AMERICAN SCIENTIST, *57*, **4**, pp. 421-457, 1969

MEMORY-SCANNING: MENTAL PROCESSES REVEALED BY REACTION-TIME EXPERIMENTS[1]

By SAUL STERNBERG

ONE OF THE oldest ideas in experimental psychology is that the time between stimulus and response is occupied by a train of processes or *stages*—some being mental operations—which are so arranged that one process does not begin until the preceding one has ended. This *stage theory* implies that the reaction-time (RT) is a *sum*, composed of the durations of the stages in the series, and suggests that if one could determine the component times that add together to make up the RT, one might then be able to answer interesting questions about mental operations to which they correspond. The study of RT should therefore prove helpful to an understanding of the structure of mental activity.

The use of results from RT experiments to study stages of information processing began about a century ago with a paper, "On the Speed of Mental Processes," by F. C. Donders (1868). It was in this paper that Donders introduced the *subtraction method*—a method for analyzing the RT into its components and thereby studying the corresponding stages of processing.

1. Decomposing RT by the Subtraction Method

To use the subtraction method one constructs two different tasks in which RT can be measured, where the second task is thought to require all the mental operations of the first, plus an additional inserted operation. The difference between mean RTs in the two tasks is interpreted as an estimate of the duration of the inserted stage, as shown in Fig. 1. This interpretation depends on the validity of both the stage theory and an *assumption of pure insertion* which states that changing from Task 1 to Task 2 merely inserts a new processing stage without altering the others.

For example, Wundt (1880, pp. 247–260) developed an application in which RTs were measured when a subject had to respond after he had identified a stimulus, and also when he had to respond after merely

[1] Most of the research reported in this paper was supported by Bell Telephone Laboratories and conducted in its Behavioral and Statistical Research Center at Murray Hill, N.J. The work reported as Exp. 4 was done in collaboration with A. M. Treisman of the University of Oxford. I am grateful to C. S. Harris, T. K. Landauer, H. Rouanet, and R. Teghtsoonian for helpful criticisms of the manuscript, and to L. D. Harmon for discussion leading to Exp. 5. R. E. Main assisted with Exps. 4 and 5, B. Barkow with Exps. 7 and 8, and B. A. Nasto with Exps. 4, 5, 7, and 8.

detecting its presence. The difference was used as an estimate of the identification time. In this instance the stages shown in Fig. 1 might be (*a*) stimulus detection, (*b*) stimulus identification, and (*c*) response organization. In an earlier application, Donders (1868) had compared mean RTs in a simple-reaction task (one stimulus and response) and a choice-reaction task (multiple stimuli and responses); he regarded the difference as the duration of the stages of stimulus discrimination and response selection.

This kind of enterprise occupied many psychologists during the last quarter of the nineteenth century. Much of their work was summarized by J. Jastrow (1890) in a popular treatise on *The Time Relations of Mental Phenomena*.

Around the turn of the century the subtraction method became the subject of criticism for two main reasons. First, the differences in mean RT that were observed in some applications varied excessively from subject to subject, and from laboratory to laboratory. In retrospect, this seems to have been caused by the use of tasks and instructions that left the subject's choice of "processing strategy" relatively uncontrolled.[2] Second, introspective reports put into question the assumption of pure insertion, by suggesting that when the task was changed to insert a stage, other stages might also be altered. (For example, it was felt that changes in stimulus-processing requirements might also alter a response-organization stage.) If so, the difference between RTs could not be identified as the duration of the inserted stage. Because of these difficulties, Külpe, among others, urged caution in the interpretation of results from the subtraction method (1895, Secs. 69, 70). But it appears that no tests other than introspection were proposed for distinguishing valid from invalid applications of the method.

A stronger stand was taken in later secondary sources. For example, in a section on the "discarding of the subtraction method" in his *Experimental Psychology* (1938, p. 309), R. S. Woodworth queried "[Since] we cannot break up the reaction into successive acts and obtain the time of each act, of what use is the reaction-time?" And, more recently, D. M. Johnson said in his *Psychology of Thought and Judgment* (1955, p. 5), "The reaction-time experiment suggests a method for the analysis of mental processes which turned out to be unworkable."

Nevertheless, the attempt to analyze RT into components goes on, and there has been a substantial revival in the last few years in the use of RT as a tool for the study of mental processes ranging from perceptual

[2] For example, Cattell (1886, p. 377) reported that "I have not been able myself to get results by [Wundt's] method. I apparently either distinguished the impression and made the motion simultaneously, or if I tried to avoid this by waiting until I had formed a distinct impression before I began to make the motion, I added to the simple reaction not only a perception, but a volition."

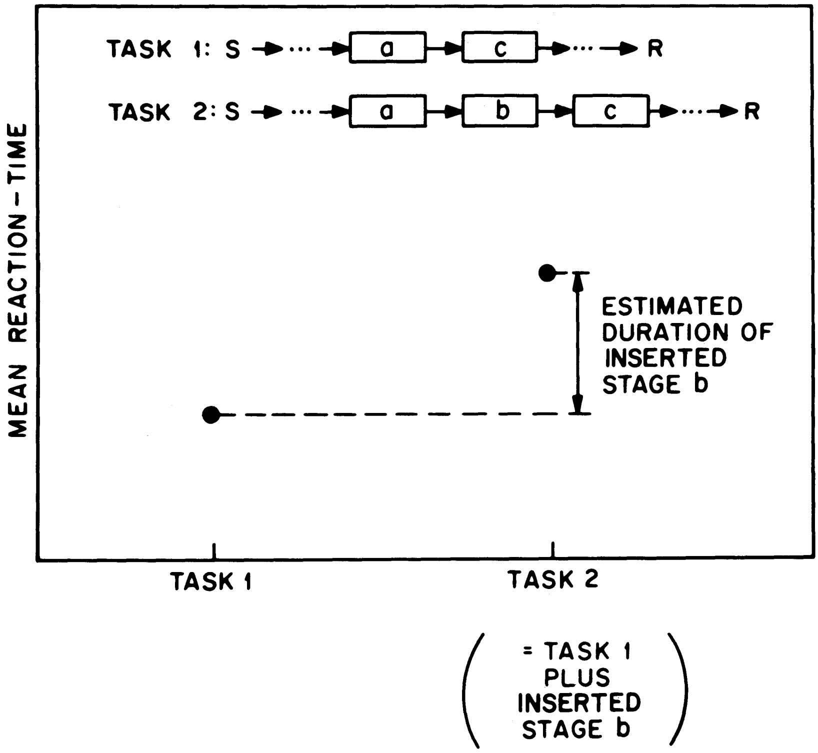


FIG. 1. Donders' subtraction method. Hypothetical stages between stimulus (*S*) and response (*R*) are represented by *a*, *b*, and *c*.

coding to mental arithmetic and problem-solving.[3] The work on memory retrieval described here is part of this revival, and is based heavily on Donders' stage theory. Modern styles of experimentation and data analysis lead to applications of the stage theory that seem to withstand the early criticisms, and to tests of validity other than introspection.

I shall describe experiments on retrieval from memory that have led to the discovery of some relatively simple search processes. My aim is to convey the general outline rather than the details of this work, so the picture I paint will be somewhat simplified; there will be little discussion of alternative explanations that have been considered and rejected. Such discussions can be found in Sternberg (1966, 1967a, b, and 1969).

The purpose of most of these experiments has been to study the ways in which information is retrieved from memory when learning and re-

[3] See, e.g., Egeth, 1966; Hochberg, 1968; Nickerson, 1967; Posner & Mitchell, 1967; Restle & Davis, 1962; Smith, 1967; Suppes & Groen, 1966.

tention are essentially perfect. The method is to present a list of items for memorization that is short enough to be within the immediate-memory span. The subject is then asked a question about the memorized list; he answers as quickly as he can, and his delay in responding is measured. By examining the pattern of his RTs, while varying such factors as the number of items in the list and the kind of question asked, one can make inferences about the underlying retrieval processes. Since the aim has been to understand error-free performance, conditions and payoffs are arranged so that in most experiments the responses are almost always correct.

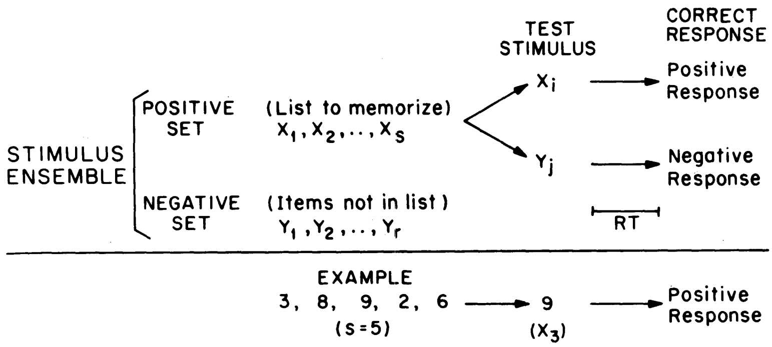


FIG. 2. Paradigm of item-recognition task (Exps. 1–5).

2. Judging Presence versus Absence in a Memorized List

The flavor of this approach will become clearer as we consider a particular experiment. Fig. 2 shows the paradigm of an *item-recognition task.* The stimulus ensemble consists of all potential test stimuli. From among these, a set of *s* elements is selected arbitrarily and is defined as the *positive set;* these items are presented as a list for the subject to memorize. The remaining items are called the *negative set.* When a test stimulus is presented, the subject must decide whether it is a member of the positive set. If it is, he makes a *positive response* (e.g., saying "yes" or operating a particular lever). If not, he makes a *negative response.* The measured RT (sometimes referred to as *response latency*) is the time from test-stimulus onset to response.

