Doing Two Things at the Same Time

How many things can we do at once? Fewer than we think, say psychologists, who are identifying "bottlenecks" in the process

Harold Pashler

People routinely do two or more things at the same time. They eat breakfast while reading the newspaper; they make calls on a cellular telephone while driving a car; and, of course, they chew gum while walking. If you ask people to assess their own ability to keep two activities going at once, they generally report difficulty only if one of the tasks is intellectually demanding; for example, they may have a hard time carrying on a serious discussion while adding up a restaurant check. If the tasks are routine, people are quite sure they can handle at least two simultaneously.

Recent research suggests that these assessments of the human capacity for parallel processing may be over-optimistic. It appears that certain mental operations are "bottlenecks" that require the exclusive use of some cognitive resource and therefore cannot be done concurrently. In particular, even the most trivial forms of decision-making and memory retrieval seem to be activities that cannot be overlapped with other operations. Much of what people perceive as parallel processing in mental life may really be more like computer time sharing, in which some mental operations are actually carried out one at a time.

At the level of neurons there is no question the brain does a great deal of parallel processing. Hundreds of millions of nerve cells must act in concert if a person is to perceive, think and act. But such concurrent neural activity does not necessarily imply parallelism at the level of thought or other mental processes. Likewise in an electronic computer, thousands of transistors may be active at once, even though only one computational process is under way. True parallel processing, in the brain or in a computer, entails the simultaneous activity of larger units than neurons or transistors.

Harold Pashler received bachelor's degrees in philosophy and psychology from Brown University, and a Ph.D. in psychology from the University of Pennsylvania. He is currently associate professor of psychology at the University of California at San Diego. Address: University of California, San Diego, Department of Psychology, 9500 Gilman Drive, La Jolla, CA 92093

Experimental psychologists have lately shown much interest in studying dual-task performance. One reason for this interest is simple curiosity about human capacities and limitations; one would like to know what is humanly possible and what isn't. Limits on parallel processing may also be of practical importance in fields such as aviation, where pilots flooded with stimuli must respond rapidly and correctly. Beyond these issues, there is an intellectual reason for interest in dual-task performance: It may provide clues about general principles of brain function that would not otherwise be evident. A common way to study a complex system is to see how it performs when it is overloaded, and this principle has already been applied fruitfully in investigating many different sorts of natural and artificial systems. By watching what happens when people attempt to do more tasks than they can readily handle at one time, psychologists hope to learn how activities in one part of the brain affect or interfere with activities in other parts.

Dual-Task Performance in the Laboratory

Studies of human performance in the laboratory show that people have some surprisingly stubborn limitations on their ability to carry out different mental operations at the same time. The nature of the limitations depends on the sorts of tasks being tested. If one simply tries to produce a fixed set of repeated motor responses while doing something else, very little interference usually arises. For example, people can tap their fingers rapidly while answering questions or while making vocal responses to tones. Much more difficulty arises, however, in producing two streams of response with incompatible timing. Stuart Klapp of California State University at Hayward showed that people have great difficulty tapping different rhythms that are not harmonically related to each other with each hand; for example, maintaining three time with the left hand and four time with the right (Klapp 1979), something familiar to just about everyone who has played the piano. Work by Steven Keele and Richard Ivry, who were then at the University of Oregon, suggests that both perception and production of timing and rhythm may depend on a single common timer in the brain, perhaps located in the cerebellum, the large cauliflower-shaped region in the back of the brain concerned with motor activities.

In the situations just mentioned, the person is asked to produce a fixed pattern of responses regardless of the sensory input received. Much more interesting things happen when the person tries to produce different responses depending on which external stimuli are presented. In the 1940s, Kenneth Craik and Margaret Vince at the University of Cambridge in England found that people had trouble responding to different stimuli presented close together in time.

Current research analyzing these difficulties usually presents people with two different "choice-response tasks." An example of a choice-response task is deciding whether a tone stimulus is high or low in pitch, and responding verbally, saying "high" or

Figure 1. Circus performers dazzle their audiences with the number of complex things they can do at once. Looking at them, one gets the impression that an infinite number of tasks can be performed simultaneously. But experimental psychologists are finding that people have some surprisingly stubborn limitations on their ability to carry out different mental operations at the same time.

"low." People typically take half to three-fourths of a second to respond in this sort of task. In a typical dual-task experiment, a subject might perform the task just described, plus a second task. Shortly after the first (tone) stimulus (*S1*), a letter would appear on a computer screen (*S2*), and the subject's task would require pressing one of several buttons depending on the identity of this letter.

The response time for the first task is often little affected by the interval between *S1* and *S2*. As the interval is reduced toward zero (simultaneous stimulation), the response time for the second task is regularly found to increase. In many cases, subjects' response time increases one millisecond for each millisecond decrease in the time between stimuli. This means that people cannot produce the second response any sooner than they do produce it, even if the stimulus for that response is made available earlier.

Even though the second response is slowed at short 51-52 intervals, there is still evidence for some overlap between the two tasks. Specifically, the total time elapsed between presentation of the first stimulus and the response to the second one is usually less than the time it would take the person to carry out both tasks independently. Results like the ones found by Craik and Vince have been found with a variety of motor responses, including vocal responses, movements of the hands, feet and eyes, and with various kinds of sensory stimuli for the subjects to respond to—tones, spots, letters and words that are seen or heard.

