

Chapter 2

Attention and Visual Perception:

Analyzing Divided Attention

Harold Pashler

Attention is one of the most active areas of experimental research in cognitive science. This chapter focuses on an aspect of attention that has been studied intensively over the past twenty-five years or so: capacity limitations on our ability to perceive multiple visual inputs. The basic questions are simple but far-reaching. How much visual information can we take in at one time? What can we do with this information? Do we recognize objects one at a time, or can we recognize a large number simultaneously? This chapter describes several different methods that investigators have used to address these questions, along with some of the most important conclusions to emerge from this research.

2.1 Attention in Ordinary Language and Psychology

Attention is a common word in ordinary language as well as the name of a held of study. Other research areas in cognitive psychology, such as memory or problem solving, have also borrowed their names from ordinary language. More often than these other terms, however, the term attention tends to be used not only as a name for a set of phenomena, but also as a theoretical construct for explaining these phenomena. As various writers have noted (e.g., Broadbent 1982), smuggling terms like attention into scientific discourse from ordinary language can cause serious confusion. One way it can do so is by leading people to assume that the prescientific notions that go along with the word must be correct. For this reason, we shall use the term sparingly in this chapter, pending an examination of the empirical evidence.

Of all the phenomena associated with the concept of attention, the most conspicuous are *selectivity* and *capacity limits*, both of which are important aspects of mental life that must be dealt with by any psychology. Our momentary experience of the world depends not just on what stimuli are presented to our sensory apparatus, but on which we choose to become

aware of. Selection is most obvious when it is difficult, as for example when we strain to listen to one voice in a crowd while ignoring other, perhaps louder voices. However, much or all of our experience reflects selection, since there are always more stimuli impinging on our sensory systems than we happen to notice at any moment. The fact that people can generally control which stimuli have the most pronounced effects on them bedeviled early stimulus-response theorists who hoped to find laws linking objectively defined stimuli and responses to them. For the most part modern cognitive scientists take it for granted that unobservable internal structures and processes are important. They thus find attention a source of puzzlement rather than of consternation.

Capacity limitations are another undeniable feature of human mental life. Certain limitations on human perception and performance are dictated by the nature of the stimuli and of the peripheral sensory apparatuses. One example of this constraint is our limited acuity for stimuli presented in the periphery of the visual field. Others follow from the structure of our bodies and motor systems: for example, we have difficulty typing and drinking coffee at the same time. In many cases, however, tasks that we could perform individually are impossible to perform at the same time, quite apart from any sensory or physical limitations. For example, few people can profitably read a novel while listening to a symphony, and only the reckless try to tune their car radio while driving in busy traffic. These sorts of capacity limits are generally termed *attentional limitations*.

The concept of attention is part of what might be called a folk-psychological theory—that is, an informal set of propositions people rely on to explain their own and other people's daily experience and behavior. The folk-psychological theory of attention postulates an internal substance or process (attention) and uses it to explain both selectivity and capacity limits. The informal theory goes something like this: Attention can be devoted to stimuli, to memories, and to activities (and perhaps other things as well). When this happens, less attention (or sometimes no attention) remains available for other experiences or activities. Attention is also thought to be precondition of conscious awareness; if one doesn't attend to a stimulus, one remains oblivious to it. Paying attention is effortful sometimes aversive (hence "paying"), but doing so yields a return: performing just about any sensory, cognitive, or motor activity improves it (except, perhaps, for automatic ones). A task that is given the benefit of full attention is accomplished more quickly and efficiently (except, again, certain automatic activities). Attention is ordinarily under the control of the "self," or the "will," although external stimuli or intruding thoughts can sometimes grab attention despite the best efforts of the will to prevent this from happening.

This intuitive theory is so compelling that one is tempted to think it self-evidently true. William James seems to have that idea in mind with his famous dictum that "everyone knows what attention is" (James 1890). If the dictum is valid, it would seem to make sense for attention researchers to take the basic notion of attention as a given and to move on to working out the details. The earliest psychologists who wrote on attention in the late nineteenth century—set about to explore attention in essentially this way. Their main focus was on how attention affects people's conscious experience of external stimulation. They debated such questions as whether attending to a stimulus increases its subjective intensity or clarity or only changes its salience. The issue of divided attention was also discussed extensively. Here, too, the existence of something called attention was taken as a given, and writers debated whether this substance was ever truly divided. Oswald Kulpe, for example, was skeptical about the idea that people could genuinely divide their attention, while James argued that people could attend to more than one thing at a given instant, but only when they apprehended all of them as part of a single object (James 1890). Others argued that people could attend to as many as five or six different things at once. For the most part, these writers relied on introspective evidence, although a few studied people's ability to count a collection of objects or to name briefly presented stimuli (e.g., Cattell 1885).

Most present-day attention researchers work in the tradition sometimes called *information-processing psychology*. Their avowed goal is not to characterize conscious experience per se, but rather to trace the flow of information among different representational systems in the mind/brain. For the most part, they place little stock in introspection as a means of achieving this goal, relying instead on recording observations of human behavior in laboratory settings. These behaviors often involve subjects' reports of stimulus objects presented to them. The reports are taken as phenomena to be explained—by postulating a sequence of underlying information-processing operations resulting in the overt report. For the most part, however, subjects are not asked to analyze their underlying mental processes (for example, where their attention was directed).

While introspection as a methodology is universally dismissed in present-day cognitive science, the basic notion of conscious experience is still generally accepted. For example, contemporary researchers often distinguish between what observers experience at the moment they look at a stimulus and what they might report about the stimulus immediately thereafter. There is much less agreement, however, about the criteria to use in making this distinction (for examples of debates that have hinged on these criteria, see Treisman and Schmidt 1982; Fagot and Pashler 1995; Cheesman and Merikle 1986). In this regard, information-processing psychologists rarely adhere to the rigorous precepts of psychophysicists, such

as Brindley (1970), who argue that subjective observations can be meaningful only when they show that observers cannot discriminate between physically distinct stimuli. (An example of an effect that does satisfy this standard of rigor is metamerism in color vision, in which blends composed of different wavelengths of light appear to have the same color as light of a pure wavelength). While some of the inferences used in studies of attention could be recast to fit this strict standard, many could not.

Even though modern attention researchers are usually dubious about introspection as a way to gain access to psychological processes, they often seem to assume that the folk-psychological theory of attention is at least partly right. For example, researchers often speak as if any task that is difficult in some way *consumes* attention. For example, when subjects hold digits in memory while doing some other task, this other task is often said to be performed under conditions of *divided attention*. Similarly, it is sometimes suggested that attention can be allocated to mental contents as diverse as positions in visual space or semantic categories in long-term memory. These ways of speaking make sense in terms of the folk-psychological theory of attention, but obviously involve assumptions that might or might not be valid.

2.2 Capacity Limits in Visual Perception

The present work focuses exclusively on one aspect of divided attention people's ability to perceive multiple visual stimuli at the same time. This question can be posed without assuming the intuitive notion of attention described above and, indeed, without even using the term attention. For the reasons described above, this practice has much to recommend it and will be followed—with occasional lapses—in the following several sections. If the intuitive notion of attention is valid, its virtues should emerge when empirical results are considered; in that case it will have proven most instructive to have put the concept to the test.

We begin with the following question: If a person is presented with several visual stimuli and attempts to perceive them all, what information can be derived from them in parallel? If the stimuli belong to familiar categories, for example, can they be categorized simultaneously? Alternatively, do we have to scan the items, recognizing first one and then the next? Or perhaps people can perceive multiple objects simultaneously, but not as quickly or efficiently as when fewer items are present. This idea is usually called *capacity-limited parallel processing* and designates *serial processing* as the most extreme form of capacity limitation. The first several sections of this chapter consider these questions and describe two approaches used to answer them. Later sections examine how capacity limits in perception of multiple objects relate to capacity limits in thought and

action. Here the broad conception of attention described in the preceding paragraphs is considered. The final section explores the relationship of capacity limits to selectivity and control and makes some observations about the intuitive notion of attention in light of the results discussed in earlier sections.