Within the item-recognition paradigm, different procedures can be used. One of them, shown at the top of Fig. 3, is the *varied-set procedure.* Here, the subject must memorize a different positive set on each trial. In one experiment (Exp. 1), for example, the stimulus ensemble consisted of the ten digits. On each trial a new positive set, ranging randomly over trials from one to six different digits, was presented sequentially at a rate of 1.2 seconds per digit. Two seconds after the last digit in the set was displayed, a warning signal appeared, followed by a visually-presented test digit. The subject pulled one lever, making a positive response, if the test stimulus was contained in the memorized list. He pulled the other lever, making a negative response, if it was not. After responding to the test stimulus the subject recalled the list. This forced him to retain the items in the presented order, and prevented him from working with the negative set rather than the positive. Regardless of the size of the positive set, the two responses were required equally often. As in the other experi-

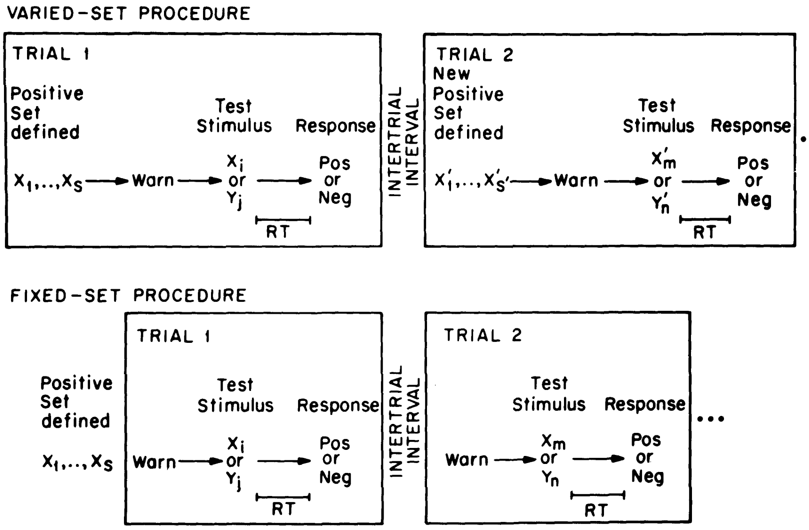


FIG. 3. Varied-set and fixed-set procedures in item-recognition. A Y represents an item in the negative set. Primes are used in representing trial 2 of the varied-set procedure to show that both the items in the positive set ($X_1, \ldots, X_s$) and its size (s) may change from trial to trial.

ments I shall describe, subjects were relatively unpracticed. The error rate in this kind of experiment can be held to 1 or 2 percent by paying subjects in such a way as to penalize errors heavily while rewarding speed.

Averaged data from eight subjects are shown in Fig. 4. Mean RT is plotted as a function of the number of symbols in memory—that is, the number of digits in the positive set that the subject committed to memory at the start of the trial.

These data are typical for item-recognition experiments. They show, first, a linear relation between mean RT and the size of the positive set. Second, the latencies of positive and negative responses increase at approximately the same rate. The slope of the line fitted to the means is 38 msec per item in memory; its zero-intercept is about 400 msec. (It happens to be true in these data that latencies of positive and negative responses have approximately the same *values*: the two latency functions have not only the same slope but also the same zero-intercept. This is not a general finding, but results from the particular conditions in this experiment. By varying the relative frequency with which positive and negative responses are required, for example, one can vary the relation between their latencies. But as relative frequency is varied the *slopes* of the two latency functions remain equal and unchanged.) Before considering the interpretation of these findings, we turn to some general matters regarding search processes.

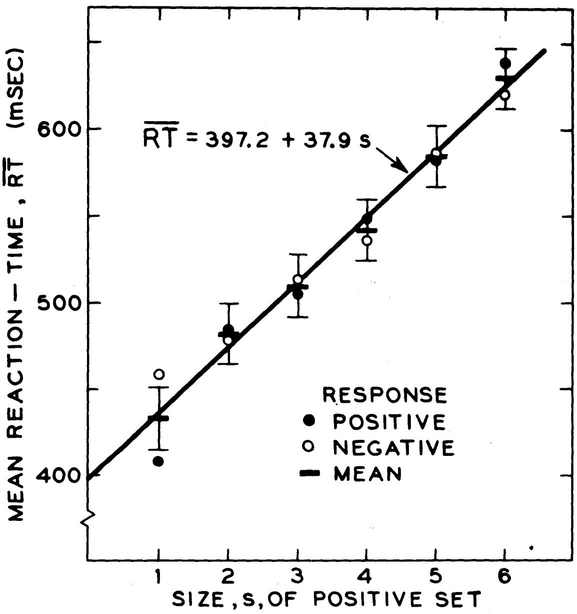


FIG. 4. Results of Exp. 1: Item-recognition with varied-set procedure. Mean latencies of correct positive and negative responses, and their mean, as functions of size of positive set. Averaged data from eight subjects, with estimates of $\pm\sigma$ about means, and line fitted by least squares to means.

3. *Two Types of Serial Search*

Let *serial search* (or *scanning*) be a process in which each of a set of items is compared one at a time, and no more than once, to a target item. Linear RT-functions, as in Fig. 4, suggest that subjects in the item-recognition task use a serial search process whose mean duration increases by one unit for each additional comparison. The purpose of the search is to determine whether an agreement (or *match*) exists between the test item and any of the items in the memorized set. Two types of serial search that might serve this purpose need to be considered. In *self-terminating serial search*, the test stimulus is compared successively to one item in memory after another, either until a match occurs (leading to a positive response), or until all comparisons have been completed without a match (leading to a negative response). In *exhaustive serial search*, the test stimulus is compared successively to *all* the memorized items. Only then is a response made—positive if a match has occurred, and negative otherwise. A self-terminating search might require a separate test, after each comparison, to ascertain whether a match had occurred, rather than only one such test after the entire series. On the other hand, an exhaustive search must involve more comparisons, on the average, than a search that terminates when a match occurs.

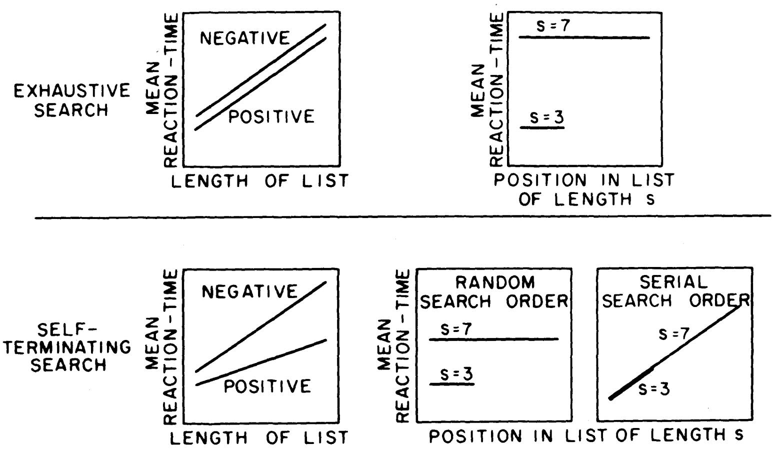


FIG. 5. Some properties of exhaustive (top) and self-terminating (bottom) serial search. Left: Theoretical RT-functions (mean latencies of positive and negative responses as functions of length of list). Right: Theoretical serial-position functions (mean latency of positive responses as a function of serial position of test item in a list of given length).

Suppose that the average time from the beginning of one comparison to the beginning of the next is the same for each comparison in the series, and is not influenced by the number of comparisons to be made. Then the durations of both kinds of search will increase linearly with the number of memorized items (*list length*). There are, however, important differences. In an exhaustive search the test stimulus is compared to all items in memory before each positive response as well as before a negative response. Hence, the rate at which RT increases with list length—the slope of the RT-function—is the same for positive and negative responses. In contrast, a self-terminating search stops in the middle of the list, on the average, before positive responses, but continues through the entire list before negatives. The result is that as list length is increased, the latency of positive responses increases at half the rate of the increase for negatives. This difference between the two kinds of search is illustrated on the left side of Fig. 5.

A second difference between the two types of search, illustrated on the right side of Fig. 5, is in the serial-position functions for positive responses. In a simple exhaustive search neither the order of search nor the position of the matching item in the list should have any effect on the RT, since all items are compared. A self-terminating search that occurred in a random order, or started at a random point, also would produce flat serial-position curves. But if a self-terminating search started consistently with the first item, and proceeded serially, then the serial-position curves would increase linearly. (If, in addition, list length influenced *only* the search process, then the curves for different list lengths would be superimposed: for example, the time to arrive at the second item in a memorized list would be independent of the length of the list.) Increasing serial-position functions are therefore sufficient (but not necessary) evidence for inferring that a search process is self-terminating.

4. *High-speed Exhaustive Scanning*

The serial-position curves actually observed in the item-recognition experiment described in Sec. 2 were relatively flat.[4] Together with this finding, the linearity of the latency functions and the equality of their slopes for positive and negative responses indicate an exhaustive search. The data show also that memory-scanning can proceed at a remarkably high rate. The slope of the mean RT-function, which is an estimate of the time per comparison, was 38 msec, indicating an average scanning rate between 25 and 30 digits per second.

Perhaps because of its high speed, the scanning process seems not to have any obvious correlate in conscious experience. Subjects generally say either that they engage in a self-terminating search, or that they know immediately, with no search at all, whether the test stimulus is contained in the memorized list.

Is high-speed scanning used only when a list has just been memorized and is therefore relatively unfamiliar? The results discussed so far (Fig. 4) are from the varied-set procedure (Fig. 3), in which the subject must memorize a new positive set on each trial, and is tested only three seconds after its presentation. How is the retrieval process changed

[4] Several investigators have, however, reported marked recency effects in item-recognition tasks: RTs were shorter for test stimuli later in the list (Corballis, 1967; Morin, DeRosa & Stultz, 1967; Morin, DeRosa & Ulm, 1967). Without embellishment a theory of exhaustive scanning cannot, of course, handle such findings. The salient procedural characteristics of experiments that produce such recency effects seem to be a fast rate of list presentation and a short interval (less than 1 sec) between the last item in the list and the test item. Findings of Posner, *et al.* (1969), indicate that in this range the time interval between successive stimuli may critically influence the nature and duration of comparison operations.

when a person is highly familiar with a particular positive set and has had a great deal of practice retrieving information from it? At the bottom of Fig. 3 is shown the *fixed-set procedure* in the item-recognition paradigm, in which the same positive set is used for a long series of trials. For example, in one experiment (Exp. 2) subjects had 60 practice trials and 120 test trials for each positive set. On the average test trial, a subject had been working with the same positive set for ten minutes, rather than three seconds. The sets were sufficiently well learned that subjects could recall them several days later. Sets of one, two, and four digits were used. There were six subjects.

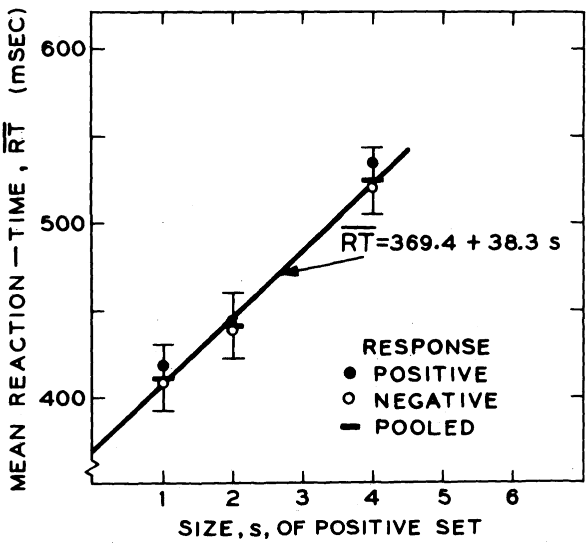


FIG. 6. Results of Exp. 2: Item-recognition with fixed-set procedure. Mean latencies of correct positive, negative, and pooled responses as functions of size of positive set. Averaged data from six subjects, with estimates of $\pm\sigma$ about pooled means, and line fitted by least squares to those means. For each set size positive responses were required on 27% of the trials.