Bottlenecks

Why would the response to the second stimulus be slowed? This question has intrigued psychologists for several decades, and it is only recently that some consensus has emerged. Most theorists proposed that performance must be limited by some kind of mental bottleneck—that is, a stage or operation in the tasks that cannot be performed simultaneously. Even these very simple kinds of tasks can be broken down into a number of stages, including at least the following three: perceiving the stimulus, choosing the response, and actually producing the response. Recent work shows that these stages are for the most part successive; one stage is largely completed before the next stage begins (Miller 1988).

The situation is analogous to one that occurs in everyday experience—going to the bank. If you walk into a bank with a single teller on duty, and another customer is ahead of you, the teller will be a bottleneck that slows down your progress. You have to wait until the teller has finished dealing with the first customer before you can complete your banking. Plotting the time you spend in the bank against the interval between the moment the customer entered the

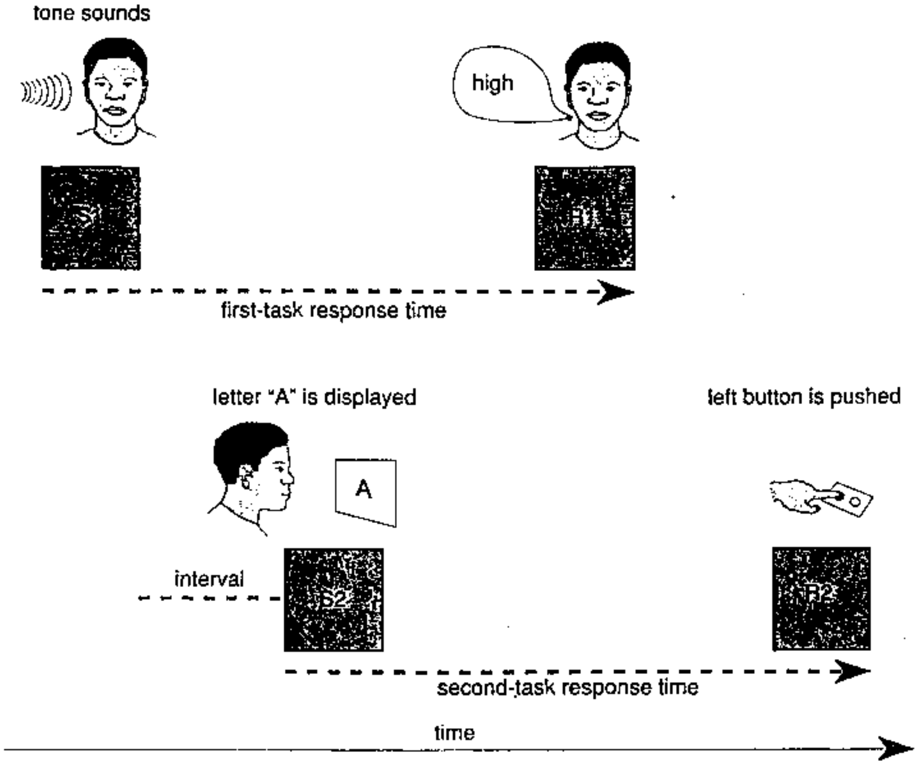

Figure 2. Experiments constructed by cognitive psychologists explore the limits of a subject's ability to perform multiple tasks simultaneously. A typical dual-task experiment is diagrammed here. The subject is presented with one stimulus, labeled *S1*, to which he is asked to make a specific response, *R1*. In the case shown, *S1* is a tone, which the subject identifies as having a high pitch; his response, therefore, is to say "high." After *S1*, the subject is presented with a second stimulus, *S2*, which in this case is a visual display of the letter "A." His response, *R2*, is to press the leftmost of several response keys. The two response times (from *S1* to *R1* and from *S2* to *R2)* are measured in the experiment. By altering the interval between stimuli or by altering the complexity of either the stimuli or the responses, psychologists have learned a great deal about the mental processes required for dual-task performance.

bank and the moment you did, one would find exactly the kind of results observed in the experiments involving dual-choice response tasks.

In the banking example, the bottleneck is caused by limits on the teller's speed. In dual-task performance, what activity is acting like the teller? Where does the bottleneck occur? One theory proposes that the bottleneck is in stimulus perception. Identification of a second stimulus has to wait until the first stimulus has been fully perceived. Another bottleneck theory suggests that people cannot produce two motor responses at the same time, even though the responses may be quite different from each other, such as speaking a word and moving a finger. Both of these theories would be consistent with the observation that response time increases as the interval between stimuli decreases. Both theories are also consistent with the evidence for some overlap between tasks, as mentioned previously

Recent studies have argued against both of these theories. Instead, current work seems to clinch the case in favor of a third bottleneck theory. The third theory, which was first suggested by Alan Welford in 1952, locates the bottleneck in the process of choosing which response is to be made. (The bottleneck may also include the initiation of the response, but not its execution.)

It may seem odd to think that choosing a response can cause a bottleneck. The task of selecting one button in response to a high pitch or a different button in response to a lower pitch does not seem to be very effortful. What is the evidence that this introspectively minimal operation constitutes such a fundamental limit on our parallel processing abilities? One kind of evidence comes from experiments in which different stages in each task are deliberately made more difficult, and the effects on the reaction times are observed.

The nature of such experiments can be understood by returning to the banking analogy. Suppose you decide to carry out some extra business with the teller, such as checking the balance in your savings account, which takes an extra minute to do. This slows down the bottleneck stage of the process (time spent with the teller), and it obviously increases the time you spend in the bank by one minute. The minute added will not depend on the interval between customer arrivals. On the other hand, suppose you take an extra minute deeding with preliminary paperwork before seeing the teller. While you are doing this, the teller is occupied with the first customer, so you will have to wait in any case. If you enter the bank right after the first customer, and if the first customer takes enough of the teller's time, then the extra time you spend filling out forms will not result in your spending any more time in the bank.