What exactly is meant by *perception* or *recognition*? Attention researchers generally use these terms to refer to deriving an internal representation that makes explicit the fact that an object belongs to some learned category, such as a particular letter, word, or type of object. Thus, a letter is not recognized by the retina just because its identity could be recovered from the pattern of retinal activity, because that representation is not explicit or usable without further processing. (Unfortunately, the term recognition is also used in memory research to refer to the judgment that something has been encountered before; that meaning will not be relevant here.) Visual perception involves a great deal more than just recognition, however. For example, our perceptual apparatus informs us about the color, position, and texture of different surfaces, perhaps the most likely three-dimensional arrangement of an object's parts, the sizes of objects and their parts, and so on. These different representations seem to be computed by at least partly independent brain areas (see Chapters 3 and 5). Partly for historical reasons, though, attention theorists have focused heavily on the recognition of familiar stimuli. Many of the issues considered here could also be raised about other perceptual operations (for a rare example in which this has been done, see Epstein, Babler, and Bowns 1992).

The question of whether visual object recognition is subject to capacity limits sounds straightforward enough. Probably the most obvious strategy would be to have observers name the objects they see in a brief display; the objects might be letters, digits, words, or pictures. Can an observer report the identity of all the elements in the display equally well regardless of how many items are present? If performance breaks down as the number is increased, one might conclude that capacity limits are present and that recognizing some of the objects depletes the capacity available for recognizing others. If it does not break down, one might infer that recognition is parallel.

The problem with this strategy is that reporting what was in the display nor only involves identifying the objects, but also storing the results long enough to report them. Even though the layperson does not usually think of memory as a limiting factor over the time scale of a few seconds, it often is. Sperling's classic studies of the perception of brief displays provide one demonstration of this (Sperling 1960). When he asked subjects to report as many letters as they could from a display, he found they could report only about four or five. This was the case whether the display was

presented for just one millisecond or for as long as one second (a factor of a thousand difference). It is also true whether the display contains six letters or many more. The main cause of this strange invariance is the fact that in order to report the information about what is present, the person must store the information in a limited-capacity memory system that is capable of holding only four or five items; getting even a few items into memory can be problematic.¹

If information must get into short-term memory and be kept there to be reported, then report tasks are a problematic way of looking at recognition per se. To get around these limitations, investigators, beginning in the 1960s, started using visual monitoring or search tasks. In a monitoring task, a subject searches a display for a target and indicates whether or not he or she believes it to be present. (In the forced-choice version of the task, the subject indicates which of several possible targets is present.) Monitoring tasks greatly reduce the amount of information that must be stored in memory (to perhaps one bit). Estes and Taylor found that even with displays so brief that people could only report four or five letters, their detection performance often showed they must have processed more than four or five letters. The term *processed* here refers to whatever operations are used to discriminate targets from nontargets (Estes and Taylor 1964). A compelling illustration of the same point is provided by Molly Potter's studies involving rapid, serial visual presentation (RSVP) of pictures (Potter 1976). When her observers tried to determine whether a target picture (e.g., a girl sitting in a bed) was present in a sequence of sixteen pictures exposed at a rate of 167 milliseconds (msec) per picture, they detected over 70 percent of them. On the other hand, when the pictures were presented first and *then* subjects were told what targets to report, their performance was much worse. It seems, then, that an observer can rapidly analyze a sequence of pictures when the purpose is to find something, even when the pictures are exposed so briefly that he or she cannot store more than a tiny fraction of what is seen. The subjective experience of looking at RSVP displays is consistent with this. At moderately fast presentation rates, people often report feeling that they comprehend each picture but somehow lose this information the moment the next picture is presented. Indeed, the quick-cut exposure of scenes favored by movie and television directors, would be hard to explain if it produced only bewilderment in a viewer.

The fact that a person's visual system can process a great deal of information in a detection task—far beyond what the person can report—

1. The interpretation of these classic studies has often been questioned and sometimes misunderstood; for a recent discussion of the nature of the short-term memory system used in Sperling's whole-report task, see Pashler, in press).

still does not tell us whether object recognition operates in parallel and whether it is subject to capacity limits. It does, however, suggest a possible way of answering that question: measuring how the number of objects in a display affects observers' accuracy of detection. Estes examined this issue with displays of letters and found that accuracy declined as the number of characters in the display (*display set size*) increased (Estes and Taylor 1964). This might sound like proof of capacity limits, but there is a further methodological hurdle. To understand it, we consider a concept psychologists refer to as the *ideal observer*. This ideal observer is a hypothetical observer subject to no capacity limits, that is, capable of processing any given item equally well, regardless of how many total items are presented. The performance of this ideal observer can be formalized as follows. Let the probability that the ideal observer mistakes a given nontarget for a target be F , and let the probability that the observer mistakes a target for a nontarget be M . Both M and F are independent of display set size N . (While the observer is ideal in the sense that it is not subject to capacity limits, it is not flawless or omniscient, as reflected in the fact that both F and M are greater than zero.)

What happens to the ideal observer's accuracy when no target is present and N is varied? One might think it would remain constant, but in fact the ideal observer becomes more and more likely to report a target falsely as N increases. The exact amount of increase depends on exactly how the ideal observer makes decisions. The simplest and most discrete decision process would consist of reporting the presence of a target if *any* channel appears to contain one. Under these assumptions, the probability of a false alarm is $1 - (1 - F)^N$. (As N grows, $(1 - F)^N$ shrinks, and $1 - (1 - F)^N$ grows.) If the decision process works in some other way (it need not be discrete, for example), the false alarm rate would follow a different equation but would still increase with N (unless some rather exotic conditions hold²). This statistically based increase in error rates produced by increases in N will be termed *statistical decision noise*.

The full implications of statistical decision noise are often neglected. For example, some of the best-known studies of visual capacity limitations have plotted increases in response times (RTs) to detect correctly various kinds of targets, as a function of N (e.g., Treisman and Gelade, 1980). The logic of these studies is intuitively very natural: if every additional item in the display increases RTs by the same amount, then one presumes that people are scanning the display one item at a time. Error rates are disregarded in this formulation, which seems reasonable since the actual error

2. The exotic condition is a so-called high-threshold model according to which the probability of mistaking a nontarget for a target is always zero, although the probability of detecting a target may be below 1.0. Since people generally *do* make false alarms, this model would have to suppose that some pure guessing occurs.

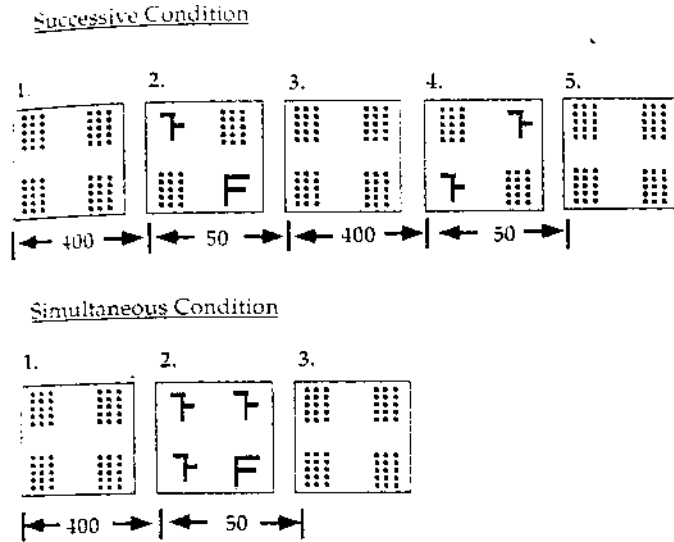
rates in these kinds of studies are very low and differences among conditions are often insignificant. It would seem, therefore, that the issue of statistical decision noise can safely be ignored in this kind of experiment. This is not the case, however. Subjects are capable of trading speed for accuracy in essentially any task—taking a little more time and consequently making fewer errors. The existence of speed-accuracy trade-offs raises some troubling possibilities. For example, when subjects are confronted with a larger display, they may compensate for statistical decision noise by taking more time to process them. If subjects actually achieve the same error rates at different display set sizes, they must be doing something of this kind. However, the typical RT experiment provides little information either way about how display set size affects error rates, because accuracy is generally very close to ceiling (Wickelgren 1977). It would require careful analysis of the joint pattern of RTs and error rates—probably combined with an unusual degree of speed stress—to unravel these issues. Given these factors, one needs to be cautious in interpreting experiments that simply look at the effects of display set size on RTs. When RTs are unaffected by display set size it is a pretty good indication of parallel processing. Enormous slopes make sequential processing seem very plausible, but the technique suffers from serious logical limitations.