Results are shown in Fig. 6, and are essentially identical to those from the varied-set procedure. The RT data are linear, the slopes for positive and negative responses are equal, and the average slope is 38 msec per digit. The small difference between the zero-intercepts in the two experiments is not statistically significant. The remarkable similarity of results from the two procedures indicates that the same retrieval process was used for both the unfamiliar and the well-learned lists.

5. Active and Inactive Memory

Evidence has accumulated, particularly during the past decade, that there are at least two systems or states of memory for encoded verbal items (e.g., Broadbent, 1958; Waugh & Norman, 1965; Glanzer & Cunitz, 1966; Atkinson & Shiffrin, 1968). The picture that is emerging is roughly as follows: The *long-term store*, or *inactive memory*, is relatively permanent and of large capacity. It receives information from the *short-term store*, a temporary *active memory*[5] of small capacity from which information is rapidly lost unless an active retention process is operating. In the long-term store, the coding of verbal items includes semantic attributes; in the short-term store, however, such items are coded primarily as acoustic or articulatory representations of their spoken names, even when they have been presented visually (see Sperling, 1960; Conrad, 1964; Baddeley, 1966; Wickelgren, 1969). The active process that regenerates the rapidly-decaying traces of a list of items in the short-term store is *rehearsal*, the overtly- or silently-spoken cyclic serial recall of stored items (see Sanders, 1961; Sperling, 1963; Posner & Rossman, 1965; Cohen & Johansson, 1967; Crowder, 1967; Atkinson & Shiffrin, 1968). Rehearsal, which also causes information in the short-term store to be entered in the long-term store, has an approximate maximum rate of from three to seven items per second (Landauer, 1962).

Whereas in the varied-set procedure of Exp. 1 the positive set must have been stored in active memory only, it is reasonable to believe that the set had entered the long-term store in the fixed-set procedure of Exp. 2. However, the similarity of results from the two procedures suggests that the same memory system was being scanned: that is, when information in inactive memory has to be used, it may be entered also in active memory (where it is maintained by rehearsal) and thus become more readily available. An experiment that tests this conjecture is described below (Exp. 5).

It appears, then, that the memory of the positive sets in both tasks is maintained by a serial rehearsal process; supporting this notion, subjects reported silent rehearsal of the sets in both experiments. But the estimated rates of high-speed scanning and the fastest silent speech differ by a factor of at least four. Rehearsal is far too slow to be identical to the scanning process. Instead, it should be thought of as a separate process whose only function in these tasks is to maintain the memory that is to be scanned.[6]

[5] An alternative term is "working memory," used by Newell & Simon, 1963, to refer to the arithmetic unit of a general-purpose computer. See also "Active verbal memory," Ch. 9 in Neisser, 1967, and "Operational memory," Sec. 4 in Posner, 1967.

[6] It is sometimes thought that the six or seven objects in the "span of apprehension" are immediately and simultaneously available, being contained in the "psychological present." And the information in active memory has occasionally been identified with this momentary capacity of consciousness (e.g., Miller, 1962, pp. 47–49; Waugh & Norman, 1965). The finding that one must scan one's active memory to ascertain its contents, rather than having immediate access to them, reveals a possible flaw in this argument.

6. Encoding of the Test Stimulus

In the scanning process inferred from these experiments, some internal representation of the test stimulus is compared to internal representations of the items in the positive set. What is the *nature* of the representations that can be compared at such high speed? Another way to phrase the question is to ask how much processing of the test stimulus occurs before it is compared to the memorized items.

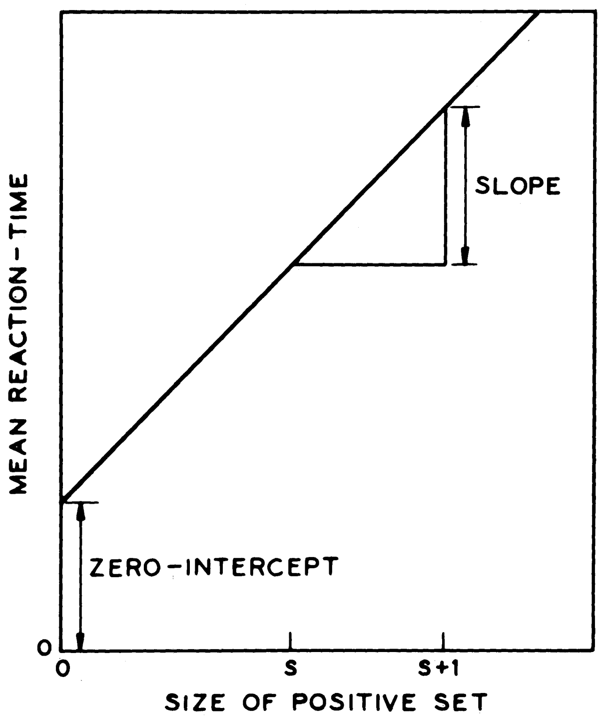


FIG. 7. Idealization of mean RT-function from an item-recognition task.

Various considerations lead one to expect a good deal of preprocessing. For example, the idea that items held in active memory are retained as acoustic or articulatory representations of their spoken names introduces the possibility that the test stimulus is processed to the point of naming, and that the name of the test stimulus is compared to the names of the items in the positive set. But two points should be kept in mind regarding this possibility. First, it would require that stored names could be scanned much faster than they could be covertly articulated, since the scanning rate is about four times as fast as people can say names of digits to themselves. Second, unlike other forms of preprocessing, such as image-sharpening or feature-extraction, preprocessing a character to the point of identification or naming would itself require the retrieval of informa-

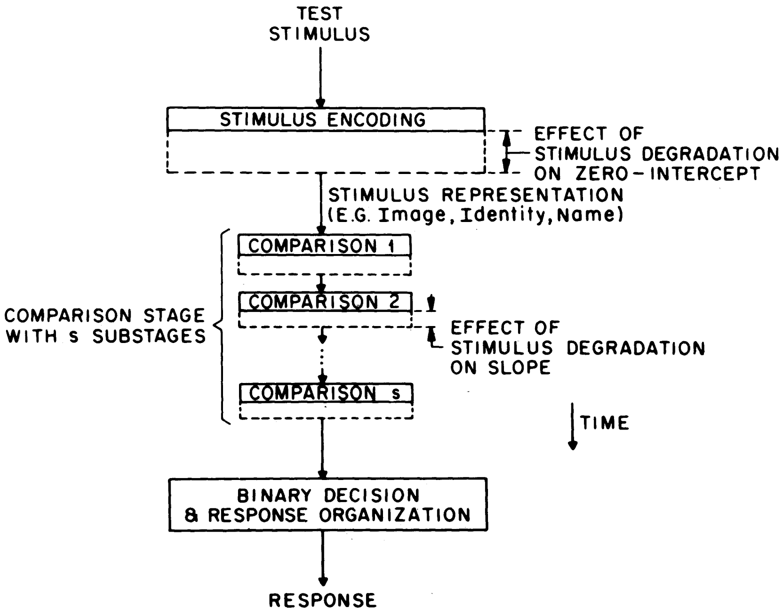


FIG. 8. Some hypothetical stages and substages in item-recognition, and two possible effects of test-stimulus quality on stage and substage durations. Height of box represents mean duration of that stage or substage.

tion from memory—that information which relates the character to its name.

In one experiment bearing on this question (Exp. 3), I degraded the test stimulus by superimposing a pattern that had been adjusted to increase the RT without substantially altering the error rate. I then examined the effect of stimulus quality on the function that relates mean RT and the size of the positive set. It is shown below that this effect would depend on the nature of the internal representation of the test-stimulus.

Figure 7 shows idealized data from a scanning experiment. The zero-intercept corresponds to the total duration of all processes that occur just once, regardless of the size of the positive set—such as the encoding of the test stimulus to form its representation, and the organization and execution of the motor response. The slope, on the other hand, measures the duration of processes that occur once for each member of the positive set—the comparison operation, and the time to switch from one item to the next.[7] Figure 8 shows a flow diagram of some hypothetical stages

[7] This analysis assumes that the mean durations of comparisons leading to matches and to mismatches are equal. Without this assumption all the statements here (and elsewhere in the paper) are correct, except that the slope of the RT-function measures the mean duration of only those comparisons that lead to mismatches, together with the time to switch from one comparison to the next. Any difference between durations of the two kinds of comparison would contribute to a difference between zero intercepts of the latency functions for positive and negative responses.

between test stimulus and response. The height of a box represents the mean duration of that stage. An effect of stimulus degradation on the *stimulus-encoding stage*, which generates the stimulus representation, would increase the zero-intercept of the RT-function. An effect on the *serial-comparison stage* would increase the slope, since a time increment would be added for each item compared.

Consider two extreme possibilities: First, suppose that the encoding stage did nothing other than transmit an unprocessed image, or direct copy, of the test stimulus. Then degradation could influence only the comparison operation, which occurs once for each member of the positive set; only the slope of the RT-function would change, as in Panel A of

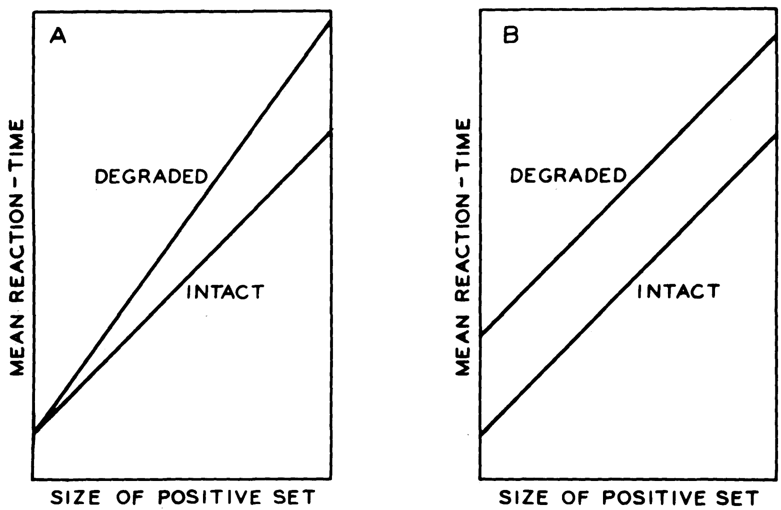


Fig. 9. Two possibilities for the effect of test-stimulus quality on the RT-function. A: Quality influences comparison stage only. B: Quality influences encoding stage only.