By applying the logic of the bank to the sequence of processing stages in the human brain, one can make some very strong tests of the different bottleneck theories of dual-task interference. If the bottleneck occurs at the response-selection stage for the second task, then mak-

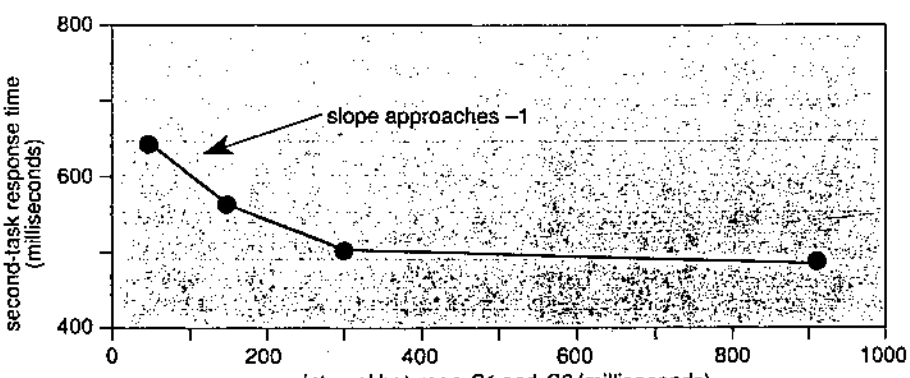

Figure 3. People have trouble responding to different stimuli presented close together in time. Here the response time for the second task is plotted as a function of the interval between *S1* and *S2*. As the interval decreases, the response slows. When the interval is small enough, the slope approaches -1, indicating that *R2* cannot be produced any sooner than it is produced, even if the stimulus *S2* is presented earlier. Observations like these have led to the proposal that some part of the process acts as a bottleneck, preventing people from producing the second response until some aspect of the first task has been completed.

ing this stage a little more time-consuming should add a constant onto the response time for the second task. On the other hand, if a stage of the second task before the bottleneck (namely, the perception of the stimulus) is made more time-consuming, this should not hold up the response time for the second task at all, assuming the interval between the stimuli is sufficiently short.

How can one make different stages of these simple tasks take longer? A variety of methods can be employed (Stern-berg 1969). For example, to make the perceptual processing of a letter of the alphabet take more time, one can simply reduce the intensity of the letter or superimpose tiny dots on it. To make the selection of a response take longer, one can change the relationship between the stimuli and the responses. For example, response selection is relatively easy if the digits 1, 2, 3 and 4 are associated with four response buttons arrayed horizontally in front of the subject (so that 1 means "press the leftmost key," 2 means "press the second key from the left" and so on). A greater challenge can be introduced if the mapping of digits to buttons is scrambled, so, for example, 3 becomes the cue to press the leftmost key The increased difficulty of choosing the response should yield a greater response time. The difference is in the process of determining which response is to be made, rather than in the identification of the stimuli or in the actual production of the key-press responses, since the altered arrangement involves no change in the stimulus or in the response.

My own experiments, in collaboration with James Johnston of the NASA Ames Research Center in California and several graduate students, have assessed responses in a range of dual-task experiments in which we increase the difficulty of various parts of each task. The results confirm predictions of the model that favors a bottleneck at the response-selection stage of the process, and they rule out competing bottleneck models (Pashler 1984, Pashler and Johnston 1989, McCann and Johnston 1992). We find that making the second task response selection more time-consuming adds a constant to the reaction time for that task. However, when we slow down the perceptual processing in that task, the effect on reaction time is reduced and even disappears when the interval between the two tasks is made sufficiently short.

Our results provide strong support for a model in which response selection for the second task cannot be initiated until after response selection—but not response production—for the first task has been completed. The evidence for this is related to, but differs slightly from, the experiments just described. Cinda Christian, a student in my laboratory, and I carried out experiments where we varied the time required to produce the first response but not the selection time. To make the first response more time-consuming, we changed the number of key presses for the response. The second task always required just a single rapid vocal response to a tone. We found that the second response was slowed as the interval between the stimuli was reduced, which is expected. The interesting tiling, however, was that although it took people longer to complete the first response when it required several key presses rather than one