2.3 A Better Method: Comparing Simultaneous and Successive

Displays

The preceding section makes it clear that diagnosing capacity limits is not quite as easy as it sounds. Fortunately, however, several strategies can provide more leverage on the problem than one gets by simply looking at error rates or RTs as a function of display set size. The first method was developed by Eriksen and Spencer (1969) and Shiffrin and Gardner (1972). Here the display set size is held constant (thereby keeping statistical decision noise constant), and the timing of the presentations is varied (see Figure 2.1). Two conditions are compared. In the simultaneous condition, all items are presented at the same time and followed by masks. In the successive condition, the N items are presented k at a time. Each item is always followed by a mask after a fixed delay, which is the same in the two conditions. If there are capacity limits, the successive condition is bound to be easier; here capacity can be shared among fewer items at any given moment. If processing is sequential—the most extreme form of capacity limit—the cost should be very substantial.

The method requires several assumptions. One is that in the successive condition, information extracted from the first display is not forgotten by the time a response is chosen. If it were, accuracy would be worse when

**Figure 2.1**

Method used in Shiffrin and Gardner (1972) experiment comparing simultaneous and successive exposures of four characters. The subject's task is to decide if a T or an F is present. In both conditions, stimuli are presented for T msec and then masked; however, in the successive condition only two stimuli are shown at any given time, whereas in the simultaneous condition all four are present simultaneously.

the target occurs in the first display rather than in the second. This outcome has not been observed (e.g., Shiffrin and Gardner 1972). Second, the method assumes that masks terminate perceptual processing fairly abruptly; if observers have hundreds of milliseconds beyond the onset of the masks to continue working on the display, the simultaneous condition would be effectively equivalent to the successive condition. Various results, including those described below, provide some justification for this assumption.

The simultaneous versus successive method has been used extensively with search tasks involving alphanumeric characters. The general finding is that detection performance is identical, or close to identical, in the two conditions. This implies that these displays can be processed with virtually no capacity limits (Shiffrin and Gardner 1972; Duncan 1930b). This finding holds even when the task involves finding the highest digit in a display of digits (Pashler and Badgio 1987). The highest-digit task is of interest because, while people might detect a single prespecified target character without actually identifying distractors, it seems unlikely they could find the highest digit without computing the identities of all the digits, since

any unidentified digit might turn out to be higher than the apparent highest.³

People are not always as accurate in detecting targets in simultaneous displays as they are with successive displays, however. The simultaneous condition falls behind when subjects are required to make more complex and numerous visual discriminations. For example, when the displays are composed of two words, the successive condition enjoys a very large advantage over the simultaneous condition. This holds true whether the task involves detecting a particular word (Duncan 1987) or detecting an unspecified member of a semantic category (Shiu and Pashler, in preparation). The fact that an advantage for successive presentations emerges when more complex discriminations are required is intuitively reasonable. It implies, however, that perceptual capacity limits are fundamentally different from verbal short-term memory (STM) capacity limits. STM seems able to hold a certain number of familiar items; whether they are words or letters doesn't make much difference.

Another condition in which successive displays enjoy an advantage over simultaneous displays is when two targets are presented and must be detected separately. Duncan carried out a series of studies involving detection of variable numbers of targets in brief displays (Duncan 1980b). In a typical experiment, subjects monitored displays of four masked letters rather like those shown in Figure 2.1. In one condition, though, subjects made a separate detection response for each of the diagonals. Since subjects responded to information in each diagonal separately, each diagonal can be referred to as a separate information *channel*. Doing this makes it easier to discuss experiments of this sort, so we will use the term henceforth. When a single target is present in one of the channels and no other target is present, subjects are almost as accurate in the simultaneous condition as in the successive condition. When two targets are present, however, one in each channel, performance is significantly worse for simultaneous presentations. The probability of detecting a target in one channel (say, channel A) depended on which happened in the other (channel B). If the subject successfully detected a target on channel B (a "hit") or falsely reported there to be a target there ("false alarm"), there was a large decrement in detection on channel A.

It turns out that detecting a target in any brief presentation on one channel generally reduces the ability to detect a target on another channel. What makes this tendency especially interesting is that it happens in the

3. Of course, one could postulate a search process that searches first for the highest potential target, then for the next highest, and so on. Given the ease and rapidity with which people perform this task upon first attempting it, this algorithm seems a little implausible.

very tasks in which people can monitor two channels simultaneously without loss (that is, when there is only one target present). Duncan showed that the effect occurs even for detection of relatively simple visual features (Duncan 1985).

Although the focus in this volume is on visual perception, it is worth noting that the results just described generalize very nicely to auditory detection performance. A representative study was conducted (Puleo and Pastore 1978) in which subjects listened for a target (defined by frequency) in one ear and a target of a different frequency in the other ear. The subject knew which ear might contain a target at which frequency. In the monaural listening condition, targets were presented only in the ear the subject monitored. In the selective attention condition, targets could occur in either ear, but the subject was instructed to detect targets in only one ear. In the divided attention condition, targets could occur in either ear, and subjects had to report on each ear separately.

When target frequencies in the two ears were similar, simultaneous targets were perceptually fused and listeners heard a single tone. (Not surprisingly, this made it hard to detect both targets.) Targets widely separated in frequency did not fuse, however. Here, detection was as good in the selective attention condition as in the monaural condition, implying that subjects could prevent target detection if they chose. In the divided attention condition, however, performance depended on what was played in the opposite ear. When no target was played, detection was comparable in selective and monaural conditions. When the other ear received a target, however, performance was substantially worse. Thus, Puleo and Pastore's results closely parallel Duncan's later findings with visual detection. Like the visual effect, the auditory double-detection decrement cannot be due to sensory interactions; a contralateral target interferes even when it is a *gap* rather than a tone (e.g., Gilliom and Mills 1974, cited by Pohlmann and Sorkin 1976). The auditory double-detection problem is also robust. Neville Moray, for example, had people monitor continuous inputs (rather than discrete trials, as in the Puleo and Pastore studies), listening for targets ranging in complexity from tones to spoken words (Ostry, Moray, and Marks 1976). Again dual-channel monitoring was about as good as single-channel monitoring unless a target was present on the other channel, in which case a large decrement occurred (Moray 1975).

We see, then, that people can often monitor several visual or auditory channels for a target in parallel without showing any capacity limitations, however, detecting a target on one channel regularly impairs detection on another channel. Unfortunately, no one seems to have asked the obvious question of whether detecting an auditory target impairs detection of a visual target, and vice versa, so it is not clear whether the underlying limitation is restricted to a given sensory modality. There is one final

intriguing parallel between the auditory and visual double-detection decrements. The effect is essentially abolished when the two targets are attributes of the same "object"—for example, orientation and texture of a line (Duncan 1984) or the loudness and timbre of a tone (Moore and Massaro 1973). The reasons for these close parallels between visual and auditory attention effects are not well understood. However, it will be suggested below that they are more likely to reflect similarly organized but separate attentional processes in different modalities than a single common attentional system serving both visual and auditory inputs.

2.4 A Converging Approach Using Reaction Times

As described earlier, many detection studies that require people to respond quickly about the presence or absence of a target analyze the reaction times in order to characterize perceptual capacity limitations. When the task is to detect a visual target that differs markedly from the background elements in a simple feature like color or orientation, RTs are barely affected by the number of (dissimilar) distractors (Green and Anderson 1956; Treisman and Gelade 1980). This essentially rules out the possibility that featural discriminations are carried out sequentially across the whole display. The parallel search may be carried out using mechanisms that normally function to segment the scene by locating discontinuities in visual features (see Chapters 2, 4, and 5).

When observers must detect a letter or digit target, on the other hand, results are more equivocal. Every additional distractor typically adds 20 or 30 msec to the total RT on negative trials. In some cases the slopes for positive (target-present) trials are about half those for negative trials, while in other cases the ratio is closer to 1:1. A 1:2 ratio for positive:negative slopes suggests a serial, self-terminating search, that is, a search that stops when a target has been found. This is because a self-terminating search will, on average, stop halfway through the array. A serial, self-terminating search might involve scanning clumps composed of more than one item at a time rather than just individuals, which results in a 1:2 slope ratio (Pashler 1987).

