Journal of Experimental Psychology: General
1980, Vol. 109, No. 1, 98–117

Dividing Attention Without Alternation or Automaticity

William Hirst
Rockefeller University

Elizabeth S. Spelke
University of Pennsylvania

Celia C. Reaves, George Caharack, and Ulric Neisser
Cornell University

SUMMARY

Spelke, Hirst, and Neisser trained two subjects to copy unrelated words at dictation as they read and understood stories. The subjects' success was interpreted as evidence against the hypothesis of a fixed attentional capacity or limited cognitive resources; instead, it was hypothesized, attention is a skill that improves with practice. However, other explanations of these results can be proposed. The present research addressed two such counterhypotheses: that capacity may be alternated between reading and writing and that the writing task may become "automatic," and require no capacity at all.

Experiment 1 was designed to see whether subjects take intermittent advantage of the redundancy of the stories to switch to the writing task. Some subjects were trained to copy words while reading highly redundant material (short stories); others were trained with less redundant encyclopedia articles. On reaching criterion, each subject was switched to the other type of reading material. Three of the four subjects trained with stories transferred their skill immediately to the encyclopedia, suggesting that they had not been using the redundancy of the stories to accomplish their task.

Experiment 2 addressed that automaticity hypothesis. Two subjects were trained to copy complete sentences while reading. Several tests then showed that they understood the meaning of the sentences: (a) They made fewer copying errors with real sentences than with random words; (b) they recalled real sentences better than random words; (c) they integrated information from successive sentences, as demonstrated by a test of recognition memory for new statements whose truth was implied by the original ones. In view of this evidence that the sentences were understood, it is hard to maintain that they were being handled in an automatic way.

These results strengthen the hypothesis that the ability to divide attention is constrained primarily by the individual's level of skill, not by the size of a fixed pool of resources. Postulated capacity limits may provide plausible accounts of unskilled performance but fail to explain the achievements of practiced individuals.

 0096-3445/80/0901-0098$00.75

Most contemporary cognitive theories assert that mental activity is constrained by fixed limits on attention. Some theories posit a single mechanism of a definite "capacity" that underlies all complex mental functioning (Broadbent, 1971), whereas others assume that a family of mechanisms draws on a single pool of central processing "resources" (Norman & Bobrow, 1975). At first glance, the phenomena of divided attention seem consistent with these hypotheses. When an unpracticed individual attempts to carry out two complex activities at once, performance on at least one of them usually suffers (Messerschmidt, 1927; Moray, 1960; Neisser & Becklen, 1975; Paulhan, 1887; Stevenson, 1976; Treisman & Geffen, 1967). Such decrements may disappear, however, when the individual is given more experience with the tasks. Underwood (1974), Ostry, Moray, and Marks (1976), and Moray, Fitter, Ostry, Faveau, and Nagy (1976) have shown that extended practice can dramatically improve performance in two-channel monitoring. Brown and Poulton (1961) showed that as people become better drivers they can perform increasingly difficult mental calculations while they drive. These findings confirm the common observation that humans can learn to perform two independent and demanding tasks at once. It is not easy to reconcile this fact with any theory of central processing limits.

In our view, mental activities are best understood as perceptual and cognitive skills. The degree to which these skills interfere with one another depends on their detailed characteristics, and these can change radically as the skills develop. Thus, the limits of mental performance at any given time are determined by an individual's history rather than by fixed cognitive structures. Interestingly, some studies of the acquisition of motor skill support this view. In a recent investigation of the abilities of runners, Ryder, Carr, and Herget (1976) concluded that no physiological upper bound on speed and endurance sets any limit on the records that can be established in track events. The attainable record depends essentially on how hard contestants are willing to train. There may be practical and psychological factors that limit how much daily training a person will endure, but there is no ceiling set by physiology.

Two years ago, we reported dramatic practice effects in a study of writing while reading (Spelke, Hirst, & Neisser, 1976). After training, our subjects could carry out both activities at once with no decrement in reading speed or comprehension. Although we used these findings to argue against the assumption of limited capacity, they are not entirely conclusive. Several hypotheses might be suggested to reconcile our results with that assumption. In the present article, we report experiments that test two of these hypotheses: rapid alternation of attention and "automatic" processing of the dictated material. As will be seen, the results of the experiments indicate that both must be rejected and thus that the limited capacity assumption is unwarranted. The experiments also examine the generality of the particular skill of divided attention that we have studied and the conditions that facilitate its acquisition. Finally, they permit us to speculate about the changes that one activity undergoes when combined with another.

Solomons and Stein (1896) first demonstrated, and Downey and Anderson (1915) confirmed, that individuals could learn to read while writing at dictation. In our earlier work (Spelke et al., 1976), we observed the developing skills of two subjects, Diane and John, over a period of 20 weeks. The subjects read short stories while copying a series of unrelated words, dictated singly. At the beginning of the experiment, their reading speed and comprehension suffered seriously, but after 6 weeks these measures rose to control levels (that is, to the values attained when the subjects read without writing). At this point in the experiment, the subjects appeared to write in an auto-

This research was supported by Public Health Service Grant MH-27721 and a grant from the Cornell University Office of Academic Funding. The first two authors contributed equally to the work.

Requests for reprints should be sent to William Hirst, Rockefeller University, 1230 York Avenue, New York, New York 10021.

matic way: Probe trials indicated that Diane and John were not detecting semantic relations among the dictated words, even though they were copying them accurately. To encourage attention to these relations, the subjects were given additional trials in which they were asked to determine whether successive words belonged to the same category or formed a sentence. They learned to perform this task with little difficulty. With additional practice, they became able to assign the words to semantic categories during dictation (to write the category rather than the word itself) and thus demonstrated understanding of the meaning of each dictated word.

There seem to be three ways in which Diane's and John's achievements might be reconciled with the hypothesis of limited capacity. First, it is often assumed that the amount of available resources can be increased by exerting extra effort (Kahneman, 1973). Perhaps increased exertion accounted for the effect of practice in our experiment. Diane and John may have made only partial use of their available resources at first. As the experiment progressed, they may have allocated more of this resource pool and thus improved their performance in the dual task condition. Second, it might be suggested that the subjects did not really perform both tasks at once, but attended to them in rapid alternation (Broadbent, 1954; Paulhan, 1887). In other words, they may have learned to "time share" their capacity. Third, it has been argued that not all tasks must draw on central processing capacity: Simpler activities may be performed automatically (James, 1890; LaBerge & Samuels, 1974; Posner & Snyder, 1975; Schneider & Shiffrin, 1977; Shiffrin, 1975; Solomons & Stein, 1896). Diane and John may have bypassed their central mechanisms and copied the dictated words without attending to them.

The first of these hypotheses seems the least plausible, and the introspective reports of Diane and John refute it directly. They began the experiment enthusiastically and seemed to be trying as hard as they could. Nevertheless, they were quite unable to combine reading and writing at this stage. As practice continued and they become more successful, the subjects seemed to expend less effort rather than more. By the end of practice, their attitude had become casual and relaxed. This trend, which also appeared in the experiments reported below and has been noted in other contexts (cf. Norman & Bobrow, 1975), is just the opposite of what the effort hypothesis must predict.

The second and third hypotheses seemed to merit experimental investigation. Experiment 1 examined the possibility that individuals accomplish the dual task by learning to switch attention rapidly between reading and writing. We tested it by determining whether the skill would transfer to a dual task involving less redundant reading material with which such a switching strategy would be less effective. Experiment 2 tested the hypothesis that people learn to write automatically in our tasks. Subjects were trained to read while they copied sentences. Their understanding of the sentences and of implications across sentences was then tested; the abstraction of complex ideas cannot reasonably be described as an automatic process.

