therefore, represents a novel variant of late selection theory.

No evidence conclusively rejects this account, but there are some reasons to doubt it. First, note that this hypothesis postulates identity-specific suppression of distractors that are excluded because of their position in the display or some other selection criterion such as color. The system could not know in advance which identities would have to be suppressed. To determine that, it is necessary to look up which items

were present in each of the irrelevant positions. It is rather odd to suppose that the identities represented in each of the *irrelevant* locations are first sought out and then suppressed, when it would seem so much simpler to look up the identity of the single *relevant* item and use that for choosing a response. This was put forth in more conventional late selection theories, such as the one proposed by Duncan (1980b, 1981). The oddity of this inhibitory scenario becomes even greater when one considers a situation with one target in a prespecified location surrounded by twenty different distractors. It seems slightly bizarre to suppose that identity codes for all twenty distractors have to be activated, found, and squelched before the appropriate response can be made. If this were true, then the number of distractors should produce a sharp increase in RTs to respond to the target even when the position of the target is known in advance, which does not happen (e.g., Colegate, Hoffman, and Eriksen, 1973).

Of course, such arguments can never be decisive, partly because they rely on an argument from lack of imagination; the mere fact one finds it hard to think why something might happen does not disprove it. Better evidence against the unselective-perception interpretation of negative priming comes from more recent findings. Ruthruff and Miller (1995) observed that when uncertainty about the location of both targets and distractors was eliminated, negative priming disappeared. Furthermore, negative priming is consistent with other scenarios quite apart from the unselective-processing account described above. Whereas negative priming in trial  $n+1$  does indicate that the distractor in trial  $n$  was identified before the response in trial  $n+1$ , it does not indicate that this happened before the response was made in trial  $n$ . Thus, negative priming cannot conclusively say how selection in trial  $n$  did or did not occur. Even a completely successful filtering operation as envisioned by Broadbent would be consistent with the data. Indeed, in his original formulation of filter theory, Broadbent (1958) remarked, "We considered . . . listening . . . as though a message which was rejected at the time of its arrival could never pass through the filter later. . . . This over-simplification was for expository purposes only" (p. 210).

Consider, then, the possibility that the distractor is analyzed some time after selection or even after response in trial  $n$ , perhaps during the

intertrial interval. Why would inhibition be observed later in that case? Since the codes (partially) activated by the distractor proved to be unhelpful in trial  $n$ , they are likely to be unhelpful in the future. Thus, inhibition may result from a learning process that works to improve subsequent performance, perhaps using some sort of error-correction learning procedure (Hinton, 1989). In this account, negative priming would reflect not the process of selection that allowed the subject to perform correctly in trial  $n$ , but rather a tuning process working to improve performance on the present and subsequent trials.

Various findings make this concept plausible. For one thing, Tipper Weaver, Cameron, Brehaut, and Bastedo (1991) found that negative priming was undiminished after a delay of about 7 seconds, which is inconsistent with the possibility that it reflects a highly transient inhibitory state. Second, Tipper and Cranston (1985) observed that negative priming was eliminated when the target item, formerly a distractor in trial  $n$ , was presented without any accompanying distractor in trial  $n+1$ . The same result was reported with the Stroop version of the effect (D. G. Lowe, 1979). This is extremely difficult to reconcile with any account involving an inhibitory *attentional* mechanism. Context specificity may be more consistent with a learning account of the sort sketched above, although it can hardly be said to be predicted by that account.

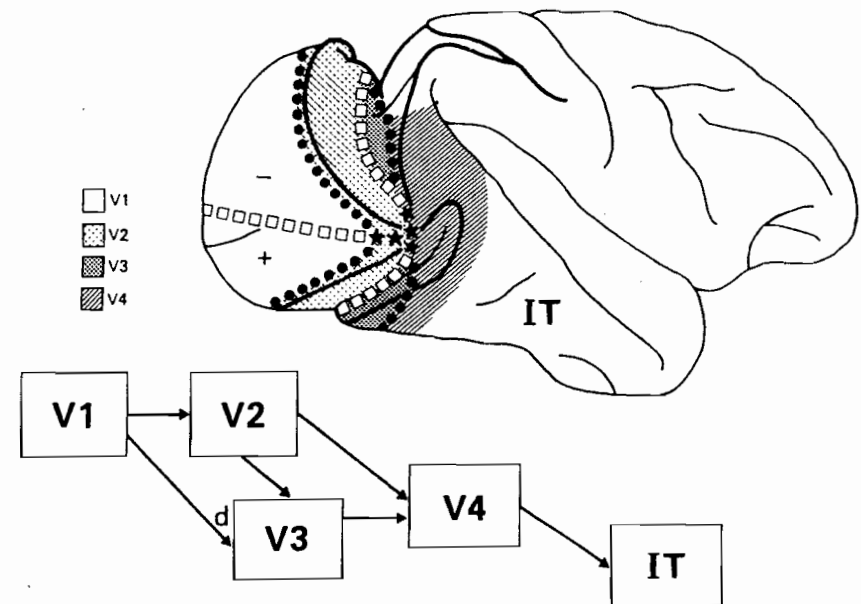
In summary, negative priming effects have been cited as bolstering the hypothesis of unselective perception by providing evidence of processing of rejected stimuli even when the more conventional effects fail to do so (Driver and Tipper, 1989). However, the effects do not indicate *when* the distractors were identified, and it would not be surprising on any account (including Broadbent's original one) if, after selecting a response, distractors were sometimes processed more extensively. The source of the inhibition that shows up in the next trial is still an open question; the suggestions raised regarding learning are merely conjectures. One useful strategy for exploring negative priming further would be to see what happens when many distractors in trial  $n$  are present. If this eliminates negative priming, as Yantis and Johnston found with flanker effects, and as the results of Neumann and Deschepper (1992) with modest-size displays suggest may well occur, negative priming will no longer constitute any real support for unselective perception. If the effect remains, negative priming might challenge some of the conclusions reached above.

**Physiological Measures** Thus far we have considered exclusively behavioral evidence about selective attention. At present, behavioral data provide the most direct evidence regarding attentional selection, but some intriguing physiological observations have emerged in the past several decades. Some of them involve humans and some involve monkeys.