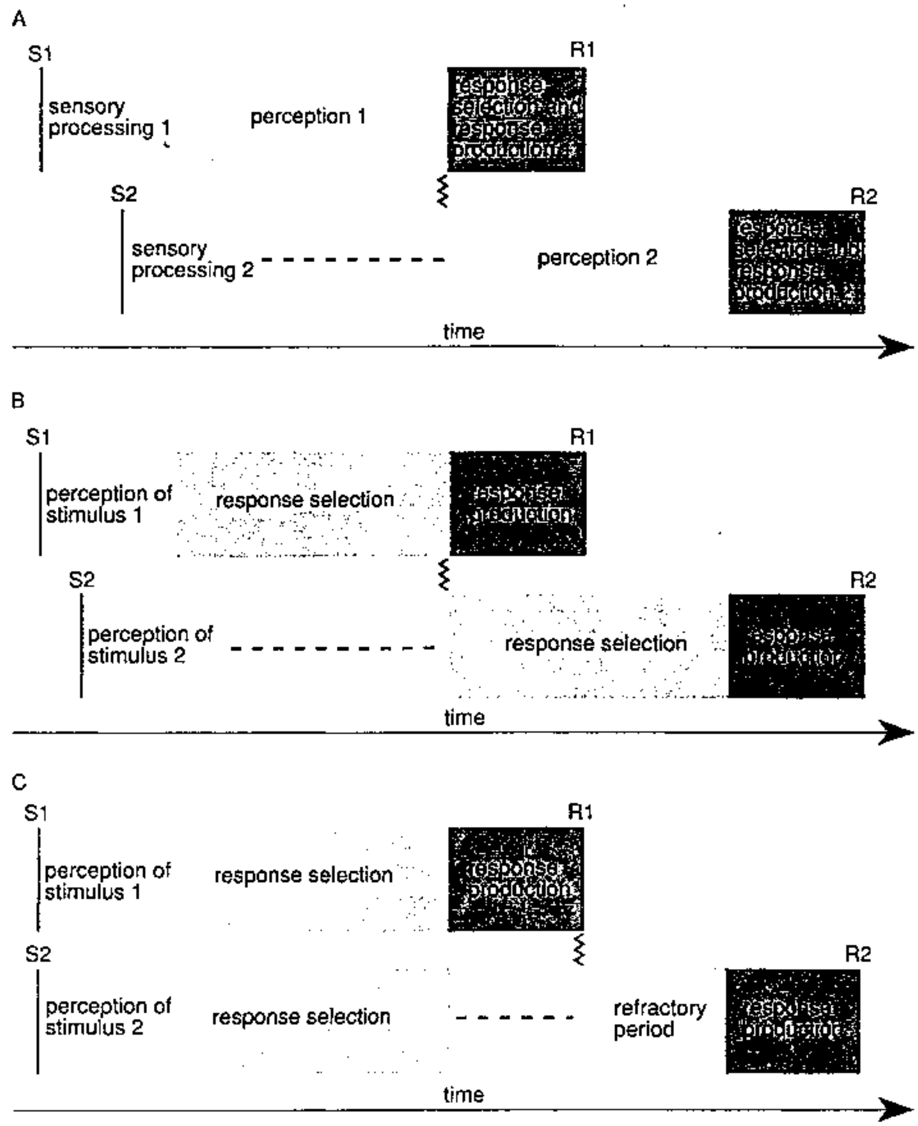

Figure 4. Bottlenecks could constrict any one of three stages in the performance of a task. If perception of the first stimulus held up further processing, then events would proceed as diagrammed in *A*. Here, the subject is potentially capable of processing two stimuli at once but must perceive them one at a time. Hence the first stimulus must be perceived before the second stimulus can be, but thereafter response selection and production of the first response can proceed while the second stimulus is being perceived. Diagram B depicts the situation that would arise if response selection caused the bottleneck. In that case, *S1* and *S2* could be perceived nearly simultaneously, but the second response could not be chosen until the first response selection had been completed. The third possibility, shown in C, allows both stimulus perception and response selection for the two tasks to proceed simultaneously, but the second response cannot be produced until the first response has been completed. Proponents of this third hypothesis suggest that the motor control system in the brain must "reset" itself after each use, and that this limits a person's ability to produce a second response immediately after the first response has been made. Increasing experimental evidence favors the model presented in *B*, suggesting that the bottleneck is at the response-selection stage.

(no surprise there), the response to the second task occurred at about the same time whether it followed a one-key-press task or a several-key-press task. In fact, when the first task involved several key presses, subjects usually made the vocal response for the second task while the sequence of key-press responses in the first task was still underway.

The process is reminiscent of an executive whose style is to delegate tasks. The experiments show that a single mental mechanism chooses the response to be made in the first task. But like the trusting executive, this device does not have to hang around to watch while its instructions are carried out. Instead, it can move on to choose the response for the second task, allowing each response to be carried out by subordinates. If the responses to the first task take sufficiently long to execute, a person may still be performing them after the response to the second task has been completed.

Taken together, these data suggest that the limitations in carrying out stimulus-response tasks concurrently are not introduced at the level of stimulus perception, nor in production of the motor response. Those mental operations can work in parallel. Rather, the problem is in deciding what the response will be, and this kind of mental operation seems to be carried out in series—that is, one task at a time.

Memory Retrieval and the Bottleneck

It seems that the bottleneck just described arises not only when the brain chooses an action to perform, but any time information is explicitly recalled from memory. A graduate student in my laboratory, Mark Carrier, and I took memory-retrieval tasks that are much more time-consuming than the choice-reaction tasks described above, and combined them with other concurrent activities. For example, subjects might learn a long list of word pairs such as FISH-DONKEY. In the memory test, subjects would be presented with a first word from one of the pairs (FISH) and their job would be to recall the second (DONKEY). Subjects were sometimes asked to do this while performing a second task that involved pushing a button in response to a tone. Their results showed that the tone task delayed the process of retrieving the second word from memory.

The issue takes on practical significance when we consider the plight of students taking an exam. Many students try to improve their performance by reading over all the questions first before answering any, with the hope that this might allow their brain to start work on all the questions in parallel.

Carrier and I tried to see whether we could find evidence of such parallel processing. We gave our subjects problems consisting of a category and the first letter of a word to be retrieved. Thus a correct answer to the cue VEGETABLE S would be squash; another example is FISH M, to which an answer might be mackerel. We first examined the rate at which our subjects could solve individual problems of this sort, and then tried giving them two such problems to work on at the same time. If they could actually carry out both independently and in parallel, then this should have been reflected in the probability with which they were able to find the answers to both problems within the time period we allowed them. In fact, subjects were much less likely to get the answer to both problems than the parallel processing model would have predicted, arguing that they were really only able to work on one problem at a time. Of course, it remains to be seen whether this result will generalize to the many different sorts of problems people try to solve.