Fig. 9. At the other extreme, suppose that the representation produced by the encoding stage was the *name* of the test stimulus. The input to the serial-comparison stage would be the same, whether or not the test stimulus had been degraded by a superimposed visual pattern; hence degradation could not influence this stage. (For the serial-comparison stage to be influenced by visual degradation, its input would have to be visual, in the sense of embodying details of the physical stimulus pattern that are not present in the mere name of the stimulus.) Only the encoding stage, then, could be influenced by degradation; and since encoding takes place just once, only the zero-intercept of the RT-function would

change, as in Panel B of Fig. 9. (The absence of a change in slope, however, does not necessarily imply a nonvisual stimulus-representation; the representation could be visual, but highly processed.)

In Exp. 3 each of twelve subjects had positive sets of one, two, and four digits, with test stimuli *intact* in some blocks of trials, and in others *degraded* by a superimposed checkerboard pattern. Intact and degraded numerals are shown in Fig. 10.

The fixed-set procedure was used. Results for the two sessions are shown separately in Fig. 11. Consider first the data from the second session, on the right-hand side of the figure. Latencies of positive and negative responses have been averaged together. The functions for degraded and intact stimuli are almost parallel, but there is a large effect on the zero-intercept, closely approximating the pattern shown in Panel B of Fig. 9. This indicates that degradation had a large influence on the stimulus-encoding stage, and that the representation generated was sufficiently processed that the serial-comparison stage could proceed as rapidly with degraded as with intact stimuli. The stimulus representation was either nonvisual or, if visual, sufficiently refined in the second session to eliminate any effect of degradation.

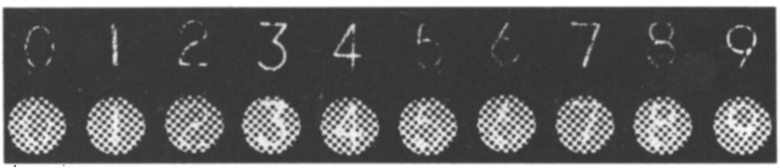


FIG. 10. Photographs of intact and degraded numerals used in Exp. 3. Numerals were about 0.6 in. high and were viewed from a distance of about 29 in. Degraded numerals were somewhat more discriminable than they appear in the black-and-white photograph, possibly because of a slight color difference between numerals and checkerboard.

The data from this session are an instance of the *additivity* of two effects on RT. There is no interaction between the effect of set size and the effect of stimulus quality; instead, the effect of each of these factors on mean RT is independent of the level of the other. Such additivity supports the theory of a sequence of stages, one stage influenced by stimulus quality and the other by set size (see Sec. 7).

Now let us consider the data from the first session, shown on the left-hand side of Fig. 11. Here, where subjects have not yet had much practice with the superimposed checkerboard, there is a 20% increase in the slope of the RT-function, as well as an increase in its zero-intercept. This pattern agrees with neither of the pure cases of Fig. 9. Stimulus quality apparently *can* influence the duration of comparison operations; hence, the output of the encoding stage must be sensitive to degradation. Findings from the two sessions imply, then, that although the stimulus representation is highly processed, it embodies physical attributes of the test stimulus, rather than being a name or identity. That is, the test-stimulus representation is visual. The memory representations of the positive set that are used in the serial-comparison stage must therefore also be visual, to make comparison possible. Hence, although items in the positive set appear to be represented as covertly-spoken names in the course of their rehearsal, this is not the only form in which they are available.

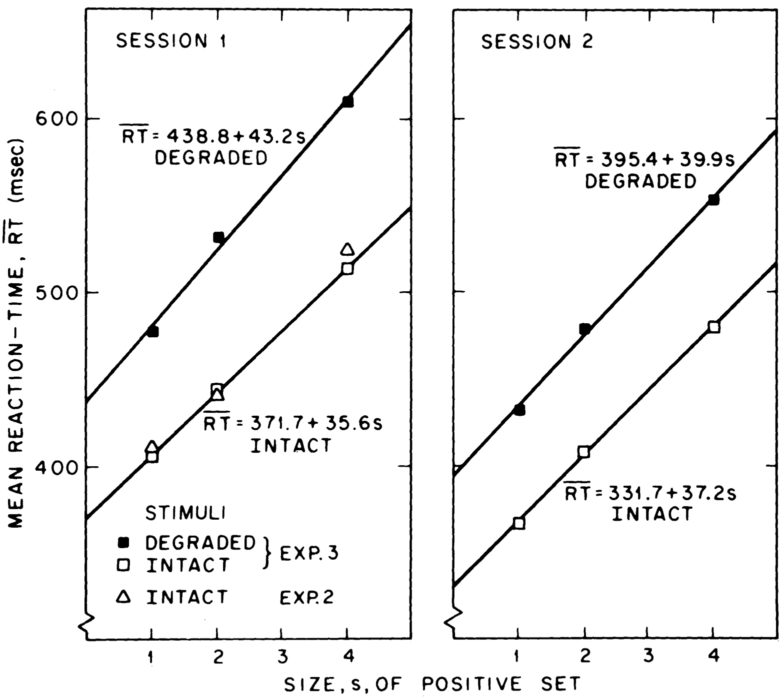


FIG. 11. Results of Exp. 3: Effect of stimulus quality on item-recognition. Mean RT, based on pooled data from positive and negative responses, as a function of size of positive set for intact and degraded test stimuli. Left-hand and right-hand panels show data from Sessions 1 and 2, respectively. Averaged data from 12 subjects, with lines fitted by least squares. In all conditions positive responses were required on 27% of the trials. Triangles show results from Exp. 2 (Fig. 6), which was similar.

What changed between the first and second sessions so as to virtually eliminate the influence of stimulus quality on the slope of the RT-function? Since the scanning rate with intact stimuli and the effect of degradation on the zero-intercept are approximately the same in the two sessions, it seems unlikely that the type of representation changed. For the present, my interpretation is that the encoding stage became more efficient at removing the effects of the fixed degrading pattern.

Additional support for the idea that the memory representations scanned in the item-recognition task have sensory characteristics, rather than being completely abstracted from the physical stimuli, comes from two other studies. In the first, Chase and Calfee (1969) created four different conditions in the varied-set procedure by representing both the positive set and the test stimulus either visually or aurally. When the set and test item were presented in different modalities, the slope of the RT-function increased by about 30%, indicating a slower scanning rate. If abstract representations were being compared in the same-modality conditions, then the change to different-modality conditions should have altered only the zero-intercept, as in Fig. 9B. In the second study, Posner, *et al.* (1969), concluded that when a single letter is presented *aurally* for memorization, the decision whether a *visual* test-letter is the "same" is facilitated by the internal generation of a visual representation of the memorized letter, which obviates the need to identify the test letter. Still further evidence will be discussed below (Exp. 4).

7. *A Test of the Stage Theory*

The work described above is grounded on Donders' stage theory. That is, as in his subtraction method, the effects on mean RT of changes in experimental conditions (factors) have been attributed to the selective effects of these factors on hypothetical processing stages between stimulus and response. How can we ensure that such inferences are not open to the classical criticism of the subtraction method, that even if information processing *is* organized in functionally different stages, factor effects may not be selective? One answer, of course, is that the test of a method's applicability is whether it produces results that fit together and make sense. But there are two other arguments as well.

The first stems from replacement of the assumption of pure insertion by a weaker and more plausible *assumption of selective influence*. Instead of requiring that a change in the task insert or delete an entire processing stage without altering others, the weaker assumption requires only that it influence the *duration* of some stage without altering others. One example is illustrated in Fig. 12. To estimate the comparison time by the subtraction method, one would have studied Task 2, in which the positive set has one member, and compared it to a Task 1. Task 1 would have been constructed to measure the zero-intercept directly, by deleting the entire comparison stage. But I suspect that there *is* no appropriate Task 1, in which deletion of all comparisons would leave the other stages of processing invariant. In this instance, then, the assumption of pure insertion is probably invalid. This is why the important RT-differences in the experiments described above were those between Tasks 2 and 3, 3 and 4, and so on, whose interpretation required only that the comparison stage be selectively *influenced* by set size. Similarly, in studying the preprocessing of the stimulus, instead of entirely eliminating the need to discriminate the stimulus (in an effort to *delete* the hypothetical encoding stage) I examined the effects of making its discrimination more or less difficult, thereby varying the amount of work the stage had to accomplish. Of course, one result of using a factor that influences but does not insert a stage is that we have no estimate of the stage's total duration.

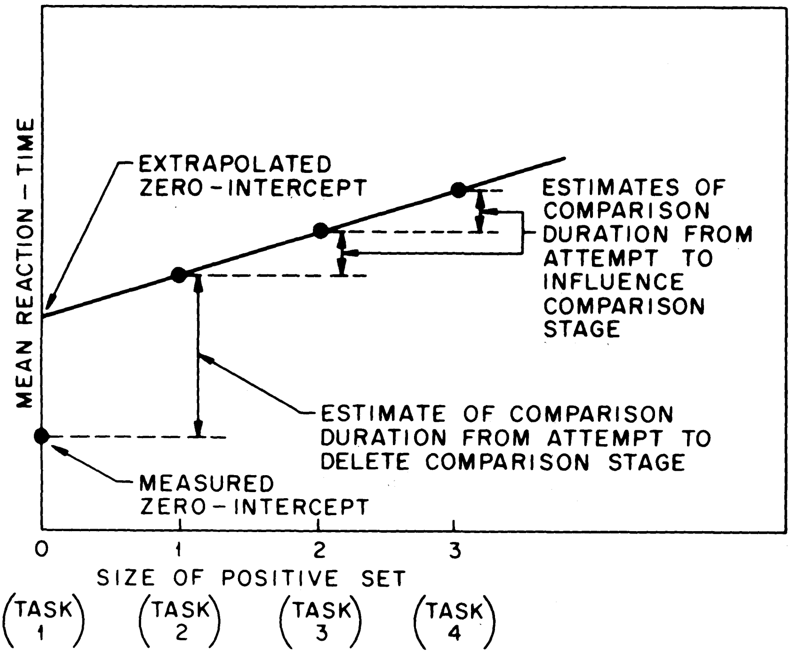


FIG. 12. Example of error from hypothetical attempt to estimate comparison time by deleting the comparison stage altogether, as in the subtraction method, and to use a measured zero intercept. Attempt fails because *deletion* of comparison stage changes the demands placed on other stages, whereas *variation* of the number of comparisons, s, ($s \geq 1$) does not.