Experiment 1

People who learn to do two things at once often report that they alternate attention between them at early stages of practice; our subjects were no exception. The conjecture that practice simply allows one to do this more rapidly, and less consciously, is not entirely implausible; it can be advanced in defense of the assumption of limited capacity. But this hypothesis is testable only if it is made specific. One can always postulate a high enough switching rate to explain any observed simultaneity, but when the hypothesis is used in this way, the fixed central mechanism becomes an article of faith. In our experiment, however, a specific and plausible conception of time sharing can be suggested. It is rooted in the possibilities offered by the reading task itself. A subject who reads intermittently, turning attention away from the text when he must copy a dictated word, might prevent

a decline in reading speed or comprehension by making use of the redundancy of ordinary prose. Redundancy can facilitate comprehension of speed on a noisy channel (Marks & Miller, 1964; Miller & Isard, 1963), selection of the message among competing alternatives (Treisman, 1964), and the recognition of words (Miller, Bruner, & Postman, 1954). It is conceivable that Diane and John improved their reading and writing performance by learning to skip redundant words in the texts they were reading. To be sure, one might have expected them to adopt the same strategy for all their reading and thus to read faster on control trials as well; no such increase was observed. A more direct measure of subjects' use of redundancy seemed desirable, however, because of the plausibility of the time sharing hypothesis at early steps of practice and its popularity in models of performance at later stages.

In Experiment 1, subjects copied words at dictation while reading either highly redundant prose (short stories) or less redundant material from the encyclopedia. After learning to combine the two tasks with text at one level of redundancy, a subject was switched to text at the other level. The time sharing hypothesis makes two predictions for this experiment: (a) Subjects trained with short stories should learn to read while writing more easily than subjects trained with encyclopedia selections; (b) subjects trained with short stories should have difficulty in performing in dual task when tested with the less redundant encyclopedia articles, whereas those trained with encyclopedia material should transfer immediately to short stories.

Experiment 1 also provided an opportunity to replicate the basic finding of Spelke et al. (1976) that people can learn to read with normal speed and comprehension while copying dictated words. Finally, it examined the generality, over changes in the reading material, of the skill of reading while writing. To these ends, one group of subjects was given daily experimental trials reading short stories while copying at dictation and control trials consisting of reading stories with no dictation. A second group of subjects was similarly trained with encyclopedia articles. Practice continued until each subject read as fast and comprehended as well on experimental trials as on control trials. At this point, a subject was switched to reading matter at the other level of redundancy and was given experimental and control trials as before. We sought to determine, first, whether all the subjects would learn to read while writing. Since the subjects were not systematically preselected, such a finding would suggest that the skill previously acquired by Diane and John is attainable by most college-level adults. Second, we examined how fully subjects trained on one type of reading material would transfer to a second type. The degree of transfer provides evidence concerning the generality of the skills that subjects acquired. Finally, we compared the learning and transfer performance of subjects in the two groups. The time sharing hypothesis predicts that subjects trained with encyclopedia articles should take longer to reach criterion, and show greater transfer, than those trained with short stories.

Method

Subjects. Eight subjects were recruited from the Cornell undergraduate population and were paid by the hour for their services. Don, Jeff, Conrad, and Mary were assigned to Group S–E (they read stories first and then were transferred to encyclopedia selections) while Paul, Tom, Debbie, and David were assigned to Group E–S (encyclopedia first, then stories). David left the experiment during the training phase for personal reasons; we report results for the other seven subjects.

Stimulus material. Most of the short stories, by American, English, and translated European writers, were those used by Spelke et al. (1976). The encyclopedia selections were articles drawn from the *Encyclopedia Britannica* (1922) and the *International Encyclopedia of the Social Sciences* (1966). Stories and articles ranged from 700 to 2,500 words. Using the Cloze technique (Carroll, 1972), we confirmed that the encyclopedia articles were less redundant than the stories.[1]

[1] Every fourth word was etched out of copies of three stories and three encyclopedia articles, and 10 preliminary subjects were asked to replace the first 50 deleted words in each passage as well as they could. The method of deletion left residual

Either a "loose" or a "strict" test of reading comprehension followed each story. The loose tests consisted of about 10 short-answer questions based on significant points in the story or article. For example, a loose-test question used with an article on stained glass was "What is Grisaille work?" The strict tests concentrated on a single episode from the story or article (a subject did not know which episode would be used for this purpose). The episodes, which ranged from 100 to 500 words, were divided into consecutive content units as in the following example from "The Infernal Parliament" by Saki (content units are separated by slashes):

The Infernal Parliament

In an age when it has become increasingly difficult to accomplish anything new or original/ Bavton Bidderdale interested his generation by dying of a new disease./ "We always knew he would do something remarkable one of these days," observed his aunts;/ "he has justified our belief in him."/ But there is a section of humanity ever ready to refuse recognition to meritorious achievement/ and a large and influential school of doctors asserted their belief that Bidderdale was not really dead./ The funeral arrangements had to be held over until the matter was settled one way or the other,/ and the aunts went provisionally into half-mourning./

Meanwhile, Bidderdale remained in Hell as a guest/ pending his reception on a more regular footing./ "If you are not really supposed to be dead," said the authorities of that region, "we don't want to seem in an indecent hurry to grab you./ The theory that Hell is in serious need of population is a thing of the past./ Why, to take your family alone, there are any number of Bidderdales on our books, as you may discover later./

"It is part of our system that relations should be encouraged to live together down here./ From observations made in another world/ we have abundant evidence that it promotes the ends we have in view./ However, while you are a guest we should like you to be treated with every consideration and be shown anything that specially interests you. Of course, you would like to see our Parliament?" (Munro, 1930, pp. 620–621.)

There were from 15 to 30 content units per critical episode. A "Wh-question" was derived from each content unit. For example, the questions, "What is increasingly difficult to accomplish these days?" and "How did Bavton Bidderdale interest his generation?" were derived from the first two content units above. Questions were written on separate index cards so that the subjects could not use information contained in later questions to help them answer earlier ones. Loose and strict comprehension tests were scored in terms of the percentage of questions answered correctly.

The dictated words were selected randomly, without replacement, from the norms of Kučera and Francis (1967).

Design and procedure. The experiment lasted 14 weeks and consisted of four stages: *pretest, training, testing,* and *transfer.* One hour-long session was given each day, 5 days per week. (An academic vacation intervened between Sessions 23 and 24.) The subjects worked as a group, sitting at one table. They aided the experimenter by timing their own reading with stopwatches.

During three days of pretesting, the subjects first read two short stories and two encyclopedia articles (without any dictation) and answered loose comprehension questions on each text. They next read two new passages of each kind and answered the strict comprehension questions. Finally, they practiced taking dictation. For the dictation practice, the experimenter dictated two lists of 40 words each. As soon as all the subjects had finished writing a given word, the next word was dictated. The subjects wrote on plain paper, moving vertically down the page for each new word. When they reached the bottom of the page, they turned to a new sheet of paper and continued to write. The time taken to dictate each complete list was recorded. Throughout the experiment, each dictation list and each story or article was administered to all the subjects, although the subjects read the selections in different orders.

Group S–E read stories in the training stage, while group E–S read encyclopedia articles. Each subject knew what kind of material the others were reading and was told that everyone would have an opportunity to read both stories and articles as the experiment progressed. There were three reading trials in each daily session of the training stage. One was a control trial, in which the subjects read without any concurrent dictation. The other two were experimental trials, in which they read while taking dictation. To control for the differing difficulty of stories by different authors, all the stories that a subject read on any one day were by the same author. On experimental trials, dictation was given until all subjects had finished reading. After each trial, the subjects took a written loose test of reading comprehension. The order in which control and experimental trials were given was varied randomly from session to session.

The training stage ended for each subject when it appeared that he or she read equally well on experimental and control trials. To determine when this occurred, reading speeds and comprehension scores in both experimental and control trials were averaged for overlapping 5-day periods (Days 1–5,

---

information about the length of the deleted word. The mean number of words correctly replaced was 25.3 for the short stories and 20.9 for the encyclopedia selections, counting only exact replacements as correct. This significant difference, $t(9) = 5.23$, $p < .001$, indicates that the stories were indeed more redundant.

2–6, 3–7, etc.). The criterion was attained when, on five consecutive 5-day blocks, the reading speeds in the two conditions were within 15 words per minute (wpm) of each other and the comprehension scores were within 5% of each other.

The testing stage immediately followed the attainment of criterion. It was intended to confirm that the subjects were indeed reading as well in experimental as in control conditions, using strict tests of reading comprehension. On each of 5 days, subjects were given one experimental and one control trial, using the type of reading matter with which they had been trained. Each trial was followed by a strict comprehension test. A subject was considered to pass these tests if reading speed and comprehension were not significantly lower (by a randomization test, $p > .2$) on experimental than on control trials. Every subject who passed the stringent testing went on immediately to the transfer stage.