A Further Test: Brief Displays

Our present hypothesis is that the mind is capable of a lot of parallel processing, but it chooses actions and recalls from memory by serial processing, which produces many of the bottlenecks described here. On the other hand, perceiving one stimulus while choosing a response to another ought to be perfectly possible to accomplish in parallel. A natural way of testing this hypothesis further is to flash a visual stimulus briefly while someone is doing another task to see whether the person can perceive and remember the visual stimulus and then *later* say what was seen. In some of our experiments, we had people respond as rapidly as possible to a high-pitched tone by pressing one button and to a low-pitched tone by pressing another (Pashler 1989). Shortly after the tone sounded (anywhere from 0.05 to 1.5 seconds afterwards) we would very briefly flash a visual display of letters, digits or other visual stimuli. These items remained on the screen for only a fraction of a second, after which they were replaced by a meaningless visual pattern in the same part of the visual field. This pattern is referred to as a mask, and it serves to wipe out any after-images that might otherwise invalidate the logic of the experiments.

To see whether subjects were able to identify the visual stimuli before the

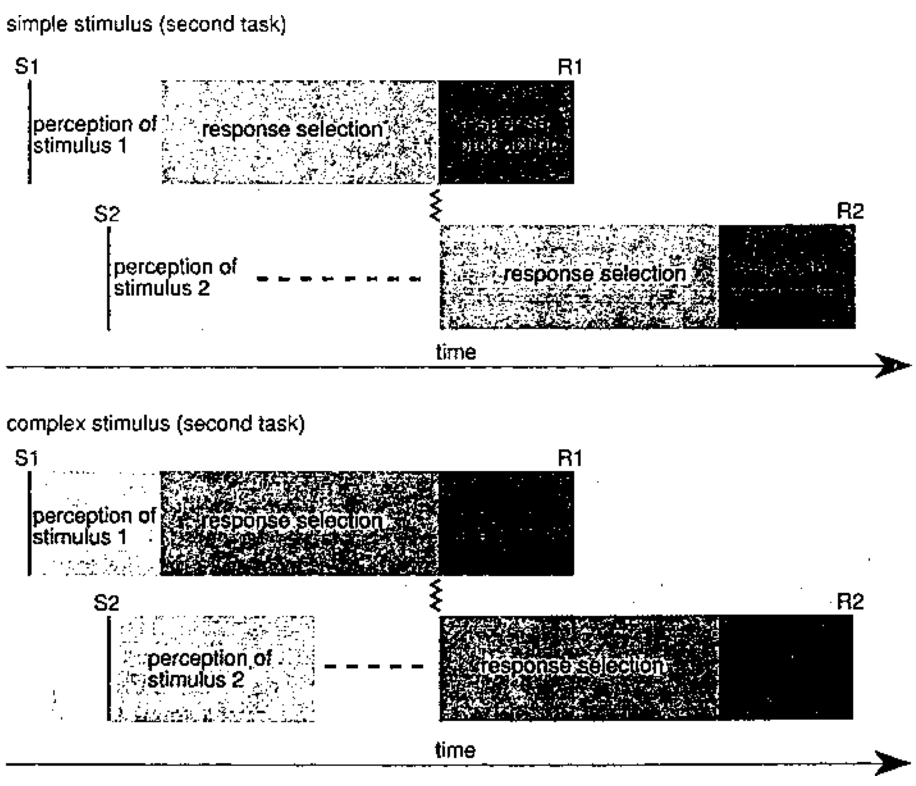

Figure 5. Response-selection bottleneck model predicts that making the second stimulus more difficult to perceive (for example, by reducing its intensity) should not change the time it takes to respond to the second task. In the upper diagram *S2* is easy to perceive, whereas in the lower diagram *S2* takes longer to perceive. Nevertheless, the increasing perception time is swallowed up by the gap between tasks, and the second response is produced at the same time in both cases.

mask appeared, we required them to report something about the stimuli. This report was unspeeded—that is, it was not given under time pressure, and we were interested only in the accuracy of the report. A variety of perceptual and symbolic judgments were tried, such as reporting what was the highest digit in an array of digits. Subjects' reports were just as accurate when the display was shown while they were performing the tone task as when the display appeared a short time after subjects had completed the first task. In other words, the subjects were perfectly able to identify the digits while performing the tone task, providing further evidence that the bottleneck responsible for dual-task interference must arise in central processes that are not concerned with perceiving.