But that seems to be of less interest than whether there is such a stage, what influences it, what it accomplishes, and what its relation is to other stages.[8]

In a given experimental situation the validity of even the weaker assumption of selective influence must be checked, however. We can distinguish those situations where one of the assumptions—influence or insertion—holds, by testing the additivity of the effects of two or more factors on mean RT (Sternberg, 1969). It is this test that provides the

[8] This alternative was preferred by Cattell, 1886, who argued (p. 378) "I do not think it is possible to add a perception to the reaction without also adding a will-act. We can however change the nature of the perception without altering the will-time, and thus investigate with considerable thoroughness the length of the perception time." But he suggested no way to test these assertions.

second and most telling way of dealing with the classical criticism. Consider a pair of hypothetical stages and a pair of experimental factors, with each factor inserting or selectively influencing one of the stages. Because stage durations are additive (by definition), the changes in mean RT produced by such factors should be independent and additive. That is, the effect of one factor will be the same at all levels of the other, when the response is measured on a scale of time or its (arithmetic) mean.[9]

In experiments with the fixed-set procedure I have examined four factors, which are listed above the broken line in Fig. 13. The additivity of five

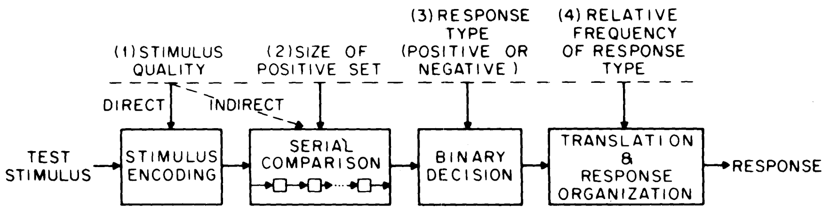


FIG. 13. Four processing stages in item-recognition. Above the broken line are shown the four factors examined. Below the line is shown the decomposition of RT inferred from additive relations between factor pairs 1&2, 1&3, 2&3, 2&4, and 3&4, the linear effect of factor 2, and other considerations. (The indirect effect of factor 1 on the comparison stage, and the resulting interaction of factors 1&2, is seen in unpracticed subjects only.)

of the six possible factor pairs has been tested and confirmed (1&2, after a session of practice, 1&3, 2&3, 2&4, and 3&4). These instances of additivity support the assumption that the factors selectively influence different stages of processing and, *a fortiori*, confirm the existence of such stages. Another instance of additivity, and the one on which inferences about the structure of the comparison stage strongly depend, is represented by the linearity of the effect of set size: the effect of adding an item to the positive set is independent of the number of items already in the set. Together with other considerations (discussed in Sternberg, 1969) these findings lead to the analysis into processing stages and substages shown below the broken line in Fig. 13.[10]

[9] Discussions of various other aspects and modern versions of the subtraction method, including considerations of validity, may be found in Hohle, 1967; McGill & Gibbon, 1965; McMahon, 1963; Smith, 1968; Sternberg, 1964; Sternberg, 1969; and Taylor, 1966.

[10] The linear interaction between stimulus quality and set size in Session 1 is attributed to an "indirect" influence of stimulus quality on the duration of the second stage, by way of its effect on the output of the first stage (see Sec. 6, and Sternberg, 1967b). Thus one may sometimes infer a separate stage even when its output is not invariant with respect to a factor that influences its duration, and when as a consequence there is a failure of additivity. In this instance the inference is justified by the form of the interaction (a *linear* increase in the effect of degradation with set size), and the structure of the comparison stage (inferred to be a series of substages).

8. *Generality of High-speed Scanning*

Let us turn now to more substantive matters, and consider the generality of the high-speed exhaustive scanning process. Binary classification of digits into sets that are small, randomly-assembled, and relatively unfamiliar is hardly a typical example of memory retrieval. But it is useful to pin down one process fairly well, and explore techniques that reveal it in a relatively pure form, in order to use it as a baseline for the study of other mechanisms.[11]

For one example of a possible alternative to serial search, consider the case where the items in a memorized set share a physical feature whose presence distinguishes them from the rest of the stimulus ensemble. Here one might expect subjects to test the stimulus for the presence of the feature rather than compare it to the items in the set one by one. Surprisingly, using letters with a diagonal line-segment as the distinguishing feature, Yonas (1969) showed that subjects start by scanning the set; only after considerable practice do they use the feature test, thereby eliminating the effect of the number of letters in the set.

Another possible alternative to serial search is an "associative" process. Consider the case in which positive items are distinguished by membership in a well-learned category. (For example, the positive set might contain digits only, and the negative set, letters.) To each member of a category is associatively linked its category label, and the binary choice depends on which label is elicited by the test stimulus. The speed of such a process might be independent of the sizes of positive and negative sets (although it might depend on various attributes of the categories that contained them, including *their* sizes; see Landauer & Freedman, 1968). On the other hand, the high speed of scanning might make it more efficient than an associative process, when one of the sets is small. In short, there may be alternative mechanisms for the same task, and which one is used may depend, in part, on which one is more efficient. If this is the case it is a great advantage to understand at least one of the alternatives in some detail.

9. *Retrieval of Nonsymbolic versus Symbolic Information*

Other questions about the generality of the scanning process are raised by its high speed, which precludes its being identified with the subvocalization of numeral names, and also by the influence of stimulus-quality on the scanning rate in Exp. 3 (Sec. 6), which indicates that the stimulus representation is not the name or identity of the numeral. The fact that numerals are patterns with extremely well-learned names may therefore be irrelevant to the scanning process. Of course, numerals

[11] The function of such an experimental baseline is similar to the use of well-understood mathematical models as theoretical baselines (Sternberg, 1963, Sec. 6.6) in which it is the discrepancies between data and model that are of interest.

have other special properties: they are highly familiar, they are symbols, they represent numerical quantities, and people are practiced at manipulating the numbers they represent. A. M. Treisman and I recently tested the importance of these properties for memory retrieval, using two different ensembles, one of nonsense forms, and the other of photographs of faces (Exp. 4). To our subjects, both ensembles were unfamiliar, nonsymbolic, unordered, and without well-learned names. We used the fixed-

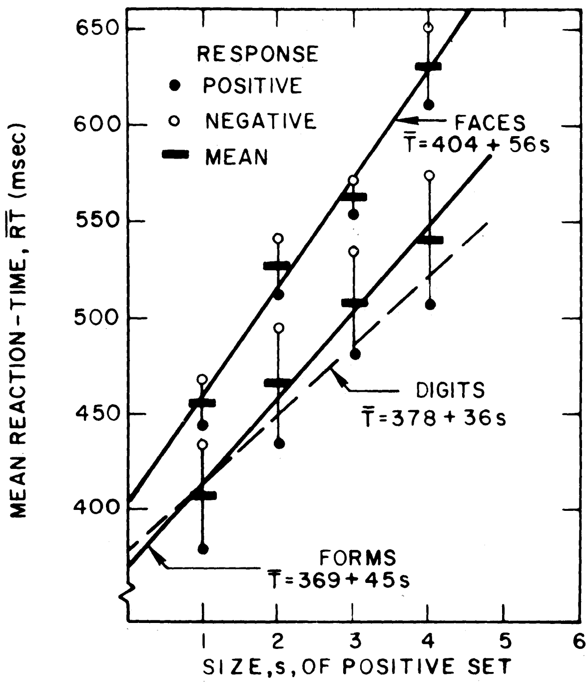


FIG. 14. Results of Exp. 4: Item-recognition with nonsense forms and photographs of faces. Mean latencies of correct positive and negative responses, and their mean, as a function of size of positive set for the two stimulus ensembles. Averaged data from eight subjects for each ensemble, with lines fitted by least squares to means. Broken line was fitted to data from a similar experiment wtih an ensemble of numerals.

set procedure with sets of size 1 to 4, but found it necessary to display the positive set before each trial in order to help the subjects, who were inexperienced, to maintain it in active memory.

RT data, shown in Fig. 14, are qualitatively the same as those for digit sets. They show linearity, suggesting a serial process, and equality of slopes for positive and negative responses, indicating exhaustiveness of search. The main difference is in the scanning rate, which seems to depend to some extent on the nature of the stimuli. Even for faces, however, the estimated rate is high—about 18 faces per second. These findings indicate that high-speed exhaustive scanning does not depend on the special properties of numerals mentioned above. They also add further support to the conclusion that the test-stimulus representation in the case of numerals is not the name of the numeral, but is some sort of visual representation.

10. Retrieval from Inactive versus Active Memory

A further question about the generality of the high-speed scanning process is raised by the conjecture (Sec. 5) that it occurs only when information is being held in active memory. The similarity of results from the varied-set and fixed-set procedures led to the idea that even when a list is contained in long-term memory, it is transferred into active memory and maintained there by rehearsal in order to be used in the

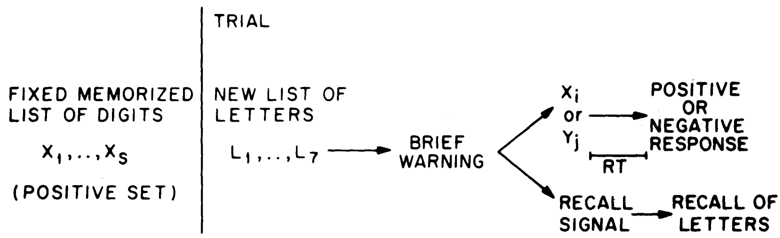


FIG. 15. Paradigm of Exp. 5: Item-recognition from active and inactive memory. Only the inactive-memory condition is shown. In the active-memory condition, also involving a fixed-set procedure (Fig. 3), no letters were presented.

item-recognition task. If that is so, one would expect some change in the process if one prevented the relevant list from being rehearsed (for example, by occupying the active memory with other material). This kind of procedure moves us closer to studying the differences between retrieval from the short-term (active) and long-term (inactive) stores, and thereby understanding the latter by using the former as a baseline.