A large fraction of the primate brain contains cells that respond to visual stimuli of some sort or another. In recent years, it has become common to divide the visual pathways in the primate cortex into two broad processing streams: the *dorsal* and *ventral pathways*. The dorsal pathway leads from the primary visual cortex in the occipital lobes into the parietal lobes. The areas along this pathway appear to be specialized for analysis of the location and trajectories of objects (Mishkin, Ungerleider, and Macko, 1983) and perhaps more specifically for the generation of plans actually to interact with these objects (Goodale and Milner, 1992). By contrast, recognition of objects of all sorts (including faces) seems to depend primarily on cortical areas lying along the ventral pathway, extending from the primary visual cortex called through extrastriate visual areas (e.g., V4 in the macaque) to the inferior temporal (IT) cortex. Lesions along this pathway produce impairment in recognition of object identities, colors, and other attributes, and often impair the apprehension of objects' shapes (Farah, 1990). Figure 2.4 shows a simplified outline of the anatomical layout of some of the structures lying along these two pathways.

Recent work has produced more fine-grained and bewilderingly complex maps of these areas (Van Essen, Anderson, and Felleman, 1992). As one moves along the ventral pathway, one encounters neurons with progressively larger receptive fields (the area in the visual field in which stimulation can affect the cell's firing rate). At the same time, the neurons tend to become more specific about the form of stimulation necessary to produce a response. For example, neurons in the primary visual cortex often respond well to oriented bars in narrow regions of the retina (Hubel and Wiesel, 1962), whereas those in IT and surrounding areas seem to respond selectively to complex stimulus configurations such as faces or facial expressions, with little regard to location in the visual field (Gross, 1971).

Moran and Desimone (1985) recorded from single neurons along the ventral pathway in alert macaque monkeys that were trained to perform



**Figure 2.4**

Overview of so-called ventral visual pathway from the striate cortex (V1) into the inferior temporal lobe (IT). This pathway, which appears responsible for object recognition, includes several extrastriate visual areas (V2–V4). Reprinted from Desimone et al. (1985, figure 1), with permission.

tasks requiring spatial selection. The experiments used a *delayed matching-to-sample task*, in which a *sample stimulus* was presented briefly, and then after an interval, either the same or a different *test stimulus* was presented in the same spot. The monkey was rewarded when it responded quickly to just those test stimuli that matched the preceding sample stimuli. In some trials, irrelevant stimuli were presented in different locations together with the test stimuli, and the monkeys' behavioral responses showed they were largely successful in ignoring these irrelevant stimuli, basing their response on the test stimulus alone.

Moran and Desimone asked whether the responses of neurons in different visual areas to a particular stimulus would differ depending on whether or not the monkey was attending to that stimulus. In the primary visual cortex, the responses did not differ. In V4, on the other hand,

responses to irrelevant stimuli were reduced by about two-thirds. Unexpectedly, Moran and Desimone also observed that this attenuation only occurred when an attended stimulus was located within the receptive field of the neuron. That is, a neuron that would ordinarily respond to a particular red bar would continue to respond when that bar was presented, but the monkey was attending to a different bar located *outside* the receptive field. In IT cortex, attenuation was always observed. Here the receptive fields were large enough that they included all stimuli, attended and unattended. Thus, the same generalization may apply to neurons in IT and V4, namely, that the response of a neuron is attenuated only when attention is directed elsewhere within the receptive field of that neuron (see Desimone and Duncan, 1995, for a review).

Corresponding experiments in auditory selection also showed a reduction in neural responses to unattended stimuli seemingly earlier in the processing stream. Benson and Hienz (1978) trained monkeys to respond to either the left ear or the right ear, depending on a visual signal, and ignore stimuli presented in the other ear. The firing rate of cells in the auditory cortex was recorded while tones and bursts of clicks or noise were played to both ears. Approximately two-thirds of the cells responded more strongly to stimuli in the attended ear. However, these attenuation effects appeared to be somewhat weaker than those observed by Moran and Desimone in extrastriate visual areas.

It is tempting to assume that the reduction in firing rate reflects attenuation of processing in something like the sense suggested by Treisman (1960). This interpretation is plausible but may not be right, because the functional significance of individual neurons' firing rate remains unclear. It has been proposed, for example, that the overall firing rate may not be as significant as the pattern of responses across many neurons (e.g., Richmond et al, 1987) or the temporal cross-correlations among individual action potentials (Gray et al, 1992). Reductions in overall firing rate do not necessarily indicate that signals are being blocked from passing beyond an early level. Visual areas in the ventral stream are highly interconnected, with large numbers of fibers running both forward toward the temporal lobes and backward toward the retina. Reduction in firing rate might therefore reflect joint activity in these forward and backward pathways, with higher as well as lower areas responding to

the stimulus. Putting possible complications of this sort aside, the findings of Moran and Desimone and of Benson and Hienz are certainly congenial to the idea of attenuation. In fact, they might even be consistent with the strong conception of filtering envisioned by Broadbent. After all, the monkeys who show partial attenuation of neural responses are not achieving the high degree of behavioral selectivity that is typical in human observers, as shown by their error rates. (This may reflect inherent limits of monkeys compared with humans, or limits on what people can train monkeys to do in the laboratory.)

Another, more plentiful kind of physiological evidence about selective attention comes from studies recording electrical potentials from the scalps of people performing attentional tasks. The continuing electrical potential recorded from the scalp is referred to as the electroencephalogram (EEG). When a stimulus is presented and a person performs some sort of task involving this stimulus, the EEG record reflects both neural events involved in executing the task and those unrelated to the activity. In a raw EEG record, this noise generally overshadows signals from stimulus-related events. Fortunately, the noise can be reduced by taking EEG records from many trials and averaging them together, time-locked to the stimulus. That is, the potentials observed 1 msec after the stimulus in each trial are averaged across trials, then potentials at 2 msec are averaged, and so on. Signals unrelated to the stimulus have no particular temporal relationship to it, so this averaging process makes them cancel each other out, assuming, of course, that they combined linearly with the task-relevant signals in the first place.

This averaging procedure is the most common way of uncovering *event-related potentials* (ERPs). The ERP waveforms generally show several positive and negative voltage peaks after the presentation of a stimulus. These peaks are termed *ERP components*, and are believed to reflect synchronized postsynaptic activity in relatively large populations of neurons. Although there is no guarantee that neural activity critical for any particular information processing operation must be detectable in the ERP, these components nonetheless provide an intriguing source of clues about neural-mental operations.

Many studies have compared ERPs elicited by attended and unattended stimuli. They differ from most of the behavioral experiments described

above in that the attended or unattended stimulus is usually, although not always, presented alone. This makes it easier to tell what stimulus elicited a given ERP component. In behavioral studies, on the other hand, both attended and unattended stimuli are usually presented at the same time, which is necessary when effects of unattended stimuli take the form of a modulation in the behavioral response to an attended stimulus.