The transfer stage also lasted 5 days. Each session included one experimental and one control trial, each followed by a strict comprehension test. During this stage, however, subjects were given a new kind of reading matter: Those in group S–E now read encyclopedia articles, while those in group E–S now read stories. Every subject who passed the stringent testing (using the same criterion as for the testing stages) was dismissed from the experiment. Subjects who failed were given additional practice with the new type of material, using loose comprehension tests, until the criterion used in the testing stage was met. Five more days of strict comprehension testing followed.

Results

The rate at which words were dictated (hence, the rate at which subjects copied them) averaged 10.2 wpm in the pretest and 10.3 wpm in the three subsequent stages.

Figure 1 (a–g) exhibits the pretest data and the course of reading speed during the training phase for each subject individually. In the first few weeks of the training stage, all the subjects except Tom read much more slowly while taking dictation than while reading alone. The differences between experimental and control trials were assessed separately for each subject by one-tailed Wilcoxon tests on Trials 6–14 (the first 9 trials after the initial week). Experimental reading speeds were significantly lower ($p < .05$ or better) for every subject except Tom. Reading comprehension was also significantly lower on experimental trials for Tom and for Jeff. Thus, all subjects began

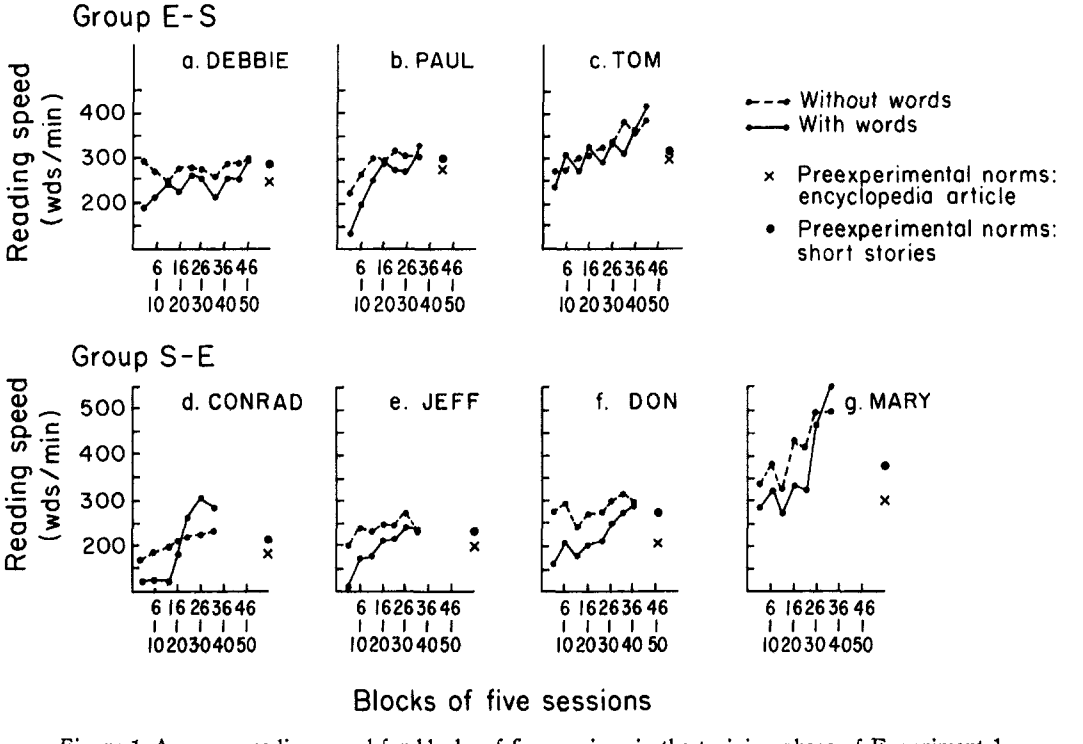

Figure 1. Average reading speed for blocks of five sessions in the training phase of Experiment 1.

Table 1
Reading Performance During the Testing Stage, Experiment 1

Group	Reading speed		Reading comprehension	
	Experimental trials	Control trials	Experimental trials	Control trials
S–E				
Don	297	329	63	69
Jeff	237	221	47	52
Conrad	225	204	55	51
Mary	400	424	59	59
E–S				
Paul	300	273	68	62
Tom	296	265	73	82
Debbie	285	261	68	63

Note. S = stories; E = encyclopedia.

by performing more poorly in the experimental condition.

This performance difference disappeared with practice; every subject achieved the criterion performance of equal reading speed and comprehension on five overlapping 5-day periods. Different subjects required differing amounts of practice before criterion was reached. The subjects trained on short stories reached criterion after the following number of days: Don, 43; Jeff, 36; Conrad, 36; Mary, 36. Those trained on encyclopedias reached criterion after the following number of days of training: Paul, 35; Tom, 45; Debbie, 50. Although the subjects trained with less redundant text took slightly longer, on the average, to learn to read while writing, random variation between subjects may account for this difference. The difference between groups is not significant ($p > .14$ by a one-tailed randomization test).

Results of the testing stage confirmed that every subject had learned to read as fast and with as full comprehension on experimental as on control trials. The mean reading speeds and comprehension scores for each subject appear in Table 1. Randomization tests failed to show any significant decrement in either reading speed or comprehension for any subject. By a two-tailed test, Tom's reading speed was somewhat greater under the experimental condition ($p < .10$).

Most of the subjects continued to read equally fast and comprehend equally well when they were transferred to the other type of reading matter (see Table 2, Group S–E and Group E–S). During the transfer stage, all the subjects comprehended equally well on experimental and control trials; all but Mary read with equal speed as well. Thus, three out of four subjects showed complete transfer from stories to encyclopedia articles and three out of three from encyclopedia articles to stories.

Mary practiced for 8 further days before reaching criterion on encyclopedia articles and undertaking another session of strict comprehension testing. Her mean speed and comprehension scores during the second period of stringent testing appear at the bottom of Table 2. Randomization tests revealed no differences between experimental and control conditions.

Discussion

This investigation replicated the principal finding of our first experiment (Spelke et al., 1976): With sufficient practice, people learn

Table 2
Reading Performance During the Transfer Stage, Experiment 1

Group	Reading speed		Reading comprehension	
	Experimental trials	Control trials	Experimental trials	Control trials
S–E				
Don	334	311	67	68
Conrad	243	228	51	41
Mary	245	305*	59	62
Jeff	237	227	57	45
E–S				
Debbie	280	293	83	82
Paul	327	343	54	48
Tom	305	287	58	51
Mary after more practice	282	286	63	67

Note. S = stories; E = encyclopedia.
* $p < .05$.

to write from dictation while reading with no loss of speed or comprehension. The conditions of training and the learning strategies of the individuals who have acquired this skill differed in various ways. The subjects of the present experiment began with markedly different normal reading speeds and levels of comprehension. Their speeds in the pretest varied from 229 wpm (Jeff) to 412 wpm (Mary); comprehension ranged from 50% (Jeff) to 78% (Tom). Three of them learned while reading texts of high redundancy and four with texts of low redundancy. Moreover, not all of the subjects appeared to pursue the same strategies during training. Don, Conrad, Mary, Paul, and Debbie maintained high levels of comprehension in the early stages of practice and sacrificed reading speed on the experimental trials in order to do so. Tom read equally fast on experimental and control trials at first, at the expense of reading comprehension. Jeff appeared to follow an intermediate strategy, reading somewhat more slowly and with somewhat less comprehension on experimental trials. Tom, Debbie, and Mary also seemed to shift back and forth between reading for speed and for comprehension during the practice phase, creating zigzag learning curves. Despite these differences, every subject succeeded in learning to read as well while writing as alone.

The experiment also indicated that the skill acquired by most subjects was general enough to transfer to a markedly different type of reading matter. Six of the seven subjects transferred perfectly from one type of material to the other; their strategies were appropriate to more than one kind of text.

Our results offer little support for the hypothesis that reading while writing requires an alternation of attention between the two tasks. The experiment tested two predictions stemming from the redundancy hypothesis: Learning to write while reading the short stories should take less time than learning to write while reading the encyclopedia, and learning to write while reading short stories should not transfer well to writing while reading less redundant material. The results were consistent with the first prediction, but they flatly contradict the second. We conclude that individuals need not rely on the redundancy of short stories in order to learn to read while writing.