The procedure in a small preliminary experiment (Exp. 5) is shown in Fig. 15. At the start of a series of trials the subject memorized a list of 1, 3, or 5 digits, which defined the positive set for the entire series. On each trial a new list of seven letters was presented sequentially, at a rate of two letters per second. A short time after the last letter, there was a brief warning signal, and then one of two things could happen. On a random third of the trials the subject saw a recall signal, and attempted to recall the seven letters. These trials were used in order to encourage the subject to attend to the letters and retain them in memory until the test event. (Observing and retaining the list of letters was intended to occupy his active memory on all trials and prevent him from rehearsing the positive set.) On the remaining trials the subject saw a test digit. He

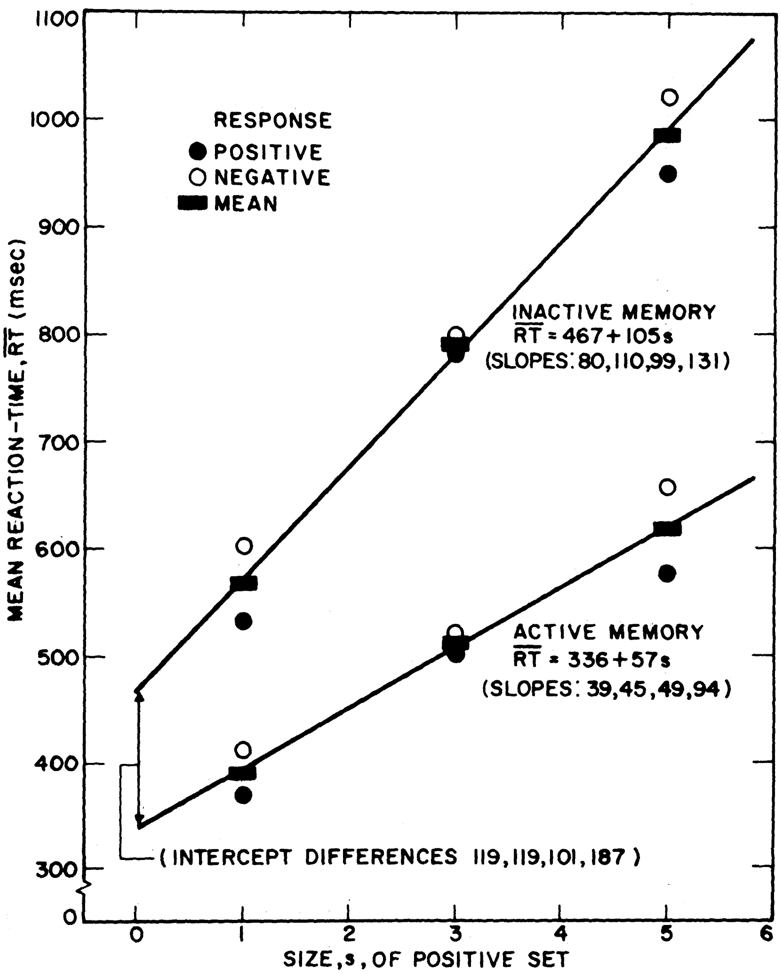
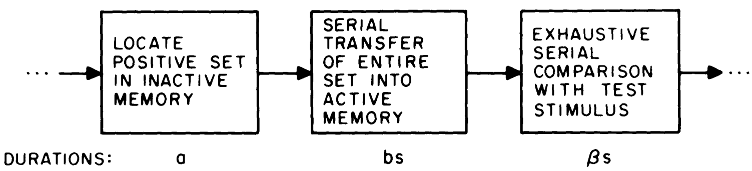


FIG. 16. Results of Exp. 5: Item-recognition from active and inactive memory. Mean latencies of correct positive and negative responses, and their mean, as functions of size of positive set, in conditions of active and inactive memory. Averaged data from four subjects, with lines fitted by least squares to means. Intercept differences and slopes for the four subjects are listed, the order of subjects being the same in each list.

was required to make a positive or negative response, based on the previously memorized digit set, as quickly as possible consistent with accuracy. This is a difficult task, and required a session of practice for smooth performance. In the series of control trials, which alternated with series of experimental trials, no lists of letters were presented.

Data averaged over the four subjects in this preliminary experiment are shown in Fig. 16. The lower set of points represents performance in the control condition, which was similar in procedure to Exp. 2. Mainly because of one exceptional subject, the fitted line is somewhat steeper than usual, with a slope of 57 msec per digit. Otherwise the data are typical. In the experimental condition the fitted line is about twice as steep as in the control condition, with a slope of 105 msec per digit. Again, the latencies of positive and negative responses grow at equal rates as set size is increased. The zero-intercepts in the experimental and control conditions differ also, by over 100 msec.

Evidently, the retrieval process is radically altered, with the effective scanning rate halved, when the information to be retrieved is not being rehearsed and is therefore not in active memory. Current notions about the functions of rehearsal include maintenance of short-term memory, and transfer of information into long-term memory (see Sec. 5). The results of Exp. 5 suggest a third role—that of making information already stored in long-term memory more rapidly accessible.



CONTROL CONDITION (ACTIVE MEMORY) $\overline{RT} = \alpha + \beta s$

EXPERIMENTAL CONDITION (INACTIVE MEMORY) $\overline{RT} = (\alpha + a) + (\beta + b)\, s$

FIG. 17. One explanation of results of Exp. 5. Left and middle boxes represent hypothetical stages that might be inserted in the inactive-memory condition. Also shown are hypothesized durations of these two stages and the comparison stage, and resulting theoretical RT-functions in which α represents the zero-intercept of the RT-function in the active-memory condition.

At this point there is little basis for selecting among potential explanations for the data from the experimental condition, but experiments are under way that may help to do so. The explanation that I favor, shown in Fig. 17, is the one that makes plausible two striking aspects of the data: despite the large effect of condition, the linearity of the RT-function and the equality of slopes for positive and negative responses are both preserved. The first two boxes in the figure represent hypothetical stages that might be present in the experimental condition but not in the control. One might be searching for the positive set in inactive memory. This would take a fixed time, regardless of the size of the positive set, and could account for the increase in the zero-intercept. The second added stage might be the serial transfer of each item in the positive set into active memory, with a fixed average time per item transferred, estimated from the data to be about 50 msec. Since all items would be transferred, whether the required response was positive or negative, the slopes of the functions for both responses would be increased by the same amount. The high-speed scanning stage, which we already know to be exhaustive, would follow. The two added stages are plausible and would account for the important features of the data. But this explanation—particularly the concept of "transferring a set of items into active memory"—needs to be made more precise and then tested.

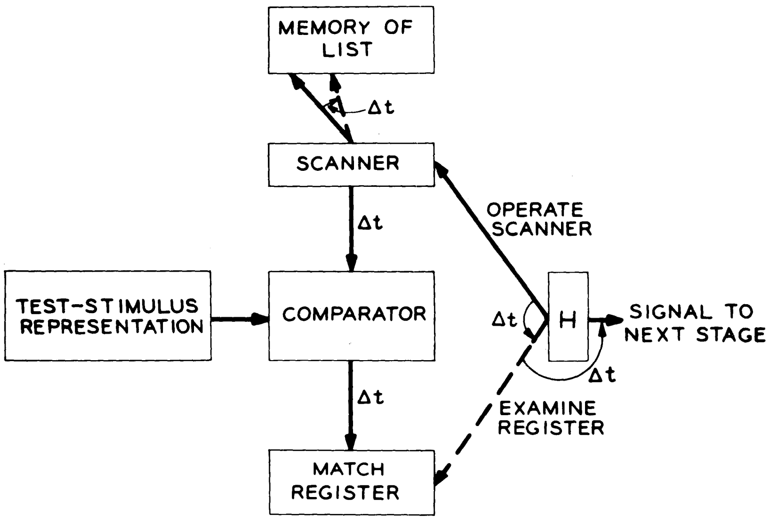


FIG. 18. A system in which exhaustive scanning could be more efficient than self-terminating scanning. Some loci of possible time delays are represented by Δts.

11. *An Explanation of Exhaustiveness*

As mentioned in Sec. 3, an exhaustive search must involve more comparisons, on the average, than a search that terminates when a match occurs. The exhaustiveness of the high-speed scanning process therefore appears inefficient, and hence implausible. Why continue the comparison process beyond the point at which a match occurs? Figure 18 illustrates a system in which an exhaustive search could be more efficient than a self-terminating one for performance in an item-recognition task. A representation of the test stimulus is placed in a comparator. When the scanner is being operated by the "central processor" or "homunculus," H, it delivers memory representations of the items in the list, one after another, to the comparator. If and when a match occurs a signal is delivered to the match register. The important feature of the system is that the homunculus can *either* operate the scanner *or* examine the register.

It cannot engage in both of these functions at once, and switching between them takes time.

In this kind of system, if the switching time is long relative to the scanning rate, and if the list is sufficiently short, then an exhaustive search (in which the match register must be examined only once) is more efficient than a self-terminating one (where the register would have to be examined after each comparison). The surprisingly high speed of the scanning process may therefore be made possible by its exhaustiveness. But such a system might have at least one important limitation. After the search was completed, there might be no information available (without further reference to the memory of the list) as to the location in the list of the item that produced the match. The limitation would create no difficulty if the response required of the subject depended only on the

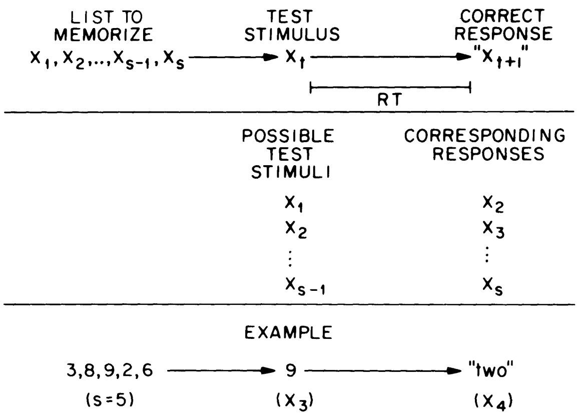


FIG. 19. Paradigm of Exp. 6: Context-recall.

presence or absence of an item in the list and not on its location, as in the item-recognition task. But the possibility that high-speed scanning does not yield location information does suggest an experiment to test this theory of exhaustiveness. Suppose we require a subject to give a response that *does* depend on where in the list a matching item is located. Then after each comparison, with information still available as to the location of the item just compared to the test stimulus (e.g., preserved by the position of the scanner in Fig. 18), it would have to be determined whether this item produced a match (by the homunculus switching from scanner to register). Scanning should then be slower than when only presence or absence has to be judged; it should also be self-terminating, since further comparisons after a match had been detected would be superfluous. Such a process will be called *scanning to locate.*

12. Retrieval of Contextual Information by Scanning to Locate

In Fig. 19 is shown the paradigm of a *context-recall task,* one of the experiments devised to test these ideas (Exp. 6). On each trial the subject memorized a new random list of from three to seven different digits, presented visually one after another. The length of the list was varied at random from trial to trial. After a delay and a warning signal, a test item was presented, randomly selected from among all the digits in the list except the last. The test item, then, was always present in the list. The correct response was the spoken name of the item that followed the test item in the memorized list. The idea was that in order to make this response—that is, to recall an item defined by its contextual relation to