The most basic finding of attention studies using ERP measures is that components arising very early after the presentation of a stimulus (within the first several hundred milliseconds) often differ as a function of whether the stimulus is attended or unattended (see Hillyard and Picton, 1987, and Naatanen, 1992, for reviews). In the auditory modality, the divergence between waveforms elicited by attended and unattended stimuli generally begins approximately 60 to 80 msec after presentation, and in some cases earlier divergence has been observed (Hillyard and Picton, 1987). The divergence is commonly known as the Nd, or negative-difference component (Hillyard et al., 1973). The time at which the Nd begins seems to depend on the difficulty of the discrimination required for selection and not response (Hansen and Hillyard, 1980): the harder the discrimination, the later the divergence.

It was initially proposed (Hillyard et al., 1973) that the Nd component simply represented a dampening of the exogenous N1 (first negative) components produced by unattended auditory stimuli (exogenous refers to components little affected by psychological state<sup>6</sup>). This would be consistent with the view that the Nd reflects reduced sensory and perceptual analysis of unattended stimuli or enhanced processing of attended stimuli. The fact the Nd begins earlier when the selection is easier fits with this interpretation. However, some investigators believe it may largely reflect an independent endogenous (psychologically influenced) component superimposed on the N1 (Teder et al., 1993; Woods, 1990). In this interpretation, when a stimulus is attended, the selection itself produces negativity, which is then added into the ERP waveform. If this is correct, the Nd effect might not reflect any change whatever in the perceptual analysis of unattended stimuli. Recent neuromagnetic studies that provide more detailed anatomical information suggest a measure of truth in both views, in that the Nd probably reflects both an endogenous selection-related effect and a reduction in exogenous responses beginning as early as 20 msec after the stimulus (Woldorff and Hillyard, 1991).

Effects of spatial selection on visual evoked responses have also been observed. Visual stimuli typically elicit two early components nicknamed P1 and N1, the first positive and negative components, generally occurring between about 100 and 200 msec after the stimulus. These effects are considered fairly exogenous, that is, they occur even when the subject passively watches stimuli flash by on the screen. Both the P1 and the N1 are smaller when a stimulus is ignored based on its location. Unlike the case of auditory attention, there seems to be little doubt that these changes reflect modulations in the magnitude of underlying exogenous components as a function of spatial attention, rather than selection-related components superimposed on the underlying exogenous ERP components.

A different and quite intriguing pattern of results is found with tasks requiring selection of visual stimuli based on color. Consider what happens when a subject is instructed to note only red bars and ignore green ones (responding on the basis of the length of the red bars). A red bar does not produce a bigger N1 or P1 than a green bar produces. Rather, as with auditory spatial selection, one sees an additional negativity superimposed on top of N1 and P1 for the color-relevant stimuli. Most remarkably, this color-selection component is greatly diminished when the stimulus is in an irrelevant spatial location (Hillyard and Münte, 1984). Thus, the ERP signatures of location and color selection are distinct and have a hierarchical arrangement, with location selection dominating color selection.

A rather appealing functional interpretation of these results would be as follows. The magnitude of N1 and P1 might reflect the degree of sensory and perceptual analysis (the more analysis, the bigger the component). Stimuli in unattended locations are not analyzed very deeply, because they are subject to early gating, hence, smaller N1 and P1; in fact, not even their color is analyzed. When stimuli are in the attended location, their color is analyzed, and if they are determined to be of the relevant color, a different selection process then takes place, indexed by the color-selection negativity.

This scheme fits the data quite well, but one should exercise some caution in accepting it. For one thing, the P1 and N1 components might not reflect perceptual analysis, but rather a selection process that fetches information from the perceptual machinery. The fact that the changes

involve modulation of the magnitude of relatively exogenous components, which arise even during passive viewing, does not rule this out; even in passive viewing, a fetching process might occur. Other alternatives might also fit the data. For example, the N1 and P1 components could reflect collateral signals sent out by visual areas signaling other brain areas to get ready to receive information, or they might have "house-keeping" functions, readying the visual areas themselves to receive and process further sensory stimulation.

These alternatives demonstrate something that is rather obvious but easily overlooked: inferences about perceptual gating rely on a set of assumptions about the functional meaning of the components and their magnitudes. Reductions in N1 and P1 imply perceptual gating only if the magnitude of the components can be assumed to indicate the extent of processing. The fact that the underlying generators of the components can be localized to extrastriate visual areas where perceptual analysis is likely to be taking place (e.g., Mangun, Hillyard, and Luck, 1993) provides some evidence for this assumption; it is, however, equally consistent with several of the alternatives mentioned above. This basic problem applies equally to both auditory and visual ERP results, of course. This important caveat notwithstanding, the various ERP results are certainly congenial to the idea of early perceptual filtering, and can reasonably be described as converging evidence for it. If it were possible to test the usual assumptions about what the magnitude of an ERP component signifies, these results might provide compelling rather than congenial support.

In summary, both invasive and noninvasive neural measures suggest that rejected stimuli may undergo less extensive sensory/perceptual processing than attended stimuli, although the interpretation of physiological measures is in some ways an open question. This obviously implies no criticism of the investigators who carried out these technically difficult studies. It seems likely that physiological measures will provide increasingly important constraints on attention theories. Before this can happen, however, new evidence is required to clarify the relationship between information processing operations and the various physiological measures. Although physiological indexes do not yet provide unambiguous information about perceptual analysis of unattended stimuli, there is little doubt that further developments combining psychological and physi-

ological measures and manipulations should provide stronger tests than either approach alone. Promising developments are emerging along these lines, such as studies combining psychophysical and single-unit recording observations in monkeys (e.g., Salzman et al., 1992).

### Consequences for Early Visual Processes

The studies discussed thus far mostly focused on the process of identifying familiar objects. The visual system provides us with a great deal of information about our environment besides which familiar objects are present. Even in an abstract sculpture garden we readily and effortlessly perceive much about the spatial layout and reflectance characteristics of visible surfaces, and can even work out the three-dimensional positions of objects that are partly occluding each other (Marr, 1982; Rock, 1983). The majority of the stimuli and tasks studied by attention researchers place few demands on these perceptual mechanisms because they use stimuli such as high-contrast letters and digits in a fixed type font. Such stimuli are ecologically valid in the sense that reading is an important human activity; on the other hand, findings regarding identification of such impoverished stimuli may or may not generalize to other more demanding perceptual tasks.

It is an interesting fact about our visual system that we can identify objects whose contours are signaled by any of a large number of different cues (see Cavanagh, 1987). One illustration of this is the random-dot stereogram (Julesz, 1971) in which a portion of a random dot pattern is shifted in the pattern shown to the left eye compared with the pattern shown to the right eye. This produces the percept of an elevated (or recessed) surface. If the surface is shaped like the block letter B, for example, one can read the B without difficulty.