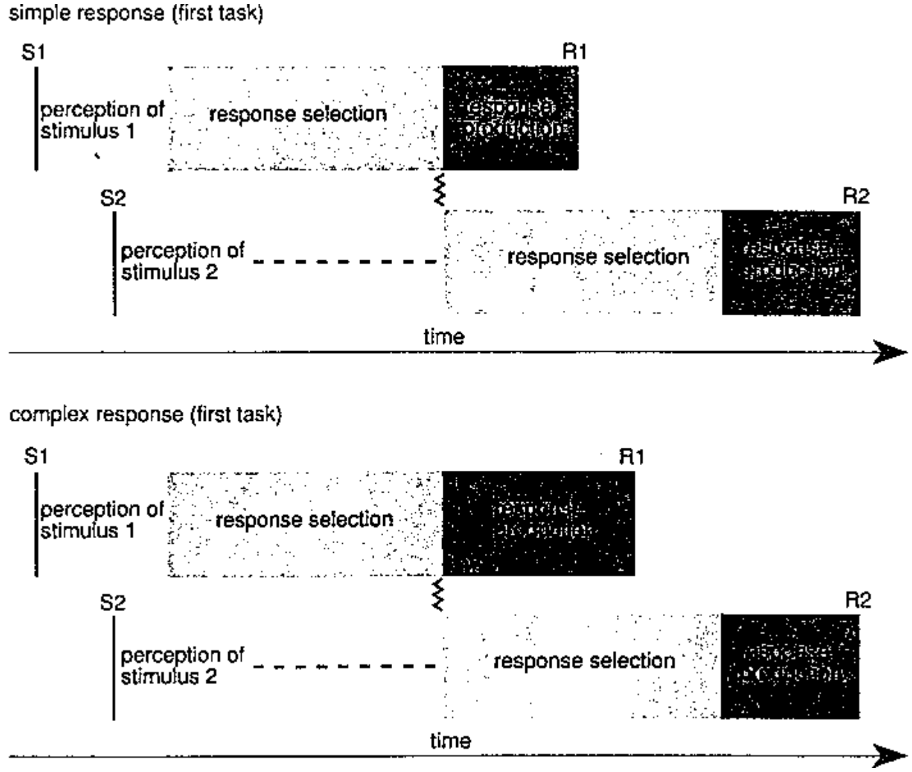

Figure 6. Increasing the complexity of the response does not change where the bottleneck occurs. When the first response is made more complex, so that it takes longer to produce it, the slowing does not propagate much onto the second task. This confirms that it is selecting the first response, not actually executing it, that constitutes a bottleneck.

Attention and the Bottleneck

People often use the term "attention" when talking about the difficulties they have in doing different tasks at the same time. But "attention" can also refer to the ability to voluntarily select one particular sensory stimulus from among a variety of competing inputs. For example, at a cocktail party, you can choose to attend to a single speaker's voice, which usually means you are left with little awareness or memory of what others are saying. Similarly, you can shift your attention among different visual stimuli without actually moving your eyes. You can verify this simply by fixing your gaze on one particular letter in a page of text. Now, have someone use a pencil to point to another letter near the one you are looking at. With a little practice, you will find that you can name the letter pointed to without moving your eyes even slightly.

"Attention" then describes both the process of selecting stimuli of interest and the limits that stand in the way of doing multiple tasks at the same time. This would seem to suggest that attention is a kind of inner resource, limited in capacity, which one voluntarily allocates to particular stimuli or activities. Do these ways of speaking reflect how the mind really works? Does the selection of sensory inputs really result from the same internal process as the bottleneck that limits a person's ability to do two tasks at the same time?

To address this question, my colleagues and I performed experiments very similar to those described in the preceding section. Here, our goal was to see whether people can shift their attention from one part of a visual display to another and at the same time select a response to a tone. If the same "attention resource" is used for selecting stimuli and choosing responses, these tasks should interfere with each other.

In our experiments, subjects made a manual response to a tone, as they did in earlier experiments, while they were required to shift their attention from one part of the visual field to another. An example of an attention-shift task is the following. A subject is presented with an arrow pointing to one of eight letters on a display, for a brief time before the display is obscured by a mask. The subject has to note and identify the letter a short while afterwards. Our results suggested there is no interference between the tone task and the visual task (Pashler 1991), even though the visual task required an attention shift. It therefore looks as though the casual use of the word "attention" is actually misleading. The serial processes of memory retrieval and response selection do not rely on the same mechanism that can focus on one stimulus or another.

In fact, sensory attention may be carried out in different areas of the brain from the likely sites of the bottleneck processes. Studies of patients with brain damage suggest that sensory attention mechanisms are located in regions to the rear of the cerebral cortex and in certain midbrain structures closely connected with those regions. On the other hand, selection of actions and memory retrieval seem to depend heavily on more frontal areas of the cerebral cortex (Requin, Riehle and Seal 1988).

Conflict with Common Sense?

The limitations that show up in the laboratory studies described above might seem to conflict with common sense. Why, one might ask, if people are unable to choose more than one response at a time, don't they appear to move like those robots in old movies that jerk first one limb and then the next? Of course, the very fact that those robots look unnatural indicates that people do not move that way; many parts of a person are easily seen to be moving at the same time. How is this to be reconciled with the conclusions drawn from our experiments?

Part of the answer is that the actual process of executing a motor response occurs after the response was selected, and this process may last for many seconds. As described above, subjects in dual-task experiments that required the person to produce a sequence of keypress responses sometimes completed the response to the second stimulus before finishing the sequence of responses