The more general question, however, is whether moderate positive slopes imply that character recognition is being carried out sequentially, or at least is subject to capacity limitations. Logically speaking, the slopes are consistent with that possibility; they are also consistent with several alternatives. Perceptual processing might, for example, be parallel, and other, later parts of the task might take longer when display set sizes increase. Even if all the items in the display are analyzed in parallel, a comparison or decision process might still be necessary to generate a response, which

might take longer for displays composed of more items. A second alternative, already noted above, is that while perceptual analysis might operate in parallel subjects may try to equalize their accuracy over different display set sizes to overcome the effects of statistical decision noise. This might be done by taking more time for larger display set sizes than for smaller ones.

Can one use RTs to sort out whether perceptual processing is serial or parallel? One fairly promising technique for doing so involves manipulating not just display set size, but also the difficulty (and hence the duration required) of perceptual analysis of each item in the display. There are several ways of prolonging perceptual analysis—for example, by reducing the intensity of the characters or degrading them with dots. In the

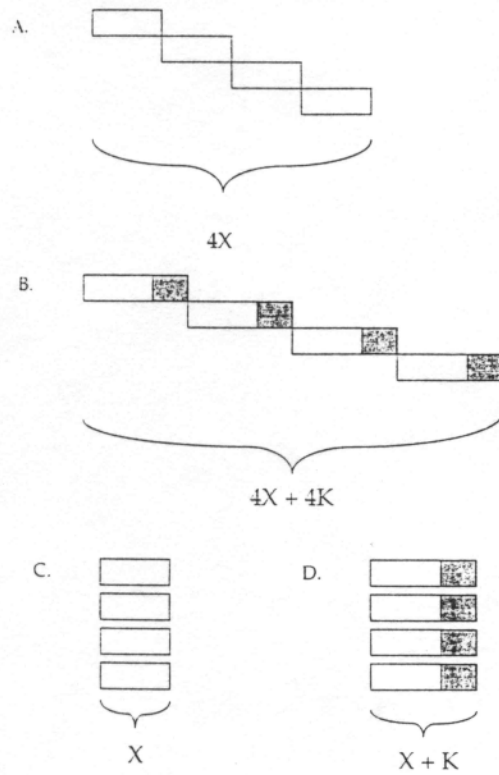


Figure 2.2
 Predicted effects of adding extra time (shaded portion, lasting k msec) to duration of perceptual analysis. (A and B) Serial Perception Model; (C and D) Parallel Perception Model. (B) and (D) represent cases in which the stimuli are degraded (adding the extra time indicated by shading). In (B) Serial Model, RTs are increased by $4 \cdot k$, whereas in (D) Parallel Model, they are only increased by k . Thus, the models make very discrepant predictions.

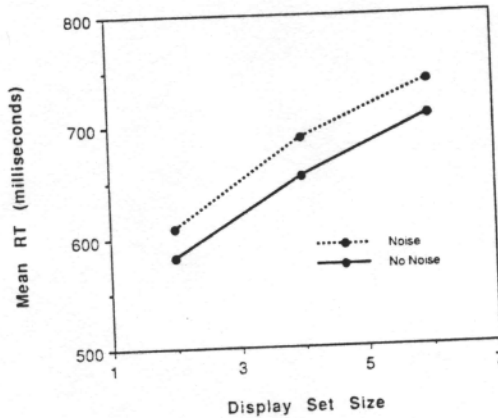


Figure 2.3

Data from experiment in which subjects name the highest digit in an array of two, four, or six digits, either all degraded or all undegraded. Degradation produces a constant increase (as predicted by the model shown in Figure 2.2, Panels C and D). (Data reproduced by permission from H. Pashler and P. C. Badgio, Visual attention and stimulus identification, 1985, *Journal of Experimental Psychology: Human Perception and Performance* 11, 105—121.)

experiments described below, the entire display was either undegraded (e.g., all letters very clear) or degraded (e.g., tiny dots superimposed on all the letters). The key test is how degradation effects on RTs vary according to display set size. As shown in Figure 2.2, parallel and serial hypotheses make completely different predictions. If perceptual analysis is sequential, slowing perception by X msec increases RTs by an amount equal to the product of X and the display set size. Thus, if, for example, degrading the display adds 20 msec to the time it takes to process each element in the display, degrading a display of four items will add a total of 80 msec to the RT, while degrading a display of eight will add 160 msec. If, on the other hand, perceptual analysis operates in parallel, the 20 msec slowing should be added into the RT regardless of display set size.⁴ Thus, the hypothesis of sequential perceptual analysis predicts that effects of display set size and visual degradation will combine in a multiplicative manner, while the parallel hypothesis predicts that they will combine in an additive fashion.

In fact, the results show the additive pattern predicted by the parallel hypothesis (Pashler and Badgio 1985). Figure 2.3 shows the results of a speeded-up task in which—rather than detecting a prespecified letter—

4. Strictly speaking, a slight (and perhaps undetectable) interaction is to be expected whenever a manipulation increases the *variance* of processing times (see Pashler and Badgio 1985).

subjects named the highest digit in a display of digits. Here too the effects combined additively.

In summary, while reaction times often increase somewhat with display size in speeded-up detection tasks involving moderate numbers of alphanumeric characters, these slopes probably represent the effects of display set size on postperceptual processes, rather than sequential perceptual analysis. The results are therefore compatible with the conclusions described in Section 2.3: moderate numbers of fairly simple objects like letters and digits can be identified in parallel. We argued in that section that recognition of more complex objects like words is probably constrained to operating sequentially (or at least is severely capacity limited). The second method described above—analyzing display set size and visual degradation effects on search RTs—has not been tried with stimuli as complex as words. Doing so would provide a useful converging test.

2.5 Interpreting Perceptual Capacity Limits

So far we have focused on the question of how much visual analysis a person can carry out in parallel while trying to perceive several inputs concurrently. Since people can process a great deal more than they can ever report, visual search turns out to be more useful for this purpose than tasks requiring overt naming of stimuli. The problem of decision noise makes it trickier than one might have expected, however, to determine whether processing is capacity limited or not. This is because accuracy falls when there are more stimuli to be processed, even if perceptual analysis is not subject to any sort of capacity limitations. We described two methods that (at least partially) overcome this problem. One involves comparing simultaneous and successive displays, and the other involves looking at the joint effects of display set size and visual degradation on RTs. The results of both methods suggest that people can identify a number of characters at the same time without capacity limits. The accuracy experiments further suggest that this is not true of more complicated stimuli such as words; here the analysis is clearly subject to capacity limitations and may operate sequentially. We turn now to the broader implications of these conclusions.

The conclusion that parallel analysis of characters is possible in the context of a search task raises several questions. The virtue of search tasks is that they get around the need to hold more than a tiny bit of information in memory. However, they also have a disadvantage: when people search for a target of some kind, they may be performing perceptual analysis that is specialized for search. Each item in the display may not be identified but merely processed in a such a way as to figure out whether it

is a target or not. This is plausible but not compelling. With a high degree of practice, subjects do seem to construct specialized detection skills limited in their generalizability to other stimulus sets (Shiffrin, Dumais, and Schneider 1981). On the other hand, our perceptual systems do not have unlimited flexibility; even when one might expect that top-down information would be used to tune perceptual machinery, this does not always happen (Johnston and Hale 1984). There is reason to suspect that visual search may depend on genuine object identification. For one thing, the highest-digit detection task produced results quite similar to those found with ordinary search tasks, as noted above. Furthermore, people often do a good job searching rapid sequential displays for such unknown members of a target semantic category as animal names (or, in Potter's RSVP experiments, pictures of animals).

It would seem reasonable, therefore, to conclude tentatively that search does reflect parallel object recognition, and that object recognition is subject to moderate capacity limits that are exceeded by a few complicated objects but not by a few simple ones. There is much more to be done to figure out the generalizability of these conclusions, however. For example, are there capacity limits in recognizing nonlinguistic stimuli like faces, tables, and chairs? Is recognizing an object like a table or a face comparable to recognizing a letter or a word in terms of its capacity demands? As the technology for presenting brief images of scenes becomes widely accessible, we will undoubtedly learn much more about the capacity limits affecting a broader range of visual tasks and stimuli.