To identify three-dimensional objects, one might suppose the visual system normally computes an internal description of three-dimensional shape, and then compares this description to descriptions of familiar objects that are stored in memory (Marr, 1982; Biederman and Ju, 1988). The visual system may be more flexible than this, however. For one thing, the ability to analyze three-dimensional shapes and the ability to identify familiar objects tend to be lost together in visual agnosia (Farah, 1990). Farah suggests that computing a description of a shape and recognizing

that shape are completely intertwined in the cortex. (Of course, the mechanisms might be anatomically intertwined but still functionally separate or sequential.) Another sign that the sequential model may be wrong is the fact that recognition of even three-dimensional objects such as faces can be achieved with images that provide little or no information about three-dimensional layout (Witkin and Tenenbaum, 1983). Therefore, one should not assume that analysis of three-dimensional form is an obligatory intermediate stage en route to object recognition, as Marr suggested.

Despite these caveats, the fact remains that a great many visual mechanisms are devoted to analyzing three-dimensional shape, and these mechanisms can provide input to object recognition. It is reasonable to ask how they are affected by attention. The unselective perception hypothesis, interpreted most broadly, would claim that they process excluded stimuli to the same extent and in the same way they process attended stimuli.

Few investigators have examined this issue, and those who have done so generally used visual aftereffects as their primary measure (e.g., Shulman, 1990). Prolonged viewing of stimuli that excite particular detectors in the visual system changes perception of subsequently presented stimuli (see Barlow, 1980). The best-known example is the motion aftereffect, commonly called the waterfall illusion. When one looked at an adapting display of continuous motion such as a waterfall for 30 seconds or so, and then views a static display such as an adjacent hillside, stationary contours appear to drift in the opposite direction. This effect probably originates at least partly in cells or synapses in area MT, the cortical area specialized for motion perception (Wenderoth, Bray, and Johnstone, 1988). Similar aftereffects can be observed for other visual dimensions such as orientation and spatial frequency (Graham, 1989).

Several studies asked whether the simple motion aftereffect can be reduced when subjects ignore the adapting field and attend to something else; they produced conflicting results. An early report by Wohlge-muth (1911) found no reduction in adaptation. Chaudhuri (1990) had subjects view a drifting texture with a small aperture in the center containing a character that changed several times a second. In the passive condition, subjects simply looked at the whole display; in the attention-diverted

condition, subjects monitored the characters looking for the occasional digit to which they made a keypress response. The aftereffect was assessed by having observers view a stationary test field. The duration of the aftereffect was substantially reduced in the attention-diverted condition. Lankheet and Verstraten (1995) reached similar conclusions using three observers (themselves and one other individual) viewing displays of two-component transparent motion. Given the subjective nature of these judgments and the conflict with Wohlge-muth's findings, effects of selective attention on the motion aftereffect require further confirmation with larger numbers of naive observers.

In a series of interesting studies Shulman assessed the attentional modulation of aftereffects involving perception of three-dimensional structure. After prolonged viewing of a cube with a three-dimensional structure that is unambiguous, an ambiguous Necker cube tends to be seen as having a perspective opposite that of the adapting figure. During the adaptation phase of Shulman's experiment, twelve subjects viewed two superimposed unambiguous figures differing in size and color. They attended to only one of the figures, reporting any changes in its color. In the test phase, subjects' perception of the Necker figure tended to show an adaptation effect induced by the attending adapting figure rather than the unattended adapting figure (Shulman, 1993). Shulman obtained very similar results for ambiguous rotations in depth (Shulman, 1991) and ambiguous staircase figures (Shulman, 1992).

On the other hand, ignoring the adapting stimulus does not seem to reduce another visual aftereffect called the McCollough effect. This is an orientation-specific color aftereffect that can be produced by viewing two grating patterns in opposite colors and different orientations, such as a green patch of horizontal stripes and a red patch of vertical stripes. After several minutes of adaptation, an observer will tend to see colorless grating patches as having a tint opposite that of the adapting patch that was presented at the same orientation. Houck and Hoffman (1986) had subjects view displays of colored grating patches, focusing their attention on some of the patches and ignoring others. The size of the adaptation effect was the same whether the patch was attended or not. The McCollough effect may stem from adaptation of mechanisms early in the visual pathway, before the site where information from the two eyes is

brought together. Therefore, it may not be surprising that this effect would be unaffected by attention, whereas those mentioned in the previous paragraphs would be affected.

The observation that attentional instructions can determine the magnitude of perceptual aftereffects reinforces the other evidence described in this chapter that questions the idea that excluded stimuli are perceptually analyzed to the same degree as attended stimuli. To get a fuller picture, similar comparisons must be performed with a wider variety of aftereffects, including tilt and size. Most of the studies described above involve motion perception, which probably depends on processes occurring in the dorsal stream of visual processing, rather than the ventral stream most critically involved in object recognition. It seems plausible, therefore, that attentional gating may reduce or suppress perceptual analysis carried out in both cortical visual streams.

### Visual Selection and Eye Movements

It has long been noted that people can attend away from the point of fixation and shift their attention without moving their eyes. Research conducted in this century attests to people's ability to shift from attending to one stimulus to another without making an overt eye movement (e.g., the partial-report experiments of Sperling, 1960). However, this does not imply that eye movements and shifts of attention are unrelated. In fact, there are obvious reasons why one would expect some functional ties between them. The acuity in the center of the fovea is far greater than at the periphery, so when visual input from a particular location in a scene is of special importance, it generally makes sense to foveate this location and improve the quality of information one acquires.

People are barely aware of making eye movements at all, so it is not surprising that they are usually unaware that the movements come in several varieties. Researchers commonly distinguish among five or more oculomotor systems, each with supposedly characteristic forms of eliciting stimuli, different movement trajectories, and different underlying neural systems. The three best-known types are saccadic, smooth-pursuit, and vergence eye movements. Saccadic eye movements are rapid and abrupt. Smooth-pursuit movements allow the eye to track relatively slowly moving stimuli up to about 20 or 30 degrees per second. With

vergence eye movements, the two eyes move in opposite directions to permit fixation on stimuli farther or nearer along the line of sight. Recently, however, some theorists have challenged this proliferation of systems, proposing a more fundamental demarcation between the fast saccadic system and the slower smooth eye movement system, and arguing that finer taxonomies reflect different ways in which these systems can be driven (Steinman, Kowler, and Collewijn, 1990).