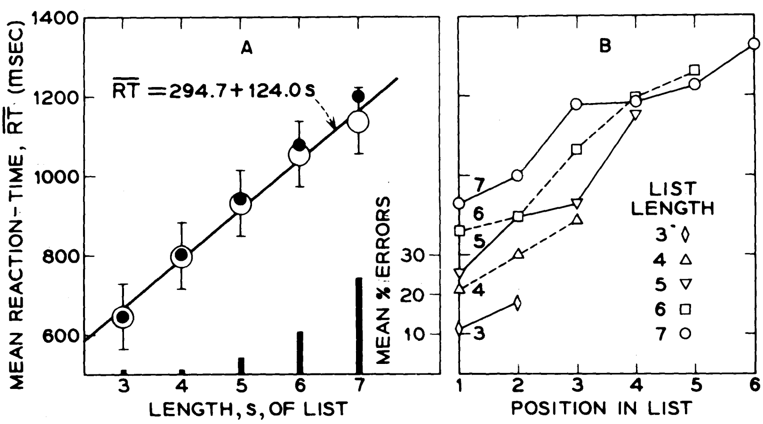


FIG. 20. Results of Exp. 6: Context-recall. Averaged data from six subjects. A: Effect of list length on percent errors (bars), on mean latency of correct responses (open circles) with estimates of $\pm\sigma$ and line fitted by least squares, and on mean RT of all responses (filled circles). B: Relation between mean RT of correct responses and serial position of the test item in lists of five lengths.

the test item—the location of the test item in the list might first have to be determined. As in the other experiments described, subjects were encouraged to respond as rapidly as possible, while attempting to maintain a low error rate.

Two aspects of the data are of particular interest: the relation between mean RT and list length; and the relation, for a list of given length, between RT and the serial position of the test item in the list.

Data averaged over six subjects are shown in Fig. 20. Consider first Panel A. The bars show the percentage of wrong responses, which rises to 25% for lists of length 7. This is much higher than one would like, given an interest in error-free performance. The effect of list-length on mean RT is roughly linear, suggesting a scanning process. (Even closer approximations to linearity have been found in other similar experiments.) With a slope of 124 msec per item, the fitted line is much steeper than the corresponding RT-function in the item-recognition task.

To interpret the slope, we have first to establish if the process is self-terminating, as expected. Evidence on this point is provided by the average serial-position functions shown in Panel B. For each list length, mean RT is plotted as a function of the serial position of the test item in the memorized list. These functions are all increasing, suggesting a self-terminating process that tends to start at the beginning of the list and proceed in serial order.

Now we can interpret the slope of the function in Panel A, if we assume that list length influences only the scanning stage. (Evidence supporting this assumption of selective influence is presented below.) Since an average of about half the items in a list have to be scanned before a match occurs, the slope represents half of the time per item, and implies a scanning rate of about 250 msec per item, or four items per second, in scanning to locate an item in a memorized list. Scanning to locate is therefore about seven times as slow as the high-speed scanning process

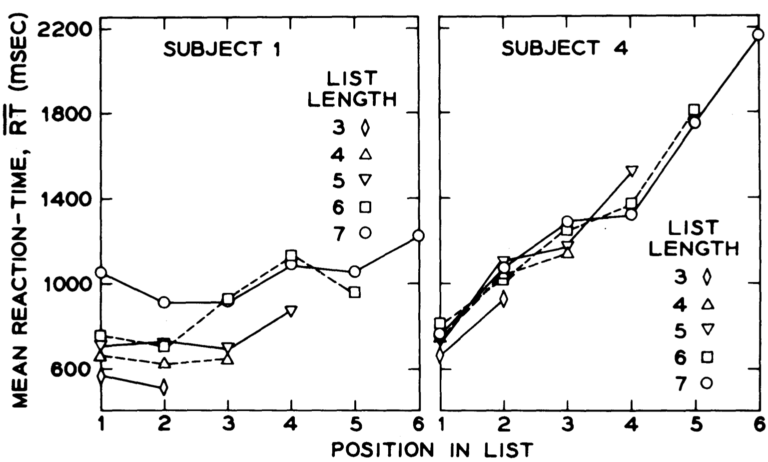


FIG. 21. Individual data from Exp. 6: Context-recall. Contrasting sets of serial-position functions in lists of five lengths, one set relatively flat and separated, the other steep and, in general, superimposed.

used to determine the presence of an item in a list. The slowness of the search, and the fact it is self-terminating, lend support to the explanation (Sec. 11) of the exhaustiveness of the high-speed process. Scanning to locate seems to be fundamentally different from scanning for presence.[12]

[12] Alternative explanations of the dissimilarity of the two kinds of scanning are possible, of course. One interesting alternative (which existing data cannot reject) is that memory representations that can carry order information are different from those that need only carry item information, and that the observed differences in retrieval result from the fact that different kinds of memory representations are being scanned. However, for this alternative explanation to apply to Exp. 1 (in which subjects had to recognize an item and then recall the entire list in order), it must be possible for both kinds of memory representation to be maintained simultaneously.

As mentioned earlier (Sec. 3), if a self-terminating process started consistently at the beginning of a list and proceeded serially, the serial-position functions would be steep and superimposed, whereas if it started at a random point they would be flat and separated. The functions shown in Panel B lie between these extremes. This is partly because they represent averages of data from several subjects. Data from two subjects in Exp. 6 who represent almost pure cases are shown in Fig. 21. The estimated scanning rates for these two subjects are almost the same, but their starting strategies appear to be radically different. Subject 1 seems

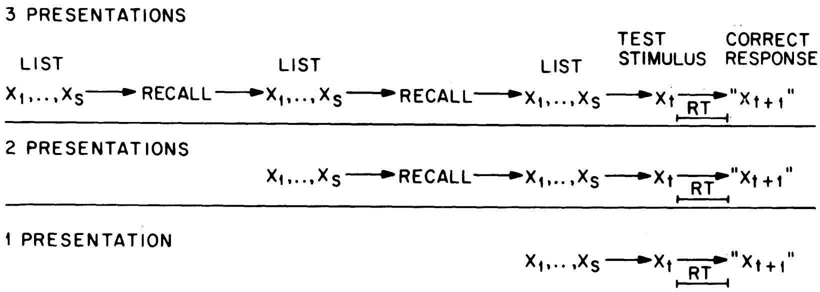


FIG. 22. Conditions in Exp. 7: Effect of learning on context-recall.

to have started at a random point. This could occur if the presentation of the test item interrupted an ongoing cyclic rehearsal process, and scanning then began at the serial position where rehearsal happened to have stopped. Subject 4, on the other hand, has the superimposed functions that would arise if he had started scanning consistently at the beginning of the list, perhaps by terminating his rehearsal before the test-stimulus appeared. Data from other subjects range between these extremes, presumably because of mixed starting strategies.

One explanation of these results is the following: In order to recall a contextual item, the subject must first determine the test item's location in the memorized list. This is achieved by a slow, self-terminating process of scanning to locate, in which the items in memory are compared successively to the test item until a match occurs. Each nonmatching item that participates contributes to the RT a component time that depends neither on list length nor on the item's position in the list. This component time is occupied by switching to the item, comparing it to the test stimulus, and determining that they have not matched. In the present context-recall task, the occurrence of a match is followed by a shift (e.g., a movement of the scanner in Fig. 18) from the item that matches the test item to the adjacent response item. For superimposed serial-position functions (as in Fig. 21) to be possible, we must assume that the duration of the shift operation (as well as other stages, such as stimulus-encoding and response-organization) is independent of the length of the list. Given this assumption, the slope of the RT-function is determined solely by the scanning rate.

The process of scanning to locate is a still more dramatic instance of having to hunt for information even when it is contained in a list that is being rehearsed. In some important sense one does not know what is in one's active memory, other than the single item to which attention is currently directed.[13]

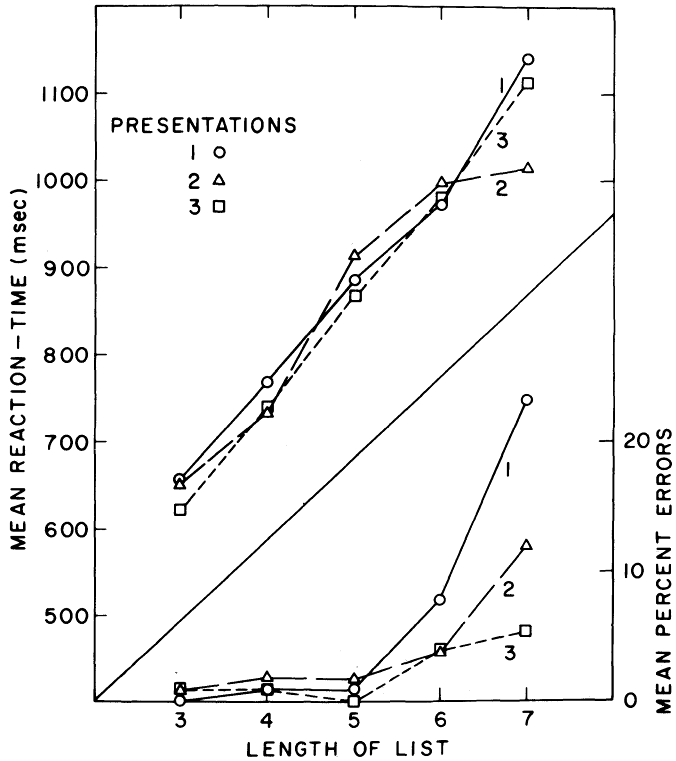


FIG. 23. Results of Exp. 7: Effect of learning on context-recall. Averaged data from six subjects for one, two, and three presentations of the list. Bottom: mean percent errors in naming contextual item. Top: mean latency of correct responses.

[13] One traditional view is that the structure of a memorized list is a chain of overlapping associated pairs of items: the subject's task in a context-recall experiment is thought of as the performance of one of the associations in the chain, and the RT measure as an index of associative strength. At the least, this view must be modified to recognize the existence of a search for the representation of the test item in the list. This search is an instance of the obligatory process (usually ignored by association theorists) that locates and activates the memory trace of a stimulus before an associative response to that stimulus can be performed (Rock, 1962). Furthermore, in this experiment, not only does the locating process produce the dominant effect, but also there appears to be *no* influence of associative strength (Sec. 13). One might therefore question whether the traditional view is at all appropriate, at least for lists contained in active memory. It has been challenged from other directions also in recent years (e.g., Slamecka, 1967).