A much deeper question is what causes the perceptual capacity limits that have been documented. At the present time little is known about this issue. Some contemporary neural-net models of visual processing predict a gradual breakdown of processing as visual load increases (Mozer 1991). These models generally predict display-set-size effects that are highly dependent on similarity (worse performance with more similar objects). The evidence for this is mixed. The most extreme manipulation of similarity is modality. Treisman and Davies showed that γ -perceptual analysis in two different modalities produces less mutual interference than processing within a modality (Treisman and Davies 1973). On the other hand, Duncan recently found no evidence that the similarity of two classification tasks within the visual modality affected the degree of mutual interference (Duncan 1993). Clearly, it will take further research to unravel the causes of the interference. From a functional, psychological perspective, we may have to be content with a description of how and when performance breaks down and the ways in which output of perceptual analyses become available for different purposes. Answers to the deeper questions of *why* these processing limitations exist may only become possible when the neural circuitry is better understood.

Another intriguing issue is what kind of internal representations people form when they look at a scene composed of multiple objects. From the studies described earlier we concluded that when people choose to do so (as in a search task), they can identify multiple objects at the same time—subject to the capacity limits described above. But beyond allowing a person to locate the search target, what purposes does this identification process serve? Does it produce a mental representation of the identities of the objects in the scene? If so, can the identity information be accessed at will? Is the knowledge of the identity of objects tied to a spatial representation of the scene as a whole, like a map in a military command center with labels pinned to it? Since we seem to have the experience of seeing a whole scene at a glance, it might seem that there must be such a representation.

There are two kinds of studies that indicate, however, that this may not be the case. The first set of studies investigates people's ability to detect changes in a scene they have been looking at for a period of time. Consider what happens when subjects look at a display for a second or two, then it disappears for a short interstimulus interval (ISI) and then reappears. If asked to say whether anything in the display has been altered, can they do so accurately? George McConkie and his coworkers explored this issue in situations where the observer makes a saccadic eye movement during the ISI (McConkie 1990). He finds that people are remarkably insensitive to changes in either text or natural scenes, except for a small portion they choose to monitor, presumably using visual short-term memory (see also Irwin 1991). This finding has a remarkable implication: our subjective sense that we see the world as stable when we move our eyes may be illusory, because our brains do not actually check out the stability (a possibility first raised by Mackay, 1962).

Of course, it is possible that the changes are missed simply because of the eye movement: a maplike representation of the scene is built up during the initial fixation, but when the eye moves the representation is flushed and a new one is created when the eye comes to rest. To test this proposition, one can ask what happens when the display simply flickers off and on, with no eye movement being made. If the ISI is very brief—say, a few milliseconds—observers see flicker or motion where the change is introduced and have no trouble using that as a cue to detect the change (e.g., Phillips 1974; 1933). When the ISI is lengthened to, say, 100 milliseconds—still a very brief offset—an observer sees flickering in all parts of the scene. Because the flicker is everywhere, it provides no clue to where changes might have taken place. Does an internal maplike representation of what was in the display allow one to perform well? It seems not: without the flicker as a cue, observers are very poor at detecting changes. This result has been observed with displays composed of unfamiliar

matrices of light and dark squares (Phillips 1974) as well as with displays of objects like letters and digits (Pashler 1988). More recently, Ling-Po Shiu and I have begun looking at change-detection by using photographs of real-world scenes. Here, too, observers often miss gross changes (e.g., the removal of a sailboat from a harbor front scene, or a change in the color of the pants worn by a person standing in the foreground). It seems, then, that while attending to the whole scene gives us the impression that we "take it all in at once," this may be in some sense illusory; at least, we don't form a rich internal representation of what objects are present in the scene and where each one is located. Or, if we do form such a representation, we don't seem able to use it to detect changes. Is it possible that people actually do form such a representation but lose it as soon as the display disappears? Or might they simply lack the ability to use their scene representation to detect change? These ideas sound reasonable enough, but they raise metaphysically troubling questions such as whether it makes any sense to say that someone knows that an object is in a certain position if the person finds nothing amiss when when it is removed.

It may be possible to get around some of these metaphysical questions by means of another method of testing, one that is similar in some ways to the RT experiments described above. An observer looks at a scene for a short while ("preview"); then a visual probe is presented next to one of the objects. The subject's job is to name the probed object, responding as quickly as possible. The task is similar to Sperling's (1960) classic partial report experiments, except that the display remains present for a short time after the probe appears, and the response is speeded up. If people really construct a mental representation of what objects are present in the scene and where each one is located, the subject should be able to retrieve the identity of the probed element from memory when the probe appears (after preview). As they should not have to identify the element "from scratch," RTs should not depend on the difficulty of identifying this item, since any slowing on that account will have been "absorbed" by the preview time. Experiments using displays of letters and various kinds of degradation did not, however, find any such absorption (Pashler 1984). (The conditions of this study are shown in Figure 2.4 and the results in Figure 2.5.) Apparently, then, subjects do have to start from scratch when the probe appears.

In summary, although we seem to be able to analyze more than one object at a time and recover their identities in parallel, we don't seem to be able to store the outputs of these analyses and access them for whatever purposes we might have (such as retrieving an identity when given its location). There are various possible interpretations of this finding. One is that we simply lack any way to hold onto the results beyond the moment they are retrieved. This could also explain the failures of change detection.

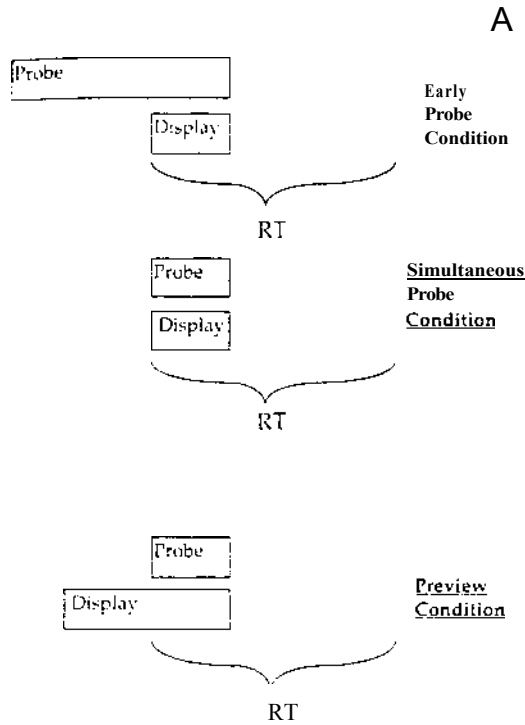


Figure 2.4
 Timing of the three conditions (early probe, simultaneous probe and preview) in the probe-and-preview experiment described in text. In all cases, subject names the probed letters as rapidly as possible. RTs are recorded from the first instant that both probe and display are present on the screen as the subject has, potentially, the information to respond from this time forward.

However, a mere memory limitation does not explain the RT results, since subjects are looking at the display when the probe appears. Another interpretation would say that location information is lost when multiple objects are analyzed in parallel. This possibility fits in nicely with certain computational models of visual object recognition (e.g., Mozer 1991). There are, however, psychophysical observations that seem to conflict with it. When people detect the presence of even a simple object in a search task they seem to be able to report accurately the location at which the detection occurred (Green 1991; Johnston and Pashler 1991). This suggests yet another formulation: our knowledge of identities and locations might be more like a set of one-way pointers from identities to locations, rather than

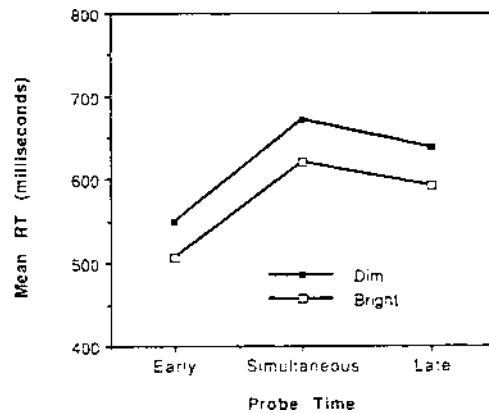


Figure 2.5

Typical results from experiments of the kind depicted in Figure 2.4. Reducing the intensity of the characters slows responses even when the probe is late (giving the subject the opportunity to preview the entire display). These data are from an unpublished experiment conducted by the author in 1991 and basically replicated in Pashler (1984), Experiment 3.

a set of bidirectional linkages (like a map). Thus, given the identity of an object, a person might be able to recover the location of the object in that location but, given the location, might not be able to recover the identity. Getting that information would require starting "from scratch" and re-processing the item alone. Another possibility is that the same neural machinery is used to analyze the whole scene and to process just a single element in it, but at a higher resolution.