Although the use of redundancy is not necessary, it is apparently possible. One of our subjects, Mary, seems to have used such a strategy. As a rapid reader, Mary must take advantage of the redundancy of prose in her ordinary reading activities. She may have used the same techniques to alternate successfully between writing and reading in this experiment. Since these techniques were less effective with the material from the encyclopedia, she required additional practice to master the second task.

In summary, a redundancy-based, time sharing hypothesis cannot be used to reconcile the results of Spelke et al. (1976) with the assumption of central limits on processing capacity. People need not borrow time from their reading to devote to the dictation task. At an early stage of practice, people may combine two activities by alternating attention between them. As they improve, however, they usually cease to do this. Experiment 2 examined a second way in which people might learn to read while writing: the hypothesis that the writing eventually ceases to draw on cognitive capacity and becomes automatic.

Experiment 2

It has often been suggested that there are two classes of mental activities: those that require attention and those that are automatic (e.g., James, 1890; Posner & Snyder, 1975; Solomons & Stein, 1896). The concept of automaticity has consistently attracted those who study the acquisition of skill (e.g., Bryan & Harter, 1899; LaBerge & Samuels, 1974; Posner & Keele, 1969; Schneider & Shiffrin, 1977; Welford, 1968), and many of them have proposed that repeated practice allows simple activities to be executed without attention.

Other thinkers, more impressed by the diversity and elusiveness of conscious experience, have found the concept of automaticity vague or unconvincing (e.g., Downey & Anderson, 1915; Norman & Bobrow, 1975;

Paulhan, 1887). To overcome such doubts, modern proponents of the concept have attempted to provide it with a more objective foundation. Various definitions have been proposed: An activity is automatic if it requires no decisions, or involves familiar and repetitive stimuli, or cannot be halted once its adequate stimulus has been registered, or causes no interference with other ongoing activities (cf. Posner & Snyder, 1975; Schneider & Shiffrin, 1977). This variety of definitions makes the automaticity hypothesis somewhat difficult to test. Fortunately, all versions of the concept seem to share a common assumption: that automatic activities must be relatively simple and routinized. This shared assumption makes it possible to test the hypothesis in a reading and writing experiment. If simultaneous reading and writing depends on automatization of the latter, it should be impossible to understand novel sentences that are dictated as one reads. As will be seen, the results of Experiment 2 disconfirmed this prediction. Although linguistic understanding is hardly a routine activity, our subjects understood the meanings of sentences and the relations between sentences while simultaneously reading an unrelated story.

Experiment 2 consisted of two stages, training and testing. Subjects had to master the dual task in the training stage before their understanding could be tested. The two successful subjects whose performance we describe here were not the first whom we had attempted to train. Experiment 2 followed an earlier effort, reported elsewhere (Spelke, Reaves, Hirst, & Neisser, Note 1), which we will call the "preliminary experiment." The two subjects of the preliminary experiment tried for 8 weeks to read while copying three-word sentences. Each sentence was dictated in a single burst. When the subjects had finished copying it, the next sentence was dictated. Reading speed, comprehension, and memory for the dictated material were tested. Unfortunately, the subjects made little progress. Thirteen more weeks were devoted to writing four-word sentences, writing variable-length sentences, and other tasks, but the subjects never read with normal speed and comprehension in any experimental condition. Many reasons for this failure can be suggested. The task is a difficult one, and we may have inadvertently selected poor subjects, or interpolated too many memory tests, or varied the procedure too often, or dictated meaningful material too early in the training period (see Spelke et al., Note 1, for details). In any case, we were unable to proceed to the testing stage in the preliminary experiment. In Experiment 2, however, two subjects were trained successfully (see below). Their understanding of the dictated material was then tested in three different ways.

One test of understanding was based on copying accuracy. It is well-known that listeners use the redundancy and meaning of a sentence to identify poorly articulated words (Miller, Heise, & Lichten, 1951) and to disambiguate lexical homonyms. This suggests that a subject who understands the dictated sentences should make fewer errors in copying them than in copying strings of unrelated randomly chosen words.

A second test was based on memory for the dictated material. People can remember verbal material much better if it consists of coherent sentences than if successive words are unrelated (Miller, 1956). Furthermore, recall of a sentence may be facilitated by cuing with a word from the sentence (Blumenthal, 1967); such a cue should be less effective in the recall of a string of unrelated words. A comparison of cued recall for the sentences and random strings that subjects had copied could therefore be used as another test of the subjects' understanding of those sentences.

Our most demanding test was based on the recognition of implications of the dictated material. Normal readers spontaneously combine the information in successive sentences; we attempted to determine whether our subjects would do this too. For this test, we dictated lists of statements that were semantically related, in a way which also implied the truth of certain other (never dictated) statements. Using the method of Bransford and Franks (1971), we presented the subjects with examples of dictated, implied, and unrelated sentences and asked each time whether they thought the sentence had

been on the dictation list. If they had integrated information across sentences, they should have found the implied sentences more familiar than the unrelated ones.

Method

Subjects. Two subjects who had served in Experiment 1, Tom and Mary, were paid to participate in further research after the conclusion of that experiment. Both began by practicing reading as they copied strings of three unrelated words (not sentences). As in the preliminary experiment, each string was dictated in a single burst. After 77 sessions, Tom had made little progress and resigned from the experiment. Mary continued to participate, mastered the task, and went on to the testing stage.

Meanwhile, we were conducting exploratory work with three new subjects on a new task: reading while copying numbers at dictation. They wrote Arabic numerals on some trials and the English names of the numbers on others. All the subjects except one experienced initial difficulty. That subject, Arlene, was able to read while copying numerals and number words with no practice at all. In a single additional week, she proved able to copy single random words while reading normally as well. Because of her evident skill with these reading/writing tasks, Arlene was enrolled in Experiment 2. At the time of the experiment, she was a second-year graduate student in developmental psychology. It may be important that she had previously worked as a secretary; she reported that during her secretarial career she had been able to type from copy while talking on the telephone.

Stimulus material. The subjects read short stories like those used in the high redundancy part of Experiment 1. Both loose and strict comprehension tests were used.

For dictation, sentences ranging in length from three to seven words were created ad hoc by the experimenters. Examples of three- and five-word sentences are "A rainbow appeared," "Doctors recommend aspirin," "The princess nodded her consent," and "A fire alarm went off." Strings of unrelated words were created by randomizing the words that made up the experimental sentences. The 30 N words in a list of 30 sentences of length N were scrambled, and new strings of length N were constructed. Examples of three- and five-word random strings are "Money rain scratch," "Is cows glad," "Spread the program daily she," and "The is married examined small."

Tests of integration of information across sentences were also prepared. In each of these "Bransford" (Bransford & Franks, 1971) tests, a list of 35 sentences was dictated. The first five were unrelated, while the rest comprised 10 sets of three sentences each. The three sentences in each triad described a single event or situation. One triad of three-word sentences was "Their house burned. Everything was destroyed. Firemen arrived late." An example of a triad of five-word sentences is "They had a baby girl. She was strong and healthy. Her father was very proud." The subsequent recognition test consisted of 30 randomly ordered sentences. Three test sentences were based on each triad: one had actually been presented (*dictated* sentence); one had not been presented but was implied by the situation described (*implied* sentence); one was not implied by that situation (*unrelated* sentence). All the test sentences, including the unrelated ones, were made up of words that had been dictated in the triad. The implied sentences for the above units were "Everything was burned" and "The baby girl was healthy." The corresponding unrelated sentences were "Firemen were burned" and "Her father was very strong." One complete dictation list and the corresponding recognition test appear in Table 3. Ten were constructed at each sentence length.

Design and procedure. Mary and Arlene practiced individually. A series of progressively harder tasks was used, each new task being introduced only after the subject had become proficient on the former one. The subjects began by reading while writing three-word random strings, then progressed to three-word sentences, then tried sentences varying in length from three to seven words, and finally were given five-word sentences. Comprehension of the reading material was tested frequently, but no tests of understanding or memory of the dictated sentences were given in the training stage.

Experimental and control trials were given throughout training. On experimental trials, the subjects wrote strings of words (or sentences) dictated by the experimenter. Each string was written horizontally across the page below its predecessor. A mixture of loose and strict comprehension tests was used. On days of loose comprehension testing, the subjects received one control trial and two to six experimental trials. On days of strict comprehension testing, they received one or two pairs of control and experimental trials. The two stories in a matched pair were always by the same author.