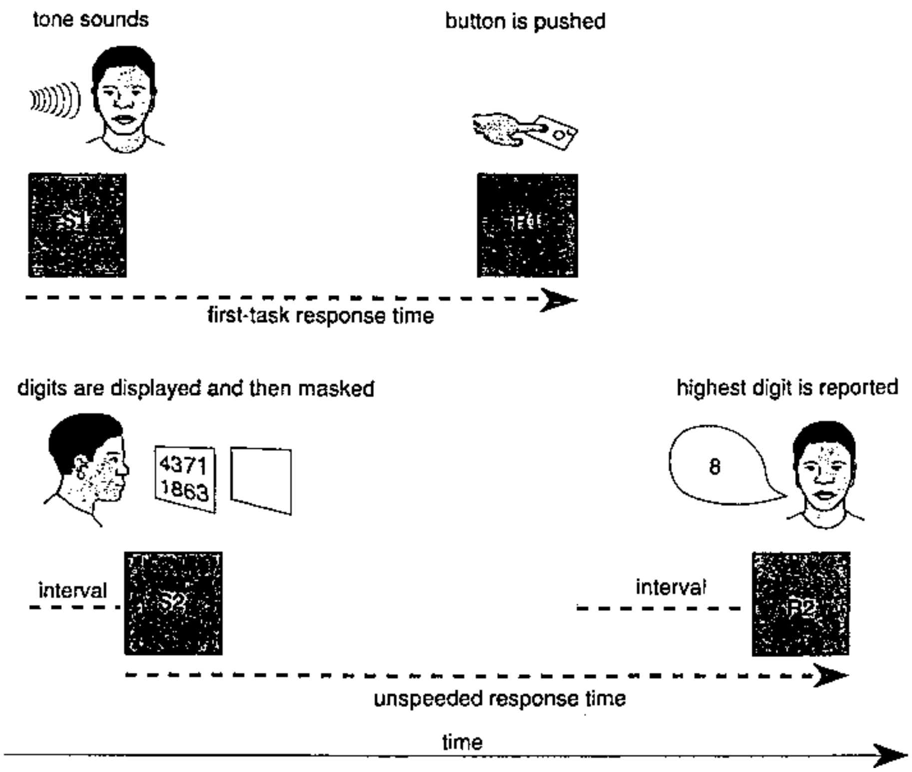

Figure 7. Further evidence for the response-selection bottleneck model comes from experiments that limit the time the second stimulus can be perceived. In an experiment like the one diagrammed here, subjects are asked to make the fastest possible response *(R1)* to the first stimulus, a tone, but can take their time in making the second response, which in this case is reporting the highest digit in a display of several numbers. The mask stops perception of the digits, so that if the interval between stimuli is sufficiently short, the numbers must be perceived during the tone task. Nevertheless, the accuracy of *R2,* the digit response, is barely affected by whether this interval is short or long, which argues once again that the bottleneck occurs at the point where subjects are selecting a response, not perceiving the stimuli.

to the first. This was true even though the selection of the second response was, we believe, delayed by the selection of the first response. The implication is that a person can talk and lift a glass toward his or her lips if the activities were not planned at the same time. More likely, they were selected and initiated at different moments.

Another reason this bottleneck may not be obvious to the casual observer is that many of the most common behaviors do not actually require a choice between responses to external stimuli. Walking, for example, consists largely of propellant motor sequences, and modifications based on sensory input are needed only intermittently when something goes awry. Even in the case of an activity like playing the guitar, response sequences may often proceed without constant adjustments based on sensory feedback, thereby making it possible for people to sing while strumming the instrument.

There is still another factor that may help to reconcile common sense with the findings of the laboratory. The tasks we have considered thus far are artificial in the sense that in each task, a single stimulus triggers a single movement or a single vocal response. Outside the laboratory, it often happens that seeing or hearing one thing summons up responses that involve several different motor movements. For example, you might reach for a ball with both hands at the same time, and if you see a friend in the distance, you might simultaneously shout and wave your arms.

Does each individual motion require a separate response-selection operation? Together with Clark Fagot, a graduate student in my laboratory, I investigated this question. We asked people to press one of three buttons, depending on the color of a square, while they named the color out loud. If the square was red, the subject would say *"red"* and press the middle response key; if the square was green, the subject would say "green" and press the right response key. We found that people were able to make each of these responses almost as quickly together as they did when they performed each task independently. There was also an extremely high correlation between the response times in the two tasks; that is, a fast response time in the first task was an accurate predictor of a fast response time in the second. Furthermore, when we used various tricks to make it more difficult to select the color response or more difficult to select the button-push response, the same amount of slowing showed up in response times for both tasks. Together, these results suggest that only a single response-selection operation was being carried out, and the two responses were produced as a unit (Fagot and Pashler 1992). It has also been noted that when each hand moves toward different targets at the same time, the movements are usually closely synchronized, so that both movements end at almost exactly the same time (Kelso, Southard and Goodman 1979).

It seems, then, that the fact that people do not move like old-fashioned robots is actually quite compatible with the conclusions described above. People may produce multiple streams of behavior at the same time in spite of very profound limitations in cognitive operations. Even if the streams overlap in time, the key mental operations that trigger them may not overlap. In some cases, multiple actions may be part of a single response stream.

Even tasks that at first are performed jerkily can become smoother with practice. People find that practiced activities require less effort and can seemingly be carried out "automatically" It is difficult to study the effects of extensive practice in the laboratory and not much is known about it. It is clear, though, that delays between two simple tasks are still observed, even after people have practiced a task for several thousand repetitions (Gottsdanker and Stelmach 1971). If practice does not really cause behavior to become automatic, it may simply allow one to choose larger units of response in a single "chunk." For example, after extensive practice, a typist may be able to select the response of typing an entire word, rather than just an individual letter. In that case, the mechanism responsible for the bottleneck may be called upon much less frequently, thereby making it possible for the typist to have other thoughts while typing.

What Causes the Bottleneck?

The results described so far suggest that even though people tend to think of themselves as being able to do multiple activities at the same time, certain kinds of mental operations (broadly speaking, finding something in memory or selecting an action) seem to be carried out one at a time, whereas other types of operations can readily overlap. But why should people have so much difficulty in choosing an action from a small number of possible responses? This would certainly seem to be a trivial process, from a computational perspective. Consider the problem of building a robot that could serve as a subject in one of the experiments described earlier, pressing the appropriate key in response to a particular letter of the alphabet. For the robot, both perceptual processing and motor control would pose a real challenge, but the response-selection stage would involve only the most trivial sort of "table lookup" operation—perhaps two lines of code in any programming language. Why is it so different for people?