Obviously, more experimentation is needed before one can assess the merits of these proposals. One conclusion seems warranted, however: visual awareness of a scene may reflect more impoverished knowledge structures than one might suspect on the basis of casual introspection. Experimental results confirm that people can take in a lot of information in parallel, but what they can do with that information seems to be limited in ways that are counterintuitive. Of course, our intuitions are based mostly on metaphors involving the objects in the physical world (e.g., maps) and recently, digital computers. Given that the brain is constructed so differently from these things, it should probably not surprise us to learn that the fate of visual information does not conform to our initial guesses.

2.6 Perceptual versus Central Processing Limitations

So far we have discussed perceptual aspects of divided attention; that is, trying to take in many different stimuli at once and figure out what they

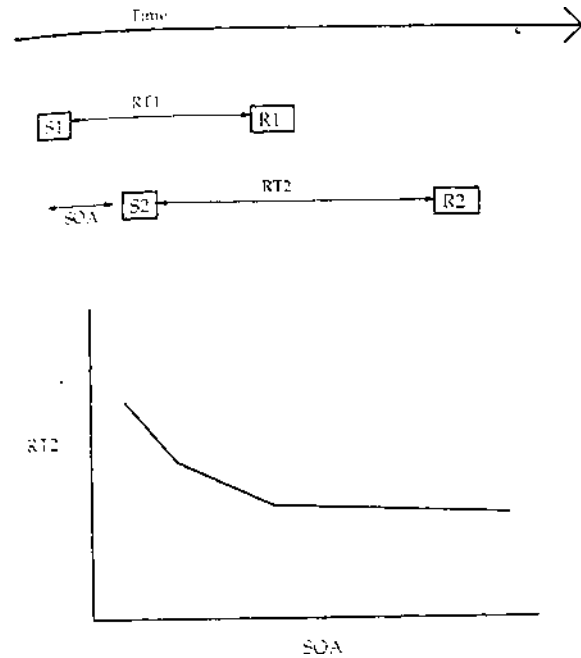


Figure 2.6
The Psychological Refractory Period (PRP) effect. Two stimuli are presented (S1 and S2) and subject makes a rapid response to each. The responses to the second stimulus are lengthened as the interval between the stimuli (stimulus onset asynchrony or SOA) becomes shorter.

are. People also have limitations in carrying out more central kinds of cognitive operations, such as thinking, deciding, and acting. Clear and striking evidence of these limitations comes from a very simple experimental situation. If a person must make a speeded-up response to each of two stimuli presented close together in time, the response to the second stimulus is almost invariably slowed. (Occasionally the response to the first is slowed as well.) The closer the stimuli in time, the greater the interference (Figure 2.6). The slowing of the second RT is usually referred to as the psychological refractory period (or PRP) effect. The PRP effect arises in a very wide range of circumstances: for example, when the two stimuli are in different modalities (Davis 1959), or when responses use diverse modalities like eye movements (Pashler, Carrier, and Hoffman 1993) or even foot movements (Osman and Moore 1993).

The chief cause of the PRP effect seems to be a central bottleneck that prevents people from selecting two responses at the same time (Welford 1952; Pashler 1993). The fundamental limitation (depicted in Figure 2.7)

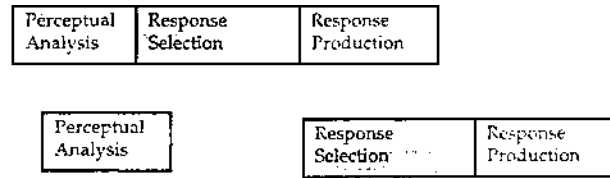


Figure 2.7

The response selection bottleneck account of the PRP effect. The selection of the response to the second stimulus waits for the selection of the response to the first stimulus to be completed, while other stages operate in parallel with each other. This represents the case in which perceptual interference is absent (e.g., where S1 is a tone and S2 is a letter); if both S1 and S2 are visual, perceptual interference can occur.

is cognitive, not "motoric." Deciding what to do in Task 2, rather than actually producing the response, is delayed by the corresponding stages in Task 1. What is the evidence for this claim? One kind of evidence comes from experiments manipulating the duration of the response production on the first task; this slows the response to Task 1 but barely delays the second at all (Pashler and Christian, submitted). Another key piece of evidence is that slowing the cognitive stages of Task 2 delays the second response by a constant amount, regardless of the interval between the stimuli (McCann and Johnston 1992). If the PRP effect were caused by a bottleneck in the producing responses, cognitive effects would decrease as the interval between tasks is shortened. However, this is not to say that there are no limits on response production; there does, in fact, seem to be an additional limitation preventing people from making two manual responses at the same time (even if one involves the left hand and the other the right hand). In practice, this probably makes no difference in a typical PRP experiment in which subjects make a single button-push response to each task because the cognitive bottleneck insures that the two responses would not overlap anyway (Pashler 1994b).

The folk-psychological theory of attention described at the beginning of this chapter attributes problems in dual-task performance to limited availability of attention. Indeed, subjects in PRP experiments often say their responses were slowed because they had trouble paying attention to the two tasks at once. Does the bottleneck that results in the PRP effect have the same underlying cause as the capacity limitations in visual detection tasks? Several lines of evidence argue that it does not. Recall that when people must perform two simultaneous and difficult visual discriminations (e.g., reading two words), accuracy is impaired (section 2.3). The degree of impairment is the same, whether the person must respond to each stimulus immediately or at his or her leisure (Pashler 1989). Thus, response selection conflicts cannot be the fundamental cause of this interference, as they

are of the PRP effect. Similarly, recall that people have difficulty detecting two simultaneous targets, as noted by Duncan (see section 2.3). This occurs even when the person only needs to make a single response to indicate the presence of both targets (for example, counting them; Duncan 1980b). By contrast the PRP effect depends on being forced to choose two actions. The dual-detection problem is avoided, however, when the subject reports on two attributes of the same object (e.g., the orientation and texture of a briefly presented line; Duncan 1984). By contrast, the PRP effect is the same whether both responses are to different attributes of the same object or to different attributes of different objects (Fagot and Pashler 1995). For example, Fagot and I found the usual PRP effect when subjects both named a character and made a button-push response to its color. Finally, as noted earlier, the PRP effect arises whether stimuli are presented in the same sensory modality or in different modalities (for example, a light and a tone). By contrast, perceptual capacity limits seem to be much less severe when stimuli are presented in different modalities (Treisman and Davies 1973).

All of these results argue that the central attentional limits found in the PRP effect have a different source than the perceptual capacity limits documented in detection tasks. There seem to be two sources of interference at work. First, there is a perceptual limitation that is most acute when we try to perceive multiple objects in the same sensory modality; this limitation is indifferent to the difficulty of (or even the need for) response choice and production. Second, there is a postperceptual bottleneck: when a person retrieves a response to one stimulus (or engages in memory retrieval; Carrier and Pashler, *in press*), he or she cannot retrieve anything else at the same time. Either or both of these conditions might arise in any given dual-task situation.⁵ The final section explores these issues in a broader context, returning to the commonsense notion of attention with which we began.

2.7 Conclusions

The results described above distinguish between several different sorts of limits arising in divided-attention experiments. What overall picture of human information-processing limitations can account for these diverse effects? Clearly, supposing that there is a single resource or mechanism called attention that gets divided up in these different contexts will not get us very far. On the other hand, we can identify some grains of truth in the

5. These are not the only factors; others include the manual-response limitation described earlier and a general problem in attaining and maintaining preparation for any task (Pashler, *in press*).

intuitive conception of attention. One is the linkage between capacity limits and selectivity that is implicit in the ordinary notion of attending. Capacity limits do indeed arise from processing that is—for the most part—subject to voluntary control. This applies to both the central bottleneck and the perceptual capacity limits. For example, there is no PRP effect when a secondary stimulus is presented but the person chooses to ignore it (Pashler and Johnston 1989).⁶ In the perceptual domain, when one ignores a portion of a visual display the contents do not tax perceptual capacity limits (Duncan 1979). Thus, selective control and capacity limits go hand in hand with regard to both central and perceptual capacity limits.