When a subject appeared to achieve normal reading speed and comprehension with a given type of material, she was given five days of strict comprehension testing. If the results indicated no significant difference between experimental and control conditions on either reading speed or comprehension (i.e., if $p > 0.20$ by randomization tests), she resumed training with a more difficult type of dictated material. If performance did not meet that criterion, she either remained with the same dictation material or returned to something easier. Decisions to return to easier material were based on the judgment of the subject and the experimenters. Mary reached criterion on three-word random strings by Session 52 and on three-word sentences by Session 71. She then practiced with

Table 3
Example of Dictated Material and Test List of Dictated (D), Implied (I), and Unrelated (U) Sentences for the Semantic Integration Test, Experiment 2

Recognition list	Test list
The rope broke	(D) Father chased him
Spot got free	(I) Spot's rope broke
Father chased him	(U) Spot chased Father
Tourists were visiting	(D) Tourists were visiting
They took pictures	(I) Tourists took pictures
The natives left	(U) The tourists left
Cookbooks contain recipes	(D) Cookbooks contain recipes
Susan owns several	(I) Susan owns cookbooks
Mary hates cooking	(U) Susan hates cooking
The Indians scouted	(D) They found trails
They found trails	(I) Scouts found trails
The paths diverged	(U) The Indians diverged
The princess arrived	(D) The guests bowed
She nodded politely	(I) The princess nodded
The guests bowed	(U) The princess bowed
It is snowing	(D) It's falling hard
It's falling hard	(I) It's snowing hard
It's not raining	(U) It's raining hard
The stores closed	(D) The stores closed
They went bankrupt	(I) Stores went bankrupt
The owners moved	(U) The stores moved
Police raided Minsky's	(D) Gangsters frequently visited
It was illegal	(I) Minsky's was illegal
Gangsters frequently visited	(U) Gangsters raided Minsky's
Paperweights are glass	(D) The desk rolled
The desk rolled	(I) The paperweights shattered
They got shattered	(U) The paperweights rolled
The dancers performed	(D) The director bowed
They were excellent	(I) Dancers were excellent
The director bowed	(U) The dancers bowed

sentences of varying unpredictable length (three to seven words), but had not reached criterion by Session 97. At this point training was shifted to five-word sentences, on which she reached criterion by Session 114. Training was then terminated, and testing began. Five-word sentences were used in the testing stage.

Arlene moved successfully through three-word random strings, three-word sentences, and variable-length sentences, reaching criterion on the latter by Session 38. She was then trained with five-word sentences exclusively, but failed to reach criterion with these by Session 52. Since it was necessary to begin the testing stage at that time, Arlene was tested with three-word sentences (which she had already mastered). A full chronology of the training stage appears in Table 4.

After the completion of training, Mary and Arlene proceeded to the tests of understanding described earlier. Just as in the training stage, they received experimental and control trials. Experimental trials were followed by either a strict test of reading comprehension or a test of understanding of the dictated material. These tests of understanding were presented in Phase 1 and Phase 3 of testing. Phase 1 consisted of the "Bransford" tests of semantic integration. Phase 2 was a simple comparison of reading while copying sentences with reading while copying random strings. Phase 3 included the tests of copying accuracy and recall. Throughout these tests, five-word sentences were dictated to Mary and three-word sentences were dictated to Arlene. A chronology of the testing stage appears in Table 5.

In Phase 1, the special lists of semantically related sentence triads were dictated. Reading was interrupted after the subject had copied the last sentence of the last triad. She immediately took the written test of recognition memory, which included dictated, implied, and unrelated sentences. She was asked to mark a sentence as "old" if she thought that it had occurred in verbatim form on

Table 4
Reading Speed and Comprehension During Training, Experiment 2

Sessions	Dictated material	Reading performance: Control Speed	Control Comprehension	Experimental Speed	Experimental Comprehension	Randomization test results: Speed ($p >$)	Comprehension ($p >$)
Mary							
1–52	3-word random strings	428	.52	426	.61	.20	.20
53–71	3-word sentences	338	.56	358	.68	.20	.20
72–97	3–7-word sentences	370	.81	345	.72	.18	.18
98–114	5-word sentences	317	.65	370	.63	.20	.20
Arlene							
1–5	3-digit Arabic numerals	479	.70	505	.67	.20	.20
6–10	3-digit numbers (written)	565	.71	545	.78	.20	.20
11–15	single words	517	.78	527	.73	.20	.20
16–23	3-word random strings	507	.66	517	.73	.20	.20
24–28	3-word sentences	529	.55	534	.58	.20	.20
29–42	3–7-word sentences	474	.45	496	.60	.20	.20
43–59	5-word sentences	611	.84	504	.76	.10	.18
60–62	3-word sentences*	582	.60	553	.69	.20	.20

Note. Each score is the mean of the last five trials with each type of material.

the previously dictated list and to mark it "new" otherwise. She also indicated her confidence in each judgment on a 3-point scale.

The first recognition test came as a surprise. The subjects had never before been asked to remember the dictated material and were not warned that such a test might be given. Thereafter, they were informed that tests of recognition memory would sometimes follow a dictation list. Nevertheless, they were asked to read in their normal manner, making no specific effort to remember the dictated material. Tests of semantic integration were given for nine more sessions, each including three trials. Altogether, the 30 trials of Phase 1 included 10 control trials, 10 experimental trials followed by the usual strict comprehension test on the story, and 10 experimental trials followed by the recognition test of semantic integration. These were given in an unsystematic order, and the subjects were not told whether any given dictation trial would be followed by a comprehension or a recognition test. The semantic integration test was not described as such, and the triad structure of its dictation list was never pointed out. At the end of the 10 days of recognition testing, the subjects were debriefed and asked whether they had noticed that half the dictation lists were made up of coherent triads of sentences. Neither subject reported noticing this; both were surprised to learn that successive dictated sentences had even been related.

A direct comparison of the subjects' reading

Table 5
Reading Speed and Comprehension During Testing of Comprehension for the Dictated Material

Sessions	Test	Reading performance: Control Speed	Control Comprehension	Experimental Speed	Experimental Comprehension
Mary					
115–124	Phase 1	315	.70	324	.58
125–135	Phase 2	332	.70	336	.64
136–140	Phase 3	295	.73	297	.71
Arlene					
63–72	Phase 1	581	.56	573	.63
73–83	Phase 2	580	.62	521	.62
84–89	Phase 3	551	.82	444	.93

performance while writing sentences and random strings was undertaken for the next 11 sessions, which constituted Phase 2. Control trials, experimental trials with sentences, and experimental trials with random strings were each followed by strict tests of reading comprehension.

The last five sessions of the experiment comprised Phase 3, which focused on copying accuracy and on cued recall. These sessions consisted of control trials, experimental trials with strict reading comprehension tests, and experimental trials with dictation recall tests. Each session lasted 2 hours, during which 5 and 10 stories were read. Mary and Arlene read 25 stories during this phase, 5 by each of five authors. For each set of 5 stories, they received one control reading trial, two experimental trials with sentences (one followed by a reading comprehension test and one by a dictation recall test), and two experimental trials with random strings (one followed by each type of test). They did not know which kind of test would follow each trial. They were encouraged to read normally throughout these sessions, making no special effort to remember what they wrote.

On the dictation recall trials, reading was interrupted after 30 items had been dictated. The subject was then given the first noun of each sentence or string as an (oral) recall cue and was asked to recall as many words from that sentence or string as she could. The cues were given in random order.

For Mary, Phase 3 occurred immediately after Phase 2. In Arlene's case, however, there was a 3-month hiatus between them. At the beginning of Phase 3, therefore, Arlene was retrained to criterion. This took two sessions, during which she received unstructured practice followed by 2 trials with strict comprehension tests. One further variation was undertaken with Arlene. Before the 25 trials similar to those given to Mary, Arlene was given 10 "control recall" trials in which she copied and recalled sentences (5 trials) or random strings (5 trials) without simultaneous reading. She knew that memory tests would follow these control trials.