Until recently, most experts wrote about eye movements as if they were completely devoid of psychological content or interest. For the most part, investigators analyzed the (sensory) input and (eye movement) output of particular eye movement systems quantitatively, trying to characterize *the* input-output transformation achieved by the system. The experimental situations were expressly designed to minimize the role of attentional factors, together with anticipation or other mental processes. However, recent work demonstrates that eye movement control has rich interconnections with selective attention. We begin with saccadic eye movements.

**Saccadic Eye Movements** We can voluntarily shift from attending to one spot in the visual field to another without producing a saccadic (or other) eye movement. Saccadic eye movements are at least potentially voluntary in the sense that one can will the eyes to move to a chosen point in the visual field and the movement will take place. Selecting a target for a willed or unwilled eye movement involves specifying an individual location in the visual field, probably based on some preliminary information about the stimuli present at that location. It seems reasonable, therefore, that shifts of attention to the intended destination of an eye movement might be involved in or necessary for preparation of a saccadic eye movement. This will be termed the mandatory shift hypothesis: a shift in attention is mandatory before execution of a saccade.

The matter seems straightforward enough, but initial studies painted a rather confusing picture. Remington (1980) used a peripheral target to summon the eye movement, and inserted a visual probe (requiring a speeded detection response) in the same or opposite direction as the eye movement to measure the allocation of visual attention (see chapter 4 for a discussion of this method). He found faster responses to probes that

were located in the same direction as the eye movement compared with probes in the opposite direction. Under these conditions, though, the peripheral target might have produced the attention shift, rather than the preparation of the eye movement. A better test, therefore, would allow the subject to prepare to make an eye movement in a prespecified direction and then probe in some way to see if attention was indeed shifted to the destination.

Around the same time, Klein (1980) carried out an experiment in which subjects made either an eye movement (the direction was fixed throughout a block of trials) or a speeded detection response to a peripheral target. In dual-task blocks, subjects did not know which task they would have to perform, although they did know where the eye movement would go, if one should be called for. Responses were no faster when the peripheral target occurred at the position in which the eye movement would have been directed, compared with when the target appeared in the opposite position. On the basis of this finding, Klein rejected the mandatory shift hypothesis. However, it is worth noting that the subject only had to make an eye movement if an asterisk appeared, and the asterisk could appear on either side of the screen. Essentially, then, the subject was asked to prepare a leftward eye movement, and also to monitor both left and right positions for an asterisk, which would indicate that the eye movement should be executed. By the mandatory shift hypothesis, these requirements are inherently problematic. Therefore, one might well expect that subjects would not prepare the eye movement until they were sure it would be necessary. Consistent with this interpretation, saccadic latencies were unusually long for prepared eye movements.

More recent studies avoided these problems and yielded strong and direct support for the mandatory shift hypothesis. Shepherd, Findlay, and Hockey (1986) presented a central arrow cue indicating whether the subject should move the eyes leftward or rightward. After a delay (stimulus-onset asynchronies 70–550 msec), a square was illuminated either to the left or right of center, and the subject made a speeded button-push response on detecting this square. In some blocks of trials the square was usually to be found in the direction indicated by the arrow cue, placing it at the destination for the eye movement. In other blocks the arrow provided no information about where the square was likely to be, and

in still others the square was usually located in the opposite direction from the one to which the arrow pointed. Detection responses were fastest when the square was in the location where the eye movement would go (the direction in which the arrow pointed). This was true even in blocks where the most probable position for the square was opposite the direction in which the arrow pointed. In control conditions where the arrow cue was present but no eye movements were required, detection responses were fastest on whichever side was most likely given the arrow. Thus, subjects were perfectly capable of using the arrow cue to shift attention either toward or away from where it pointed, but not when making an eye movement in the direction it pointed.

The findings of Shepherd et al certainly indicate that subjects cannot simultaneously move their eyes in the direction indicated by an arrow and (fully) shift their attention in the opposite direction, even when it would be to their advantage to do so. Furthermore, the fact that attention shifts toward the eye movement destination appear when the arrow says nothing about the likely position of the probe, is quite consistent with the mandatory shift hypothesis.

It would obviously strengthen the case if preparing an eye movement entailed an attention shift when the subject was not presented with such complicated demands. The probe-RT dependent measure used in these studies also raises problems of interpretation (see chapter 4). It would seem that an optimal and straightforward test of mandatory shift would combine rapid fixed-direction eye movements with an attentional measure that involved perceiving forms in a brief display.

Hoffman and Subramaniam (1995) recently carried out just such a test. They gave their subjects ample time to prepare to make an eye movement toward a prespecified one of the four corners of an imaginary square, and a tone provided a go signal for this eye movement. After a very brief interval from the tone onset (0, 50, or 100 msec), a display of four letters was briefly flashed. Subjects made the eye movement and then reported which of two possible targets was present in the display (search task). The targets for the search task were equally likely to be in any of the four positions. Nonetheless, subjects were substantially more accurate when targets were presented at the destination of the eye movement. Saccade latencies were very rapid, indicating that subjects were preparing

the eye movement diligently. In another experiment, the subject always made saccades in a given fixed direction when the tone sounded (this direction did not change throughout the whole session). At the beginning of each trial, however, and shortly before the tone, an arrow cue indicated the most likely position of the target for the search task. As predicted by the mandatory shift hypothesis, target detection was most accurate when the target was located at the destination of the eye movement. Even more compelling was the fact that when the letter target was in a position *other than* where the eye moved, detection accuracy did not depend on whether or not this position had been cued. Deubel and Schneider (1996) confirmed these findings using a slightly different design. Kowler et al (1995) added further support, demonstrating that when people make a saccade in the direction of a pointer, the requirement to identify a letter not located at the target prolonged saccadic RTs by 50 to 70 msec.

These conclusions may be relevant to more naturalistic tasks such as reading. Inhoff and Brihl (1991) had subjects read passage presented in an upper line of text on a CRT while ignoring another passage contained in a line of text immediately below. Subjects were given multiple-choice questions about both relevant and irrelevant passages. Evidence suggested that subjects picked up material from the irrelevant material, but only on those occasions when they inadvertently fixated on that material.

In summary, evidence reveals that preparing to move the eye to a particular position in the visual field requires that visual attention first shifts to that position. Data suggest that this shift occurs during the time the system is actually preparing to produce the eye movement. Besides its theoretical interest, this conclusion raises intriguing questions about human performance in tasks requiring rapid responding. For example, drivers and pilots often glance toward instrument panels. Based on the results described, one might expect that a driver would be unable to detect an oncoming pedestrian for 100 msec or more preceding a glance, even if he or she was staring right at the person.<sup>7</sup> Since saccades often occur two or three times per second, one is led to wonder if, for a good portion of the time during which a person surveys a scene, the person is not really using information from the point of fixation.