A second grain of truth in the ordinary notion of attention is that selecting a stimulus for one purpose often makes information derived from it available for other purposes. For example, when one selects an object in order to respond to its color, information about its form can be stored in short-term memory with minimal interference, while it is difficult to store the form of a different object at the same time (Pashler 1994a). What is selected for one purpose is thereby made available for other purposes (and what is not selected for one purpose is thereby unavailable for other purposes). Were it otherwise, the notion of attending might be useless indeed.

If the folk-psychological notion of attention is not supported, however, is there some alternative conception that fits with the results described earlier? At this point, the most that can be ventured is a very tentative sketch drawn with broad strokes. The following paragraphs offer a brief version of one such sketch, which assumes that selective control works by excluding rather than enhancing particular inputs. There is some evidence for this assumption, but this is not the place to describe it.

Associated with each sensory modality is a set of filtering mechanisms that shut out unwanted stimuli. In the case of vision, information is excluded on the basis of location—although there must be some interaction between the exclusion of a location and the segregation of figure and ground, since people can attend to one of several spatially overlapping objects (Neisser and Becklen 1975). This filtering process operates "early"—that is, prior to object recognition but probably after some initial featural analysis necessary for figure/ground segregation. Stimuli that are allowed passage by this mechanism may activate multiple codes corresponding to various levels of description of the stimulus. This activation may also store the results of all analyses of the object in visual short-term memory, where they are at least partially insulated from being masked by subsequent visual input. The machinery that carries out these analyses can

6. This statement applies when Task 2 is a choice task, but not when it is a simple RT task (see Pashler, in press).

be overloaded: if too many stimuli are admitted, accuracy suffers. In one sense, then, stimuli that pass through the filtering mechanism could be said to compete for perceptual "capacity," while those blocked do not. When an object activates a description that matches with a central cognitive activity (e.g., if it is a target in a search task, for example), the filtering mechanisms constrict around the location of the object. Obviously this notion of matching is very vague; there is little data on how central processes determine detection responses. We do not know, for example, whether a person can think about a sheep *without* searching for sheep. In any case, this constriction makes it difficult to detect two simultaneous targets, even when the discrimination between targets and distractors is so easy that enough capacity is available to search the entire display without capacity limits (Duncan 1985).

Central access to the outputs of these analyses seems far less complete than one might expect, even given the existence of capacity and acuity limits. Although spatial location may be implicitly coded for certain purposes, our conscious awareness of a scene does not seem to generate a maplike representation of the identities of objects in the scene and their locations. At least, if it does, we do not seem able to use the representation to see if anything is changed in a scene, nor can we look up the identity of the object present in a given location.

The perceptual limitations described in these few paragraphs are, of course, only half the story of divided attention. In addition, there is a central bottleneck that arises quite separately from these perceptual capacity limits. This bottleneck is seen most clearly in the PRP effect. Whenever an action is selected or information is retrieved from memory, other operations of the same type are delayed. These delays are mysterious from a computational standpoint, since looking up a few responses seems like a trivial operation, compared to the remarkable computational feats achieved by our perceptual and motor-control system.⁷

What, then, about attention—the concept we banished from the discussion back in section 2.1? Can attention be equated with any of the mechanisms proposed in the preceding paragraphs? One might refer to the machinery that "gates" stimuli as attention and say that an object is attended to when it is allowed access to perceptual machinery. This fits in reasonably well with ordinary usage, although it provides no justification for some common ways of talking. For example, if attention is identified with perceptual gating, it is hard to justify speaking of holding onto a

7. Alan Allport (1989), argues that the key problem of divided attention is preventing incompatible actions, but central interference seems to arise even when tasks involve decision making without any overt responses. The key problem seems to be memory retrieval, not action per se.

memory load while doing some other task as a divided-attention manipulation. Nor should one speak of "attending to a semantic category," "attending to a color," or many other common locutions. Since these are such common ways of talking, it may be best not to use the term *attention* for any of the mechanisms revealed by studies of performance or physiology; instead we might simply regard attention as a pretheoretical concept that is being replaced by empirically based concepts.

In summary, it seems increasingly clear—and perhaps hardly surprising—that the limitations we describe as involving attention in fact reflect the operation of a number of distinct control mechanisms and performance limitations. The ordinary conception of attention, taken for granted by many psychologists as well as laypersons, contains some elements of truth but seems to underestimate the functional separation between perceptual and cognitive limitations. The fine-grained study of performance in divided attention tasks—the examination of concurrent perceptual, cognitive, and motor activities—is a relatively new enterprise. Nonetheless, it seems to be revealing some fundamental limitations about what is humanly possible and offering up several clues to essential features of human beings' cognitive/neural architecture.

Suggestions for Further Readings

The book edited by R. Parasuraman and D. R. Davies (1984) contains a useful collection of articles on many different aspects of attention—including important topics in attention not discussed here, such as selective attention, set, and vigilance—although many recent developments are not included. A very clearly written analysis of capacity limitations and how they can be demonstrated is by John Duncan (1980a). For those interested in learning more about decision noise and how it can be modeled in the context of threshold perception Norma Graham (1989) is lucid and definitive. A large literature describes the use of reaction times to analyze visual capacity limits; see, for example, Treisman and Gelade (1980) and Wolfe, Cave, and Franzel (1989), both of which focus on perceptual combination of features rather than recognition of familiar forms like letters. James Townsend and F. Gregory Ashby (1983) describe a wide variety of formal methods for discriminating parallel from serial processing. For a brief description of how limitations in dual-task performance have been analyzed using the psychological refractory effect, see Pashler (1993); for a more detailed review of studies of dual-task interference, see Pashler (1994b).

Problems

2.1 Consider the task of determining whether a display contains a barely detectable dark gray square (presented against a black background). In the *location-uncertain* condition the observer knows that the square might occur in any of four possible positions, in a *location-certain* condition, the observer is told in advance where the square will occur if it is presented. If detection performance is better in the location-certain condition than in the location-uncertain condition, does that imply that people can allocate some sort of attentional capacity to a specified location and that this attentional capacity improves the processing of faint sensory input?

2.2 Consider the following hypothetical experiment. The subject sees a display of digits and simultaneously hears a spoken digit. The task is to verify that the display contains at least one token of each of the digits from one to the spoken digit. For example, if subjects hear FIVE their task is to verify that there is at least one 1, one 2, one 3, one 4 and one 5 in the display. Consider only the RTs on positive trials—that is, the trials in which subjects correctly respond YES—and, specifically, the effects of two variables: (1) what the spoken digit is; and (2) the visual quality of the display (degraded or normal). Suppose it should turn out that there is a *multiplicative interaction* between the spoken digit and visual quality: that is, RTs are slower when the display is degraded, and the amount of slowing is in direct proportion to the particular spoken digit (e.g., three times as much slowing to respond to THREE than ONE). Offer a straightforward interpretation of this finding and relate it to some of the ideas described in the text.

2.3 Casual observation suggests that people can drive while having a conversation but that doing so becomes risky if the traffic is heavy or the road is narrow. How would you reconcile this general impression with the findings about central and perceptual capacity limits described in the text?

References

- Brindley, G. S. (1970). *Physiology of the retina and visual pathway* (2nd ed.). London: Edward Arnold.
- Broadbent, D. E. (1982). Task combination and the selective intake of information. *Acta Psychologica* 50, 253-290.
- Cattell, J. M. (1835/1947). The inertia of the eye and brain. In A. T. Poffenberger, ed., *James McKeen Cattell, 1860-1944, Man of Science*. Lancaster, PA: Science Press.
- Cheesman, J., and P. M. Merikle (1956). Distinguishing conscious from unconscious perceptual processes. *Canadian Journal of Psychology* 40, 343—367.
- Davis, R. (1959). The role of "attention" in the psychological refractory period. *Quarterly Journal of Experimental Psychology* 11, 211-220.
- Dennett, D. C. (1991). *Consciousness Explained*. Boston: Little, Brown.
- Duncan, J. (1979). Divided attention: The whole is more than the sum of its parts. *Journal of Experimental Psychology: Human Perception and Performance*. 5, 216-228.
- Duncan, J. (1980a). The demonstration of capacity limitation. *Cognitive Psychology* 12, 75-96.
- Duncan, J. (1980b). The locus of interference in the perception of simultaneous stimuli. *Psychological Review* 87, 272-300.
- Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General* 113, 501-517.
- Duncan, J. (1985). Visual search and visual attention. In M. I. Posner and O. S. M. Mavin, eds., *Attention and Performance XI*, 85-105. Hillsdale, NJ: L. Erlbaum.
- Duncan, J. (1987). Attention and reading: Wholes and parts in shape recognition—A tutorial review. In M. Coltheart, ed., *Attention and Performance XII: The Psychology of Reading*, 39-61. London: L. Erlbaum.
- Duncan, J. (1993). Similarity between concurrent visual discriminations: Dimensions and objects. *Perception and Psychophysics* 54, 425-430.
- Epstein, W., T. Babler, and S. Bownds (1992). Attentional demands of processing shape in three-dimensional space: Evidence from visual search and precuing paradigms. *Journal of Experimental Psychology: Human Perception and Performance* 18, 503-511.
- Eriksen, C. W., and T. Spencer (1969). Rate of information processing in visual perception: Some results and methodological considerations. *Journal of Experimental Psychology: Monograph* 79, 1-16.