Control experiment. Four untrained subjects participated in a control experiment to establish baseline norms for performance on the semantic integration and copying accuracy tasks. They wrote without simultaneous reading for a total of eight sessions. During the first three sessions, they simply copied sentences and random strings at dictation. For the next five sessions, they received the semantic integration tests in the same order and with the same instructions as Mary and Arlene, except that they were not asked to read concurrently. An equal number of trials on which unrelated sentences were dictated were also given, as they had been for Mary and Arlene. The subjects were told they were participating in a control condition for a reading and writing experiment and that the variables of interest were spelling and penmanship. Three Cornell summer school students and one prospective graduate student in psychology were paid to participate in the control experiment.

Results

Copying rate. The rate at which sentences were dictated (hence, copied) depended on their length. During training, the dictation rate for three-word random strings averaged 22 wpm for Mary and 24 wpm for Arlene. Three-word sentence dictation proceeded at average rates of 20 wpm (Mary) and 25 wpm (Arlene). Variable-length sentences (three to seven words) were dictated at rates of 25 wpm (Mary) and 27 wpm (Arlene). Five-word sentences were dictated at rates of 29 wpm (Mary) and 27 wpm (Arlene). During testing, Mary copied five-word sentences at the rate of 31 wpm, and Arlene copied three-word sentences at the rate of 24 wpm.

Reading performance. Initially, Mary found it much more difficult to read while copying three-word strings than to read while copying single random words. Fifty-two sessions were required before Mary achieved criterion with random-string dictation. This was much longer than her total training time had been with single words. When she finally achieved equal performance in the experimental and control conditions, however, her speed was at the same level as she had attained with single words, 400 wpm (see Table 4). Arlene, who had required no practice to learn to read while copying single words, also required no practice to learn to read while copying three-word strings.

When the dictation material was changed from random strings to sentences, Mary required 14 additional trials before the experimenters judged that strict comprehension testing could be given. Strict comprehension testing revealed that criterion was achieved on the last five trials, although reading speed on both experimental and control trials was 70–90 wpm below that attained with random strings and continued at this new level throughout the remainder of the experiment. Arlene, on the other hand, was able to copy three-word sentences from dictation without additional practice. Her

reading speeds and comprehension scores were at levels comparable to those attained with random strings: Speed was a bit higher and comprehension lower (see Table 4). Criterion was met on the first five sessions.

As Table 4 reveals, both subjects experienced some difficulty with sentences of variable length. Arlene eventually achieved criterion. Her overall speeds and comprehension scores appeared to be declining, however, and she reported that the task was difficult and unpleasant. Mary did not achieve criterion in 26 practice sessions. Both subjects were then shifted to five-word sentences. Mary achieved criterion in 17 sessions. Arlene did not reach criterion, and she expressed a wish to conclude the experiment. She was therefore returned to three-word sentences for the testing stage.

Comparisons of reading performance on experimental and control trials were continued during the testing stage. They revealed that the tests of understanding had little effect on reading performance. Table 5 contains the relevant data. Mary and Arlene each were given 10 control trials and 10 experimental trials with strict comprehension tests during Phase 1, when the tests of semantic integration were being administered. Randomization tests on their reading speeds and comprehension scores confirmed that they were still at criterion: Each subject read as quickly and comprehended as well on the experimental as on the control trials (all $ps > .20$). The introduction of the recognition tests did not disrupt the skill of reading while writing.

Each subject was given 11 more control trials and 11 more sentence-dictation trials with strict comprehension tests during Phase 2, in addition to the trials on which random strings were dictated. The addition of the random-string trials did not seem to disrupt Mary's performance on the sentence trials; she was still at criterion when control and sentence-dictation trials were compared. However, Arlene now read more slowly on sentence-dictation than on control trials, although the difference was not significant ($p > .20$). The introduction of random-string dictation trials may have caused a disruption in Arlene's reading.

Table 6
Reading Speed and Comprehension During Dictation of Sentences and Random Strings, Experiment 2, Phase 2

	Sentence dictation		Random-string dictation	
Subject	Speed	Comprehension	Speed	Comprehension
Mary	328	.66	321	.71
Arlene	521	.62	562	.58

Finally, Mary and Arlene each were given five control and five sentence-dictation trials with strict comprehension tests during Phase 3, when copying accuracy and recall were being tested. Mary's reading speed and comprehension in these conditions did not differ (both $ps > .20$). Arlene tended to read more slowly on experimental trials ($p < .10$); and she also displayed slightly higher comprehension ($p > .20$). It appeared, therefore, that continued dictation of random strings and the presence of recall tests on some trials did not greatly disrupt the reading-while-writing performance of either subject.

It is worth examining the possibility that Mary and Arlene may have used the redundancy of the dictated sentences to facilitate their reading and writing. In that case, their reading performance might have been expected to decline when they copied random strings. Table 6 shows that this did not occur. There were no differences in Mary's reading speed or comprehension between sentence dictation and random-string dictation trials in Phase 2 (by randomization test, both $ps > .20$). Arlene actually read more slowly on sentence-dictation trials than on random-string dictation trials, but comprehended slightly more. Her greater reading speed during random-string dictation was not significant ($p > .12$). Just as most of the subjects of Experiment 1 had not used the redundancy in the reading matter to improve their reading performance, the subjects of Experiment 2 did not appear to use the redundancy in the dictation material for that purpose.

Copying accuracy test. Table 7 presents

Table 7
Number of Copying Errors With and Without Concurrent Reading, Sentence, and Random-String Dictation, Experiment 2, Phase 3

Subject	No. words per string	With reading		Without reading	
		Sentence dictation	Random-string dictation	Sentence dictation	Random-string dictation
Trained					
Arlene	3	27	58	7	26
Mary	5	8	55	—[a]	—[a]
Control					
Steve	3	—	—	22	43
Andy	3	—	—	1	12
Rita	5	—	—	9	171
Robin	5	—	—	4	59

[a] At the point in the study when this control was run, Mary had left Cornell.

the frequency of copying errors for Mary, Arlene, and the four control subjects. Under normal conditions, there are fewer copying errors with sentence dictation than with random-string dictation (for each control subject, $p < .002$, binomial test). The same difference appears on experimental trials for both Arlene and Mary (each $p < .001$, binomial test). Indeed, Arlene and Mary were just about as accurate as the control subjects who were not reading, even though copying accuracy had been stressed in the instructions to the controls. (However, Arlene herself was still more accurate when she was not reading; this difference was significant, $p < .001$, for both types of dictated material.) Arlene and Mary evidently used the structure or meaning of the dictated sentences to disambiguate homonymous or imperfectly heard words, even while they were simultaneously reading with full speed and comprehension.

Table 8
Number of Words Recalled, Experiment 2, Phase 3

Subject	Sentence dictation	Random-string dictation
With reading		
Arlene	49	7
Mary	30	3
Without reading		
Arlene	98	32

Cued recall test. Results of the tests of cued recall are given in Table 8. Both subjects recalled more words on sentence-dictation trials than on random-string trials; $\chi^2(4) = 21.7$, $p < .001$, for Mary, and $\chi^2(4) = 31.5$, $p < .001$, for Arlene. The same pattern of results appeared on Arlene's control trials, when she copied and recalled sentences without reading. There, too, her cued recall scores were higher on sentences than on random strings, $\chi^2(4) = 39.0$, $p < .001$. Although Arlene's overall recall with either procedure was higher on control trials than when she had been reading simultaneously (both χ^2s: $p < .001$), the relative advantage of sentences over random strings was no higher on the control trials. These results again indicate that Mary and Arlene were sensitive to the syntax or meaning of the dictated sentences.

Whenever Arlene recalled one of the two to-be-remembered words in a three-word sentence, she was likely to recall the other word as well. On experimental trials, Arlene recalled a total of 20 sentences fully and 9 partially; on control trials, she recalled 40 sentences fully and 18 partially. Mary, who was dealing with five-word sentences, rarely recalled them in their entirety. She recalled a total of 2 sentences fully and 10 partially.