The most straightforward interpretation of the bottleneck in response selection is that the brain contains a single piece of machinery, akin to the central processing unit of a computer, that takes a single retrieval cue (the output of the process of perceiving a stimulus) and looks up in its memory the corresponding action to be performed. This is an appealingly simple scenario. However, it conflicts with the results of another type of behavioral experiment that involves only a single task but multiple stimuli. Suppose a subject is instructed to press a button if he or she hears a high tone *or* sees the letter "X." One can compare the time required to respond to one stimulus (X or high tone) with the time to respond to both (X and high tone). It turns out that subjects make the appropriate response more quickly when offered redundant cues than they do when only a single cue is presented. Jeff Miller (1982) of the University of California at San Diego earned out some ingenious analyses of the results of experiments of this kind and concluded that the brain uses both cues jointly and simultaneously to retrieve the appropriate response. He termed this process *coactivation.*

Outside the laboratory, people often use several (sometimes quite weak) retrieval cues to converge on an appropriate concept in memory. This ability allows people to carry out tasks that are very difficult for digital computers. An example of such a task is solving a crossword puzzle. Here, one can take different kinds of cues that have never been experienced together and use them all to search memory (for example: novelist, South American, name ends with "Z"). It is difficult to see how this problem could be solved if one could search one's memory with only one cue at a time. To find the answer one cue at a time, one would first have to compile a list of all novelists, then of all South Americans, and then of all names that end in "Z" and then find the name common to all three lists. This process seems especially unlikely in light of the very limited temporary memory capacity people have.

If one imagines the brain working something like a digital computer or a clerk in a reference library, one is confronted with a paradox. The system appears able to take in two inputs and use them simultaneously to choose a single appropriate response from memory, but it cannot take two stimuli and use them simultaneously to choose two different responses. How can the system "decide" to process the stimuli sequentially only when they lead to different responses, especially before it has looked up what the responses are to be?

The resolution of this dilemma is not yet clear, but it may have some fundamental implications. One implication is that the metaphor of search and retrieval may be misleading. Some recent studies of neural networks (Hinton 1984) suggest an alternative view that might resolve the paradox. If memory retrieval involves establishing a characteristic pattern of activity in a large ensemble of neurons, then having multiple cues available may cause the appropriate final pattern to emerge more quickly. However, the neuronal pattern that enables one response to be performed may not be the pattern that enables other responses to be performed at the same time. Thus, the selection of two different responses would constitute a bottleneck even though coactivation would also occur.

Obviously, this account is nothing more than conjecture, and other alternatives will need to be considered. Fine-grained analysis of the behavior people produce when they try to perform more than one activity at the same time seems sure to lead to a deeper understanding of how human information-processing mechanisms work and how they differ from man-made information-processing devices. Ultimately, the results of behavioral experiments will have to be integrated with an understanding of how the brain actually carries out these computations. If the history of neuroscience is any guide, though, a biological understanding of simultaneous mental operations will require an analysis of these limits at the functional psychological level. Although this understanding is still sketchy, detailed study of response timing and accuracy with human subjects performing concurrent simple tasks is beginning to reveal some parts of the picture.

References

Carrier, M., and H. Pashler. Submitted. The attention demands of memory retrieval.

Fagot, C., and H. Pashler. 1992. Making two responses to a single object: Exploring the central bottleneck. *Journal of Experimental Psychology: Human Perception and Performance,* 18:1058-1079.

Gottsdanker, R., and G. E. Stelmach. 1971. The persistence of psychological refractoriness. *Journal of Motor Behavior* 3:301-312.

Hinton, G. E. 1984. Distributed representations. Technical Report CMU-CS-84-157. Pittsburgh: Department of Computer Science, Carnegie-Mellon University.

Keele, S., and R. Ivry 1991. Does the cerebellum provide a common computation for diverse tasks: A timing hypothesis. In *Developmental and Neural Basis of Higher Cognitive Function,* ed. A. Diamond, pp. 179-211. New York: New York Academy of Sciences.

Kelso, J. A. S., D. L. Southard and D. Goodman. 1979. On the coordination of two handed move merits. *Journal of Experimental Psychology: Human Perception and Performance* 5:229-238.

Klapp, S. 1979. Doing two things at once: the role of temporal compatibility. *Memory and Cognition* 7:375-381.

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.

Miller, J. O. 1982. Divided attention: evidence for coactivation with redundant signals. *Cognitive Psychology* 14:247-279.

Miller, J. O. 1988. Discrete and continuous models of human information processing: theoretical distinctions and empirical results. *Acta Psychologica* 67:1-67.

Pashler, H. 1984. Processing stages in overlapping tasks: evidence for a central bottleneck. *Journal of Experimental Psychology: Human Perception and Performance* 10:358-377.

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. 1991. Shifting visual attention and selecting motor responses: distinct attentional mechanisms. *Journal of Experimental Psychology: Human Perception and Performance* 17:1023-1040.

Pashler, H., and J. C. Johnston. 1989. Interference between temporally overlapping tasks: chronometric evidence for central postponement with or without response grouping. *Quarterly Journal of Experimental Psychology* 41 A:19-45.

Requin, J., A. Riehle and J. Seal. 1988. Neuronal activity and information processing in motor control: from stages to continuous flow. *Biological Psychology* 26:179-198.

Sternberg, S. 1969. The discovery of processing stages: extensions of Donders' method. In *Attention and Performance* II, ed. W. G. Koster, pp. 276-315. Amsterdam: North Holland.

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.