- Estes, W. K., and H. A. Taylor (1964). A detection method and probabilistic models for assessing information processing from brief visual displays. *Proceedings of the National Academy of Science* 52, 446-454.
- Fagot, C, and H. Pashler (1995). Repetition blindness: Perception or memory failure? *Journal of Experimental Psychology: Human Perception and Performance* 21, 275-292.
- Graham, N. (1989). *Visual patten: analyzers*. New York: Oxford University Press.
- Green, M. (1991). Visual search, visual streams, and visual architectures. *Perception and Psychophysics* 50, 388-403.
- Green, N. F., and L. K. Anderson (1956). Color coding in a visual search task. *Journal of Experimental Psychology* 51, 19-24.
- Irwin, D. E. (1991). Information integration across saccadic eye movements. *Cognitive Psychology* 23, 420-456.
- James, W. (1890/1950). *Principles of Psychology*, vol. 1. New York: Dover.
- Johnston, J. C., and B. L. Hale (1984). The influence of prior context on word identification: Bias and sensitivity effects. In H. Bouma and D. G. Bouwhuis, eds., *Attention and Performance X: The Control of Language Processes*, 243-256. Hillsdale, NJ: L. Erlbaum.
- Johnston, J. C., and H. E. Pashler (1991). Close binding of identity and location in visual feature perception. *Journal of Experimental Psychology: Human Perception and Performance* 16, 843-856.
- Kahneman, D., and A. Treisman (1984). Changing views of attention and automaticity. In R. Parasuraman and D. R. Davies, eds., *Varieties of Attention*, 29-62. New York: Academic Press.
- Mackay, D. M. (1962). Theoretical models of space perception. In C. A. Muses, ed., *Aspects of the theory of artificial intelligence*. New York: Plenum.
- McCann, R. S., and J. C. Johnston (1992). Locus of the single-channel bottleneck in dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance* 18, 471-484.
- McConkie, G. (1990). Where vision and cognition meet. Paper presented at the H. F. S. P. Workshop on Object and Scene Perception, Leuven, Belgium.
- Moore, J. J., and D. W. Massaro (1973). Attention and processing capacity in auditory recognition. *Journal of Experimental Psychology* 99, 49-54.
- Moray, N. (1975). A data base for theories of selective listening. In P. M. A. Rabbitt and S. Domic, eds., *Attention and Performance V*, 119-135. New York: Academic Press.
- Mozer, M. C. (1991). *The Perception of Multiple Objects: A Connectionist Approach*. Cambridge, MA: MIT Press.
- Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General* 106, 226-254.
- Neisser, U., and R. Becklen (1975). Selective looking: Attending to visually specified events. *Cognitive Psychology* 7, 480-494.
- Osman, A., and C. M. Moore (1993). The locus of dual-task interference—Psychological refractory effects on movement-related brain potentials. *Journal of Experimental Psychology: Human Perception and Performance* 19, 1292-1312.
- Ostry, D., N. Moray, and G. Marks (1976). Attention, practice and semantic targets. *Journal of Experimental Psychology: Human Perception and Performance* 2, 326-336.
- Parasuraman, R. and D. R. Davies eds., (1984). *Varieties of Attention*. Orlando: Academic Press.
- Pashler, H. (1984). Evidence against late selection: Stimulus quality effects in previewed displays. *Journal of Experimental Psychology: Human Perception and Performance* 10, 429-448.

- Pashler, H. (1987). Detecting conjunctions of color and form: Reassessing the serial search hypothesis. *Perception and Psychophysics* 41, 191-201.
- Pashler, H. (1988). Familiarity and visual change detection. *Perception and Psychophysics* 44, 369-378.
- Pashler, H. (1989). Dissociations and dependencies between speed and accuracy: Evidence for a two-component theory of divided attention in simple tasks. *Cognitive Psychology* 21, 469-514.
- Pashler, H. (1993). Doing two things at the same time. *American Scientist* 81, 48-55.
- Pashler, H. (1994a). Divided attention: Storing and classifying briefly presented objects. *Psychonomic Bulletin and Review* 1, 115-118.
- Pashler, H. (1995b). Dual-Task interference in simple tasks: Data and theory. *Psychological Bulletin*.
- Pashler, H. (in press). Structures, processes, and the flow of information. In E. Bjork and R. Bjork, eds., *Handbook of Perception and Cognition*, vol. 10, *Memory*. Orlando, FL: Academic Press.
- Pashler, H., and P. C. Badgio (1985). Visual attention and stimulus identification. *Journal of Experimental Psychology: Human Perception and Performance* 11, 105-121.
- Pashler, H., and P. Badgio (1987). Attentional issues in the identification of alphanumeric characters. In M. Coltheart, ed., *Attention and Performance XII*, 63-81. Hillsdale, NJ: L. Erlbaum.
- Pashler, H., M. Carrier, and J. Hoffman (1993). Saccadic eye movements and dual-task interference. *Quarterly Journal of Experimental Psychology* 46A, 51-82.
- Pashler, H., and J. C. Johnston (1989). Chronometric evidence for central postponement in temporally overlapping tasks. *Quarterly Journal of Experimental Psychology* 41(nl-A), 19-45.
- Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception and Psychophysics* 16, 283-290.
- Phillips, W. A. (1983). Short-term visual memory. *Philosophical Transactions of the Royal Society*, London, B302, 295-309.
- Pohlmann, L. D., and R. D. Sorkin (1976). Simultaneous three-channel signal detection: Performance and criterion as a function of order of report. *Perception and Psychophysics* 20, 179-136.
- Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Learning, Memory and Cognition* 2, 509-522.
- Puleo, J. S., and R. E. Pastore (1978). Critical-band effects in two-channel auditory signal detection. *Journal of Experimental Psychology: Human Perception and Performance* 4, 153-163.
- Shiffrin, R. M., and G. T. Gardner (1972). Visual processing capacity and attentional control. *Journal of Experimental Psychology* 93, 78-82.
- Shiffrin, R. M., S. T. Dumais, and W. Schneider (1981). Characteristics of automatism. In J. B. Long and A. D. Baddeley, eds., *Attention and Performance IX*, 223-238. Hillsdale, NJ: L. Erlbaum.
- Sperling, G., (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, Whole No. 498, 1-29.
- Townsend, J., and F. Gregory Ashby (1983). *Stochastic modeling of elementary psychological processes*. Cambridge, Eng.: Cambridge University Press.
- Treisman, A., and A. Davies (1973). Dividing attention to ear and eye. In S. Komblum, ed., *Attention and Performance IV*, 101-117. New York: Academic Press.
- Treisman, A., and G. Gelade (1980). A feature integration theory of attention. *Cognitive Psychology* 12, 97-136.
- Treisman, A., and H. Schmidt (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology* 14, 107-141.

- Welford, A. T. (1952). The "psychological refractory period" and the timing of high speed performance—A review and a theory. *British Journal of Psychology* 43, 2-19.
- Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics *Acta Psychologica* 41, 67-85.
- Wolfe, J. M., K. R. Cave, and S. L. Franzel (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance* 15, 419-433.
- Yantis, S., and J. C. Johnston (1990). On the locus of visual selection: Evidence from focused attention tasks. *Journal of Experimental Psychology: Human Perception and Performance* 16, 135-149.