Semantic integration test. The subjects'

judgments that a test sentence in the semantic integration test was "old" or "new," together with their confidence ratings of those judgments, were converted into a 6-point scale. A 6 on this scale refers to a very confident "old" judgment, a 1 to a very confident "new" judgment. Figure 2, Graph a shows the recognition scale scores for the first day of testing; Figure 2, Graph b shows the average scores for all 10 days. The results in the left panel of each graph show that the four control subjects did indeed integrate information across successive sentences. Not only were dictated sentences judged as more familiar than implied ones but implied sentences seemed more familiar than unrelated ones. Like the subjects of Bransford and Franks (1971), our controls demonstrated memory for ideas that had been implied across sentences but never explicitly stated. The difference between dictated and implied sentences was in the predicted direction for all four control subjects and individually significant at the .001 level for three of them. The difference between implied and unrelated sentences was also in the predicted direction for all four subjects; it reached the .001 level for two of them and the .02 level in a third (individual Friedman tests). All the subjects noticed the triadic structure of the dictation lists.

Although Mary and Arlene were reading stories as they copied the sentences, they exhibited the same pattern. The dictated sentences seemed most familiar and the unrelated ones least familiar, with the implied sentences in between. This pattern appeared in both subjects on the first day of testing (Figure 2, Graph a), when the subjects could not have expected the dictation list to have a triadic structure, as well as in all 10 test days taken together (Figure 2, Graph b). Taken over the 10 days, the differences between the three types of test sentences are clearly significant. For Arlene, the overall $\chi^2_r(3, 10) = 31.26$, $p < .001$ (Friedman test). Separate tests showed that the difference between dictated and implied sentences and that between implied and unrelated sentences were significant, with $\chi^2_r(2, 10) \geq 6.25$, $p < .02$. For Mary the test across all three sentence types was also significant, $\chi^2_r(3, 10) = 10.75$, $p < .01$; the difference between dictated and implied sentences itself was marginally significant, $\chi^2_r(2, 10) = 3.61$, $p < .10$, and that between implied and unrelated sentences clearly significant, $\chi^2_r(2, 10) = 5.29$, $p < .02$.

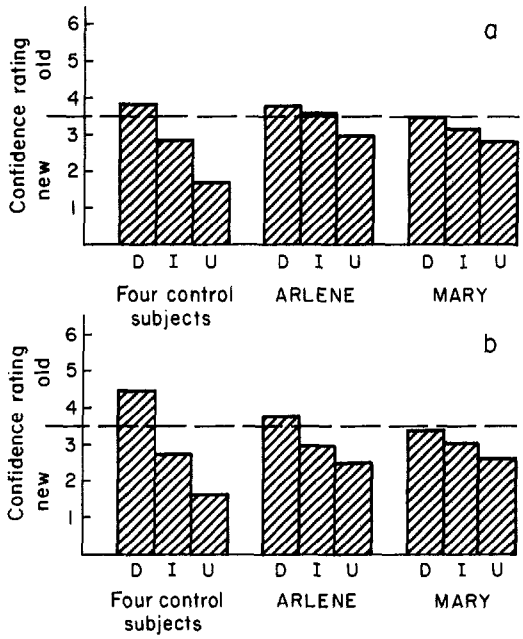

Figure 2. Average confidence rating for the control and experimental subjects for the dictated (D), implied (I), and unrelated (U) sentences: (a) the results from the first semantic integration test; (b) the average over all 10 semantic integration tests.

Although Mary and Arlene showed the same pattern of results as the control subjects, the differences between their ratings of the various types of sentences were not as marked. In part, this was because they were less willing than the controls to use extreme values on the confidence rating scale. This trend is clearly evident in Figure 3, which shows the frequencies with which the six points of the scale were used by experimental and control subjects. When the controls called a sentence "new" or "old" they were fairly sure of themselves. Arlene and Mary, however, found most of the sentences somewhat unfamiliar and rarely exhibited much

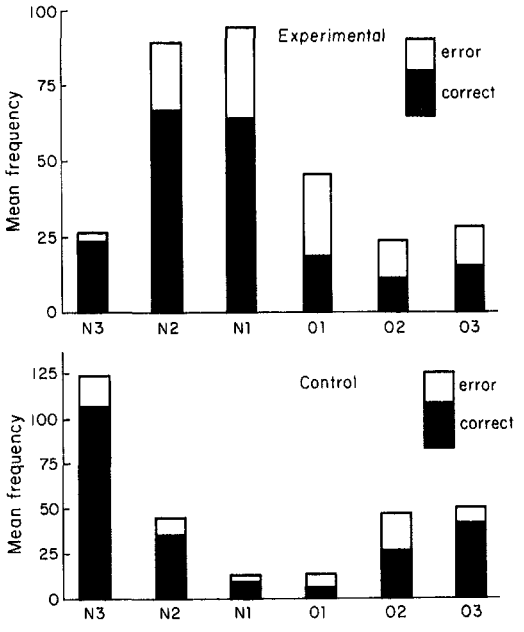

Figure 3. Mean frequency of response of the experimental (Mary and Arlene) and control subjects for 10 semantic integration tests: whether they judged the sentence new (N) or old (O), were correct or in error, and their confidence ratings (1–3).

confidence in their judgments. Moreover, as noted earlier, neither of them noticed the triadic structure of the dictated lists.

Discussion

The results of the testing stage provide clear evidence that Arlene and Mary were sensitive to the structure and meaning of the sentences they copied. This is somewhat surprising, since they had never been asked to understand or remember any dictated material; the subjects of Spelke et al. (1976) had not noted the meanings of dictated words until they were asked to do so. The data of Phase 3 suggest a possible reason for this difference. Understanding the meanings of sentences makes it possible to copy them more accurately; understanding the meanings of single words may not convey any corresponding advantage.

Although there is no doubt that our subjects understood what they were writing as well as what they were reading, their understanding seemed less than complete under these conditions. Arlene herself recalled fewer words from experimental trials than from the special control trials in which she copied without reading. (This difference, however, may reflect only the superiority of intentional over incidental learning: She knew that a recall test would follow every control trial, but not whether such a test would follow any particular experimental trial.) Moreover, Mary and Arlene both expressed lower confidence in their judgments on the semantic integration test than the controls did, and neither of them reported any awareness of the semantic links between successive sentences. Their performance was not automatic in any objective sense—they understood what they were writing—but it apparently involved a minimum of conscious monitoring.

In summary, the hypothesis that skilled activities become automatic cannot be used to rescue the notion of limited capacity. Understanding new sentences and integrating ideas are just the kinds of achievements for which a central processor is intended. They could not be automatic according to any plausible version of this hypothesis. Mary's and Arlene's writing was automatic, however, in some sense: Most of their feats were not accessible to their own awareness. The hypothesis that skilled activities become automatic may not be entirely wrong.

General Discussion

The difficulty of dividing one's attention does not stem simply from the fixed properties of some central mechanism nor from the limited size of a pool of processing resources. Although it is difficult at first to do two things at once—or one thing well, for that matter—people can learn to do indefinitely many things indefinitely well. So far as we know now, performance is constrained only by obvious peripheral factors (e.g., the mechanics of the effectors, visual acuity, the articulation of joints, the mass of body parts) and by an individual's willingness to practice.

Our results specifically rule out two versions of the limited capacity hypothesis as

accounts of the skill of reading while writing. First, skilled subjects do not simply direct their attention to the dictated material during moments when the reading matter is redundant. Decreases in the level of text redundancy did not hamper the performance of three trained subjects in Experiment 1. To be sure, this finding does not completely rule out the time sharing hypothesis: One might argue that time sharing is too rapid to detect and so efficient that it permits thorough performance of each component task. This version of the hypothesis would render it untestable and, we think, useless. If people can switch their attention among indefinitely many activities with unrestricted speed, the concept of limited capacity will not help us understand the constraints on human performance.

Although no realistic concept of time sharing can explain the performance of our skilled subjects, we do not want to deny that time sharing happens. Subjects often report that they alternate attention in the early stages of practice. These alternations may occur simply because at that point, the subjects have not yet learned to combine writing and reading and nevertheless must carry out the dual task somehow. It is also possible, however, that alternation actually facilitates the acquisition of skill, perhaps by allowing each activity to be coordinated with the other. This hypothesis would explain why the subjects trained with short stories reached criterion a little more quickly than those trained with encyclopedia articles, if that finding is not just a matter of random sampling variation. The short story subjects may have been better able to alternate the tasks at a stage of practice when it was useful to do so.