An intriguing study by Blanchard et al. (1984) lends credence to this seemingly implausible speculation. The researchers monitored eye move-

ments while subjects read computer-displayed text. A short time (50–120 msec) after the subject began to fixate a critical word (e.g., “tomb”), the display flickered to a mask (Xs replacing letters) for 30 msec, at which point new text appeared with a replacement word (e.g., “bomb”) occupying the position formerly occupied by the critical word. Both the critical word and the replacement word made sense in the context of the sentence the subject was reading. Subjects were asked if they saw anything change and then given a recognition test for words from the passage. In the great majority of trials they appeared to have processed only one of the words, with the replacement affecting neither their recognition judgments nor their saccadic latencies. When the replacement occurred after 50 msec exposure of the first word, the word they recognized was usually the second word, but when the replacement occurred after 120 msec, it was usually the first.

Blanchard et al postulated a brief crucial period during a fixation in which the visual information is used. Given the research reviewed above, one explanation for this would be that during the later part of the fixation, planning the next eye movement requires a shift of visual attention away from the currently fixated word. Although this account may be correct, the data of Blanchard et al suggest that the picture cannot be quite that simple, since in 12% of trials in which the second word was reported, the replacement word was present only during the last 30 msec of the fixation.

**Smooth-Pursuit and Vergence Eye Movements** The most striking difference between smooth-pursuit eye movements and saccades, aside from the fact the former occur more slowly, is that one cannot generate smooth eye movements by a sheer act of will in the absence of appropriate stimulation. The reader can verify this by simply staring at a wall and trying to make his or her eyes pan from left to right. The result is a series of short abrupt (saccadic) eye movements, rather than a smooth-pursuit movement.

The usual trigger for smooth-pursuit eye movements is a stimulus that slowly drifts along in the visual field. This does not mean that the smooth-pursuit system is devoid of psychological interest. On the contrary, these movements are intertwined with mental events in an

intriguing way. For one thing, they can be driven by anticipation as well as direct sensory input. When a person expects a spot to begin moving at a certain time, the eye often begins moving before the spot does (Kowler and Steinman, 1979). Kowler (1989) carried out some elegant experiments pitting the expectation of the future path of the target against experience with its trajectory; the results showed that expectation rather than habit governed the eye movements. It may sound odd to say that these eye movements cannot be initiated voluntarily but that they are nonetheless driven by conscious expectations. Can one not form an expectation by sheer will? Apparently not, and other involuntary response systems also seem to illustrate the point. For example, changes in GSR are produced by conditioned stimuli that predicted shock in shock-conditioning experiments. These responses cannot be produced voluntarily and they depend on conscious expectations; however, when expectation is pitted against experience, expectation determines the response (Brewer, 1974).

More recent research, however, shows that anticipation of the trajectory of a stimulus is not a sufficient condition for eliciting a pursuit eye movement. An object must also be *attended* to be pursued. Kowler et al. (1984) showed subjects two superimposed fields of dots, one drifting and the other stationary. Observers could voluntarily choose which set of dots to track. This was true even when the density of the dots was so great that individual dots tended to be obscured by those from the other field. Variation in the relative luminance of the two sets of dots, and other stimulus factors, made little difference to the result. If subjects were indeed controlling which set of dots they tracked by selectively attending to the dots (in the same sense we spoke about attention in other contexts), then one would expect this to have perceptual consequences. Kowler and Zingale (1985) used essentially the same displays, but occasionally turned off a subset of one of the fields of dots. The subjects' task was to report when dots disappeared from either of the two fields. They were substantially faster and more accurate at doing this when dots disappeared from the field that was being tracked than from the untracked field. It might be objected that this difference could reflect the retinal smear of the untracked field. Kowler and Zingale recorded some plausible arguments against this possibility, but further investigation would be useful.

Older observations about voluntary control support the attentional interpretation. For example, try having someone move one hand in front of your eyes while you fixate directly on the hand; meanwhile, attempt to *prevent* your eye from tracking the hand. You will probably find that you can do this, but only by focusing your attention on objects in the surrounding field, as Mach noted in 1906 (cited by Kowler and Zingale, 1985). This would seem to imply that one can attend to outer regions of the field without attending to the central object (e.g., the hand). Can numerous noncontiguous elements actually elicit smooth pursuit even when the foveal stimulus follows a different trajectory?

Collewyn, Curio, and Grusser (1982) examined a form of pursuit eye movements induced by flashing, rather than moving stimuli, termed Sigma-pursuit movements (SPMs). In one condition, two horizontal rows of dots were presented in such a way that either one alone would induce SPMs at different velocities. The striking finding was that observers were able to foveate one of these inducing rows and yet, by attending to the other row, to produce SPMs appropriate for the other row.

Steven Yantis and I informally explored the limits of attentional control of ordinary smooth-pursuit eye movements using similar designs. We had people view two columns of spots moving from the top to the bottom of a CRT screen while a column of spots in between the two columns moved upward. The object was to attend to the outer columns and ignore the central column. In many cases, subjects' eyes appeared to move at the rate of the outer dots although the dots that were being fixated were moving in the opposite direction (Yantis and Pashler, unpublished observations). Most subjects found the task to be rather difficult. In summary, various observations, starting with those of Ernst Mach, demonstrate the fact that which stimuli guide pursuit eye movements depends to a great degree on voluntary control, which seems to be closely tied to attentional selection.

In vergence eye movements, the two eyes move in opposite directions to fixate on stimuli nearer or closer to the observer. Erkelens and Collewyn (1991) presented subjects with a pattern composed of multiple dots or sets of lines to each eye. The positions of various elements in the field differed slightly (to put it more technically, the elements had different

retinal disparities). When the observers were instructed to attend to a particular form and its image in one eye was perturbed, that eye moved to maintain convergence on the form. These attention-dependent vergence movements could occur even when unattended stimuli were projecting onto the fovea.

In summary, it appears that the three most well-known types of eye movements—saccadic, smooth-pursuit, and vergence—are closely tied to visual selective attention. Preparation of saccadic eye movements seems to require shifting visual attention to the destination of the movements during a period of several hundred milliseconds immediately preceding the saccade. Combined motion and/or displacement signals from the elements in the display that are attended seem to combine to determine the path of smooth-pursuit eye movements. Vergence eye movements work to foveate attended, but not unattended, elements in the visual field. As Kowler and Zingale (1985) point out, the fact that smooth-pursuit eye movements provide an outward manifestation of selective attention suggests they may be a valuable tool for the study of visual attention. Researchers have barely begun to exploit this tool.

### **Bimodal Selective Attention**

People are always (or almost always) subject to stimulation in different sensory modalities. In light of that fact, it is surprising that the relation between selection in one modality and selection in another has been relatively little investigated. The standard experimental paradigms in attention research use either visual stimuli or auditory stimuli, but not both, and most researchers have stuck to one or the other. Two issues are related to bimodal (or potentially multimodal) stimulation that are of obvious importance. The first is the relative difficulty of what one might call intramodal selection, selecting a stimulus in one modality and excluding other stimuli in the same modality, compared with intermodal selection, selecting a stimulus in one modality while excluding stimuli in a different modality. The second issue is whether simultaneous intramodal selection in different modalities is possible, and if so, how selection in one modality relates to selection in the other. For example, are they linked in space or some other way?

Intramodal versus intermodal selection can be addressed with filtering tasks that require selection of stimuli in one modality and rejection of stimuli in another. Anyone who has read a newspaper in a crowded subway or tried to work within earshot of a stereo can confirm that this issue has real-world validity. Obviously, we can accomplish this kind of selection with some success. What is not obvious is how effectively we can do it compared with excluding stimuli within the modality we are focusing on.

Various different hypotheses seem reasonably plausible. On the one hand, people might be capable of simply turning off processing in an unwanted modality more effectively than they can prevent processing of unwanted stimulation in a modality that they must simultaneously monitor. After all, we have seen that greater differences between selected and excluded stimuli tend to make selection easier, and intermodal selection could naturally be viewed as an even more gross basis for selection than, say, pitch or location. On the other hand, it has sometimes been suggested that rejected stimuli are more extensively processed when the selected channel contains little information to process, and thereby leaves resources idle (Helmholtz, 1867/1968; Kahneman, 1973, chapter 2). If this line of thought is valid, and if auditory and visual processing resources are at to some degree separate (a point that will be supported by evidence described in chapter 3), then excluding stimuli in a different modality might be less effective rather than more.

The few investigations that have been carried out do not seem to demonstrate that people are especially effective in shutting out stimuli in a particular modality when everything in that modality is irrelevant. To give one example, Greenwald (1970) played tapes to subjects while they read visually presented digits aloud. The tape contained either another digit or a tap at the same time as each visual digit. Subjects' responses to the visually presented digits were slowed by about 35 msec when accompanied by another digit, and the effect did not depend greatly on the rate at which the digits were spoken. Irrelevant speech substantially, although not massively, disrupts immediate memory for visually presented materials (Salame and Baddeley, 1982). Unfortunately, no data allow one to compare intra modal and intermodal selection. Doing so would undoubtedly be difficult, but the importance of the issue warrants trying.

A second broad issue about bimodal selective attention concerns the relationship between the mechanisms of selectivity within different modalities. Can people simultaneously select, say, the spoken messages coming to the left ear, ignoring the right ear message and the words colored in a CRT display, ignoring the green ones? Or is selectivity possible within only one modality at a time? Can someone process the visual stimulus coming from one location and the auditory stimulus coming from another location, ignoring distractors in each modality at the other location? The issue spans the topics of selective and divided attention; involving divided attention between modalities and selective attention within each modality.

In many mammalian species, even infants respond to unexpected sounds with an orienting response involving the whole body, triggered by the binaural localization of the sounds (Kelly and Potash, 1986). This response is hardly surprising from a functional point of view; sounds outside an animal's field of view may be very significant and vision is critical in determining this significance. Some authors even suggest that the evolution of auditory localization was largely driven by the demands of orienting attention (Heffner and Heffner, 1992). When one observes overt behavior, therefore, one sees plentiful evidence of cross-modal influences on attention. It does not follow, however, that these linkages are inherent and obligatory, much less that there is only a single supramodal spatial attention controller. The linkages that show up in orienting might merely reflect a default setting that can be overridden, or they might be restricted to unexpected sounds coming from locations out of view. To distinguish these possibilities, one must construct situations in which people have an incentive to avoid attending in the same way to visual and auditory stimulation.

Reisberg, Scheiber, and Potemken (1981) had subjects listen to a word list from one loudspeaker while another loudspeaker delivered distractor messages. Subjects were not able to remember as many of the relevant words and showed more intrusions of irrelevant words when they fixated their eyes on the loudspeaker playing distractors than when they fixated on the relevant loudspeaker or a silent speaker. However, the effect was small, and in a set of follow-up studies, Wolters and Schiano (1989) were unable to replicate the effect. Eye position (which of course is not the

same as visual attention) is not apparently a potent determinant of auditory selection.

A more direct approach to this question is to present several signals in two modalities at once, and see if people can select auditory stimuli from one part of space while they select visual stimuli from another part. Driver and Spence (1994) carried out an experiment using two speakers and two CRT displays. One speaker and display sat at approximately the same location to the left of the subject, and the other pair sat at a corresponding position to the right. Two tasks were required: shadowing the message coming from one speaker while ignoring the message coming from the other, and monitoring a continuous visual display of characters on one of the CRT displays. A significant decrement in shadowing performance occurred when the visual display was on the opposite side as the speaker, rather than the same side. However, the magnitude of the difference was not especially large. Alfonso (1992) performed similar experiments in my laboratory. Using a slightly easier shadowing task, she found that subjects could readily attend to the visual stimulus on one side and the auditory stimulus on the other. Unlike the findings of Driver and Spence, there were no significant costs of attending visually to one side visually and auditorily to the other in this experiment, but the trend was in the same direction.

Another way of approaching the same issue is to present a cueing stimulus in one modality and a target requiring a response in another (the mechanisms behind these spatial cueing effects are explored in detail in chapter 4). In several studies of this type, responses to a visual target on the left or right of a screen were speeded up by auditory cues from speakers positioned on the corresponding side (Buchtel and Butter, 1988) and also by tactile cueing (Butter, Buchtel, and Santucci, 1989). As with the orienting behaviors, this result does not demonstrate an obligatory linkage: the cues predicted where the target would occur, and therefore subjects would have reason to select on this basis even in there were no intrinsic connections. (If the odor of camphor had indicated to subjects that a target on the right was likely, performance might have been improved; this would not implicate an obligatory olfactory-visual linkage.) Spence and Driver (1996) reported, however, that unpredictable auditory cues do produce a brief facilitation of responses to visual stimuli

in adjacent positions. Of interest, they did not find visual cues to have the same effect, and suggested that this asymmetry may reflect the different functional roles of vision and audition alluded to above. Together, this evidence implies that some partial auditory-visual yoking is probably obligatory. Alfonso's experiment may have been insensitive to some small effect of this type.