A second version of the limited capacity hypothesis, the notion of automaticity, also fails to explain our results. Individuals need not write automatically, without understanding, in order to read at the same time. To be sure, skilled writing while reading is not the same as writing alone: The subjects misspelled words more, remembered them less well, and were less aware of relations between sentences when they were reading than when they were not. They remained able to extract the meaning of the dictated material, however, and could integrate semantic information over a series of novel sentences. No theory based on a concept of capacity can easily claim that such activities are automatic.

The introspective reports of subjects in these experiments suggest that at certain stages of practice (perhaps intermediate stages), one loses much awareness of what one writes. Diane and John, by their own report, were oblivious to semantic relations among dictated words at one stage of our first study (Spelke et al., 1976). Mary and Arlene reported no awareness of the semantic relations among successive sentences, although they used those relations to integrate information. The meaning of these reports, however, is as elusive today as it was in the laboratory of Solomons and Stein (1896). The subjects' introspections may reflect their expectations about the material they wrote rather than the manner in which they processed it. Diane and John, for example, had been given only unrelated words to copy at the time that semantically related lists were first dictated. They may have failed to report those relations because they did not expect them; those relations might indeed have been detected in ways for which we did not test. Similarly, Mary and Arlene were given only unconnected sentences to write during the training stage; they were surely not prepared for the semantically related sentence triads that followed. The fact that they integrated information across the related sentences, while reporting no awareness of the relations, suggests that their introspections may have been shaped by their expectations about what they were doing and not by what they actually did.

If alternation or automaticity does not underlie divided attention skills, what does make them possible? In our opinion, actions change qualitatively when they are practiced. A skilled individual has learned to detect new stimulus constellations and execute new patterns of action, not just to do old things intermittently or unconsciously. The experienced bird-watcher who scans treetops for a woodpecker is not automatically processing

the same feature that she once examined in a conscious way; the commands that a skilled typist issues automatically to his fingers are not the same as those that governed his behavior as a novice. Component perceptual or motor processes are changed by becoming embedded in larger schemes and may lose their independent existence entirely. When an activity ceases to exist in its original form, one may cease to experience it and may describe it as automatic. Such practiced actions, however, are no less complex, subtle, or attended than unpracticed ones.

The study of the acquisition of skills of divided attention has hardly begun. We do not yet understand why some activities are easy to combine while others are so difficult. We do not know what happens during practice to make divided attention possible. We do not know if concepts like the alternation of activities, or their automaticity, will figure in any useful principles of skill acquisition. Perhaps some answer to these questions will emerge from further studies of practice effects in divided attention tasks. When we know something about the changes that tasks undergo when one combines them, and about what makes such combinations easy or difficult, we will be in a better position to theorize about the nature and limits of human attention.

Reference Note

1. Spelke, E., Reaves, C., Hirst, W., & Neisser, U. *Skills of divided attention II: A sentence copying experiment.* Unpublished manuscript, 1977. (Available from E. Spelke, Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania 19174.)

References

Blumenthal, A. L. Prompted recall of sentence. *Journal of Verbal Learning and Verbal Behavior*, 1967, *6*, 203–206.

Bransford, J. D., & Franks, J. J. The abstraction of linguistic ideas. *Cognitive Psychology*, 1971, *2*, 331–350.

Broadbent, D. E. The role of auditory localization in attention and memory span. *Journal of Experimental Psychology*, 1954, *47*, 191–196.

Broadbent, D. E. *Decision and stress.* London: Academic Press, 1971.

Brown, I. D., & Poulton, E. C. Measuring the spare "mental capacity" of car drivers by a subsidiary task. *Ergonomics*, 1961, *4*, 35–40.

Bryan, W. L., & Harter, N. Studies on the telegraphic language. *Psychological Review*, 1899, *6*, 345–375.

Carroll, J. B. Defining language comprehension: Some speculations. In J. B. Carroll & R. O. Freedle (Eds.), *Language, comprehension and the acquisition of knowledge.* Washington, D.C.: V. H. Winston, 1972.

Downey, J. E., & Anderson, J. E. Automatic writing. *American Journal of Psychology*, 1915, *26*, 161–195.

Encyclopedia Britannica. Chicago: Britannica, 1922.

International encyclopedia of the social sciences. New York: Macmillan and Free Press, 1966.

James, W. *The principles of psychology* (Vol. 1). New York: Holt, 1890.

Kahneman, D. *Attention and effort.* Englewood Cliffs, N. J.: Prentice-Hall, 1973.

Kučera, H., & Francis, W. N. *Computational analysis of present-day American English.* Providence, R.I.: Brown University Press, 1967.

LaBerge, D., & Samuels, S. J. Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 1974, *6*, 293–323.

Marks, L., & Miller, G. A. The role of semantic and syntactic constraints in the memorization of English sentences. *Journal of Verbal Learning and Verbal Behavior*, 1964, *3*, 1–5.

Messerschmidt, R. A quantitative investigation of the alleged independent operation of conscious and subconscious processes. *Journal of Abnormal and Social Psychology*, 1927, *22*, 325–340.

Miller, G. A. The magic number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 1956, *63*, 81–97.

Miller, G. A., Bruner, J. S., & Postman, L. Familiarity of letter sequences and tachistoscopic identification. *Journal of General Psychology*, 1954, *50*, 129–139.

Miller, G. A., Heise, G. A., & Lichten, W. The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology*, 1951, *41*, 329–335.

Miller, G. A., & Isard, S. Some perceptual consequences of linguistic rules. *Journal of Verbal Learning and Verbal Behavior*, 1963, *2*, 217–228.

Moray, N. Broadbent's filter theory: Postulate H and the problem of switching time. *Quarterly Journal of Experimental Psychology*, 1960, *12*, 214–220.

Moray, N., Fitter, M., Ostry, D., Faveau, D., & Nagy, V. Attention to pure tones. *Quarterly Journal of Experimental Psychology*, 1976, *28*, 271–283.

Munro, H. H. The infernal parliament. In H. H. Munro, *The complete stories of Saki.* New York: Viking Press, 1930.

Neisser, U., & Becklen, R. Selective looking: At-

tending to visually specified events. *Cognitive Psychology*, 1975, *7*, 480–494.

Norman, D. A., & Bobrow, D. G. On data-limited and resource-limited processes. *Cognitive Psychology*, 1975, *7*(1), 44–64.

Ostry, D., Moray, N., & Marks, G. Attention, practice and semantic targets. *Journal of Experimental Psychology: Human Perception and Performance*, 1976, *2*, 326–336.

Paulhan, F. La simultanéité des actes psychiques. *Revue Scientifique*, 1887, *13*, 684–689.

Posner, M. I., & Keele, S. W. Attention demands of movements. In *Proceedings of the 16th International Congress of Applied Psychology, Symposium on Work and Fatigue, August (1968)*. Amsterdam: Swet & Zeitlinger, 1969.

Posner, M. I., & Snyder, C. R. R. Attention and cognitive control. In R. Solso (Ed.), *Information processing and cognition: The Loyola Symposium*. Potomac, Md.: Erlbaum, 1975.

Ryder, H. W., Carr, H. J., & Herget, P. Future performance in footracing. *Scientific American*, 1976, *234*(6), 109–119.

Schneider, G. E., & Shiffrin, R. M. Controlled and automatic human information processing: I. Detection, search and attention. *Psychological Review*, 1977, *84*, 1–66.

Shiffrin, R. M. The locus and role of attention in memory systems. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance V*. New York: Academic Press, 1975.

Solomons, L., & Stein, G. Normal motor automatism. *Psychological Review*, 1896, *3*, 492–512.

Spelke, E. S., Hirst, W. C., & Neisser, U. Skills of divided attention. *Cognition*, 1976, *4*, 215–230.

Stevenson, J. H. Effect of posthypnotic dissociation on the performance of interfering tasks. *Journal of Abnormal Psychology*, 1976, *85*, 398–407.

Treisman, A. The effect of irrelevant material on the efficiency of selective listening. *American Journal of Psychology*, 1964, *77*, 533–546.

Treisman, A., & Geffen, G. Selective attention and cerebral dominance in perceiving and responding to speech messages. *Quarterly Journal of Experimental Psychology*, 1967, *19*, 1–17.

Underwood, G. Moray vs. the rest: The effect of extended shadowing practice. *Quarterly Journal of Experimental Psychology*, 1974, *26*, 368–372.

Welford, A. T. *Fundamentals of skill*. London: Methuen, 1968.

Received October 13, 1978 ■