The issue of supramodal attention has also been considered by neuropsychologists. Patients with damage to the parietal lobe frequently show a tendency to react weakly or slowly or not at all to information coming from the side of space opposite to their lesion (Heilman, Watson, and Valenstein, 1993). In many cases they are able to perceive a stimulus on their "bad" side as long as no other stimulus is presented, but when stimuli are presented to both sides, they do not perceive the stimulus on the bad side. This symptom, which may occur in various sensory modalities, is referred to as extinction. De Renzi, Gentilini, and Pattacini (1984) found that the severity of extinction in one modality did not very strongly predict the severity of extinction in another modality. By itself, this does not rule out the idea that a single center is responsible for orienting attention to space governing *all* sensory modalities, since the lack of correlation might be due to disconnection of attentional input to such a center, as pointed out by Farah et al. (1989). They found that extinction, as indicated by delayed responses to targets presented in the bad hemifield, could be induced by auditory as well as visual stimulation in the good hemifield. Unlike the experiments of Butters and colleagues, Farah et al. induced extinction with cues that did not predict the stimulation, so these results are not subject to the sorts of problems described earlier. These various results again suggest a fairly strong linkage between the machinery controlling attention to space in the two modalities.

Further study of this issue necessary before one can draw strong conclusions about how selectivity in different perceptual modalities is related. At present, evidence exists for some degree of obligatory linkage. Deliberately attending to objects in one location, or detecting auditory stimulation in a location, seems to favor selectivity in the other modality to that location. On the other hand, it seems clear people can decouple selection in two modalities to a substantial degree, arguing against a single supramodal spatial attention controller.

### The Time Course of Selection

In many of the filtering studies the subject was informed about the criterion of selection by a stimulus of some sort. For example, in von Wright's studies, a bar marker next to a character indicated to the subject that this stimulus should be attended and other stimuli should be rejected. How long does it take to use such a cue to initiate selective processing by location? The best parametric information on the time course of selection comes from pioneering studies by Charles Eriksen and his colleagues. These researchers systematically examined spatial cueing in selection tasks involving arrays of characters; generally, subjects were asked to name or make a selective response to the cued character. These studies are especially useful because they provided converging evidence involving both response time and accuracy measures. Furthermore, unlike the more widely copied design involving simple detection of just a single spot (see chapter 4), the experiments required subjects to identify objects and to filter out distractors, which makes the attentional demands of the task much more apparent.

Eriksen and Collins (1969) presented subjects with brief circular displays of six letters, followed or preceded by an indicator of some kind that specified which letter should be reported. Various kinds of spatial cues were used, including indicators on the opposite side of the display from the cued position. Accuracy improved as the cues were presented earlier, up to around 100 msec before the display. Another experiment showed that the benefit seemed to reach a maximum value at approximately 200 msec. The RTs to vocalize the cued letter declined as the SOA from cue to display was increased to about 250 msec, and flattened off thereafter (Colegate, Hoffman, and Eriksen, 1973). Uninformative warning signs and spatial cues also were given at the same time as the display to show that this was not due to the alerting effects of the cues.

With displays of twelve items (Eriksen and Hoffman, 1973), RTs decreased substantially between probe-display SOAs of 150 and 250 msec. It seems possible, then, that the more spatial precision required, the longer it takes to set the selectivity in place. Eriksen and Rohrbaugh (1970) were able to find only equivocal support for this idea, however. Sperling and Reeves (1980) developed another, very ingenious, approach

to measuring the time required for attention shifts. Subjects monitored a stream of sequentially presented characters looking for a target character. When they detected this target, they immediately attended to a second stream of characters and tried to remember the first several items they could “catch.” By examining which items were reported from this second stream, Sperling and Reeves were able to compute the time between the onset of the character triggering the switch and the completion of the switch. This method yields slightly longer estimates of shifting time than the other approaches described, presumably because the task imposes greater perceptual and cognitive demands.

Together with some informative results, one unfortunate misimpression that came out of this research program deserves mention. In some articles, the authors suggested that the focus of attention might encompass a fixed extent, such as 1 degree of visual angle (e.g., Eriksen and Eriksen, 1974). This conclusion, based on observations regarding flanker effects and noise characters, has often been quoted. It is easily refuted without making detailed measurements in the laboratory. Consider a spatial-selection task like the one used by Eriksen and colleagues in which a bar probe sits next to one item in a display of eight letters. Observers with good eyesight can see such a display from such a long distance the *entire* display lies within 1 degree of visual angle, yet they can choose only the probed item. Readers with good visual acuity can verify this by holding this text several feet away and reading aloud the last letter in each word. Relative distance may affect the efficiency of selective attention and the exclusion of irrelevant stimuli, but spatial attention is far too flexible to be described by a single parameter in this way.

## Conclusions

### Selective Attention in Vision and Hearing

This section focused on perceptual filtering tasks in which numerous stimuli are presented, and an observer or listener attempts to identify just those that satisfy a simple, physical selection criterion such as spatial location, pitch, or color. Research has ranged over a wide variety of such tasks involving both visual and auditory modalities. Several basic generalizations can be stated with reasonable confidence. The first of these was

well known to Cherry, who helped launch the modern era of attention studies, although it was not clear when he was writing how robust this generalization would prove to be.

1. When people try to focus on certain stimuli and ignore others, they generally notice and report only relatively gross physical properties of the rejected stimuli.

Generalization 1 seems to hold true over a wide range of different observation conditions, not only with speech presented to the two ears (Cherry’s design), but also with such diverse stimuli as musical passages and overlapping video images. However, over the almost four decades that have elapsed since Cherry made his observations, a variety of subtle measures have been used to show that rejected stimuli are sometimes processed to a semantic level, findings sometimes described under the heading of breakthrough of the unattended. These examples illustrate a second generalization.

2. Under conditions that allow effective selection, as indicated by people’s reports and lack of memory for unattended stimuli, it is often possible to show that *some* semantic analysis of rejected stimuli still occurs, at least in *some* trials.

At first, findings of this sort were taken to support the concept that identification of stimuli is wholly unselective, as late selection theorists proposed. However, many subsequent studies, using both auditory and visual stimuli, eroded this interpretation. A third generalization appears well supported.

3. When favorable conditions for attentional selectivity are provided, evidence of unselective processing is unconvincing.

Examples of factors that provide optimal conditions for selection include easy discrimination between relevant and irrelevant stimuli and adequate time to maintain a consistent criterion of selection. There is also moderate support for a further generalization.

4. Increasing the number of *rejected* stimuli reduces and even eliminates their effects, as indicated by the indirect measures.

The findings of Yantis and Johnston (1990) and Kahneman and Chajczyk (1983) provide the most clear-cut examples of this generalization, but there are others. This result argues against the hypothesis that