13. Independence of Learning and Retrieval from Active Memory

One problem with Exp. 6 is the high error rate, and its marked increase with list length (Fig. 20A). This makes the RT data somewhat suspect and violates the aim of studying error-free processes. Moreover, it raises the possibility that the level of learning of the list, which is clearly lower for longer lists, might be contributing to the increase of RT with list length. (For example, suppose that a list embodies a chain of associations and that the recall of a contextual item involves the performance of one

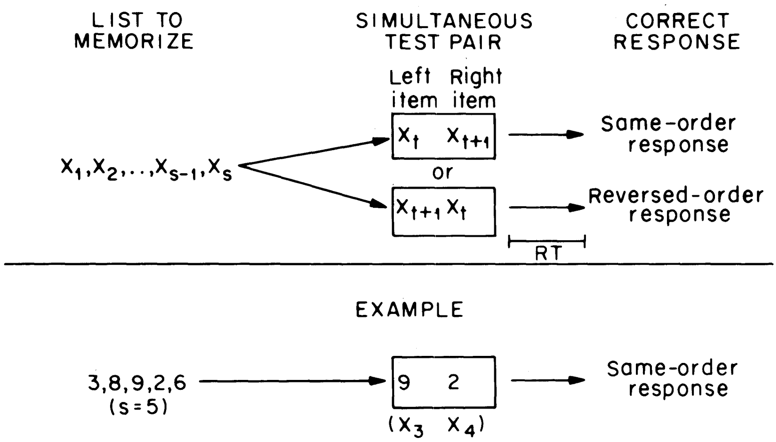


FIG. 24. Paradigm of Exp. 8: Context-recognition.

of the associations. If the associations in a longer list are weaker, then at least one of the sources of the effect of list length on RT might be an increased associative latency.) In an experiment (Exp. 7) devised to look into these matters, the list was presented once, twice, or three times, as shown in Fig. 22, to vary how well it was learned. In the one-presentation condition, at the bottom of the figure, the list was presented, and there followed a test stimulus and response, just as in Exp. 6. Again, the list changed from trial to trial, and contained from 3 to 7 digits. In the two-presentation condition, each trial included an additional presentation of its list and an attempt to recall it. In the three-presentation condition there was still another presentation and recall of the list.

Results from six subjects are shown in Fig. 23. At the bottom, the percentage of errors in naming the succeeding digit is shown as a function of list length, for each condition. Added presentations reduced the error rate by a factor of three. At the top of the figure, mean RT is shown as a function of list length, for each of the three conditions. Despite the change in level of learning indicated by the error data, the pattern of RTs shows no systematic change with number of presentations.

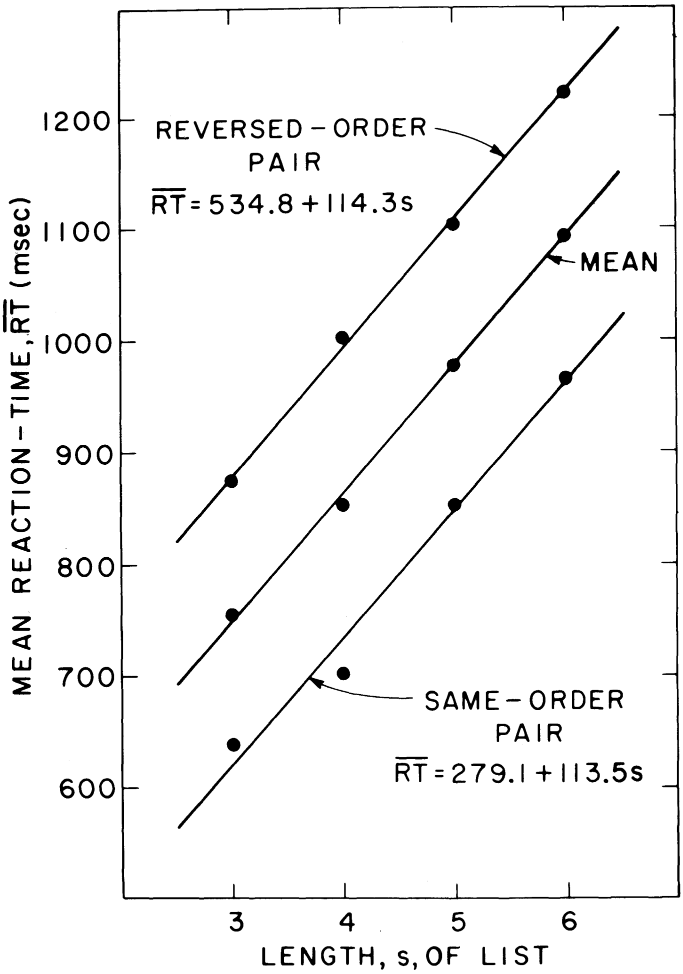


FIG. 25. Results of Exp. 8: Context-recognition. Averaged data from six subjects. Mean latencies of correct same-order and reversed-order responses, and their mean, with lines fitted by least squares.

This invariance indicates that differences in level of learning that are associated with list length do not contribute to the influence of list length on mean RT; and further, that within the limits of the experiment, the rate of scanning to locate is independent of how well a list has been learned.[14] The invariance with level of learning, which is similar to that of the high-speed scanning process over fixed-set and varied-set proce-

[14] If a list could only be either perfectly learned or not learned at all, this conclusion would not be justified, since restricting the latencies analyzed to those of correct responses would entail the selection of lists that had been learned to the same degree (perfectly) in the three conditions. This objection does not apply here, mainly because correct responses in conjunction with *partially*-learned lists were frequent.

dures, is consistent with the interpretation of the context-recall data presented in Sec. 11, and adds to the evidence that factors well known to influence learning may have no effect on active-memory functioning. Finally, given the invariance of the retrieval process, the strong influence of number of presentations on error rate suggests that the errors result primarily from faults in learning and retention, rather than in retrieval.

14. Recall versus Recognition of Contextual Information

In explaining the difference between findings from the item-recognition and context-recall tasks (Exps. 1–5 *versus* Exps. 6–7) I have emphasized that in one case the response depends merely on presence of an item in the list, and in the other case on its exact location. For an explanation in these terms to be valid, however, certain other differences between the tasks must be shown to be unimportant: one task involves recall, and the other recognition; one requires that for production of the response a memory representation be converted into a particular form—its name—and the other does not; and whereas the number of response alternatives in one task grows with list length, the other always requires a binary choice.

The last experiment to be described (Exp. 8) was designed to evaluate the importance of these factors and to examine further the generality of the process of scanning to locate. A recognition procedure was used to study the retrieval of contextual information; the resulting *context-recognition task* is shown in Fig. 24. On each trial the subject attempted to memorize a list of from 3 to 6 different digits, presented visually, one after another. To increase accuracy, the list was actually presented twice, with a recall attempt after the first presentation. The test stimulus was a pair of simultaneously presented digits that had appeared successively somewhere in the list. The subject's task was to decide whether the left-to-right order of the pair was the same as its temporal order in the list, or reversed. He made his response by pulling one of two levers (as in Exps. 1–5).

This experiment seemed to be somewhat risky, since there appeared to be a variety of strategies open to the subject. One possibility was that before its order could be tested, the pair might have to be located in the list by means of a scanning process. This process would be revealed by the relation between RT and the length of the list. Suppose that the test pair is located in the list by scanning for the location of one of its members, according to the self-terminating process described in Sec. 11. One would then expect that in the context-recognition task mean RT for both same-order and reversed-order responses would increase linearly with list length, and at equal rates, and that the rate of increase would be the same as in the context-recall task.

The same six subjects who performed the context-recall task of Exp. 7

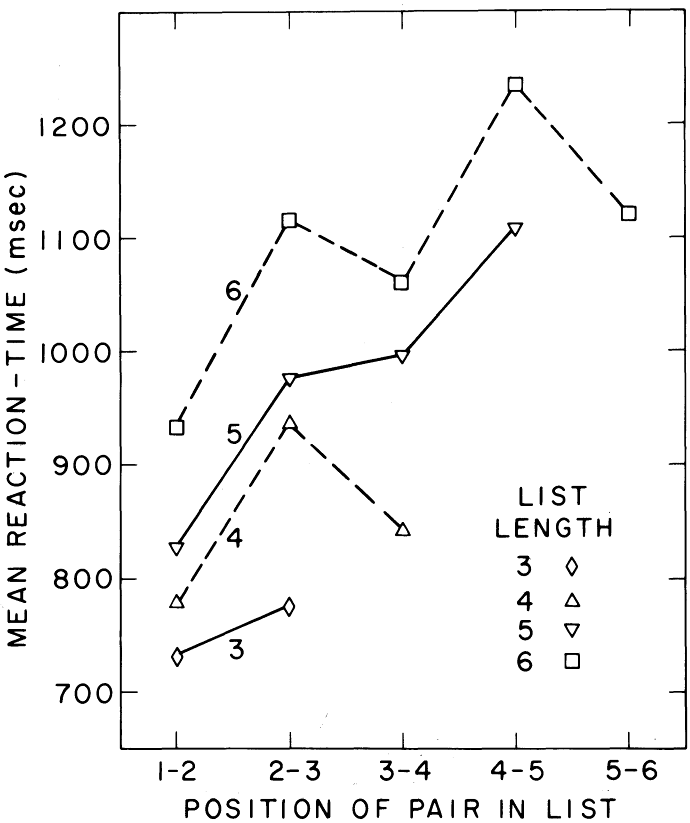


FIG. 26. Further results of Exp. 8: Context-recognition. Relation between mean RT of correct responses and serial position of the test pair in lists of four lengths. Data were averaged over six subjects and over same-order and reversed-order responses.

also served (in a balanced order) in the recognition task. RT-functions for both responses (Fig. 25) are linear, supporting the notion that in this task, also, performance involves a scanning process. For both responses the slope of the fitted line is 114 msec per item.[15] The equality of slopes is consistent with the idea that both responses depend on first locating one of the members of the pair in the list. That this is accomplished by means of a self-terminating process is suggested by the serial-position data: for all subjects, and for both responses, mean RT increased with the serial position of the pair in the list. Averaged serial-position data are shown in Fig. 26.

RT-functions from the context-recognition and context-recall tasks (Exps. 7 and 8) are compared in Fig. 27. The fitted lines are parallel,

[15] Although equal in slope, the RT-functions for the two kinds of response differ by about 250 msec in intercept. The several ways in which one might account for this difference are not discussed in this paper.

supporting the idea that the same search process (scanning to locate) underlies performance in both tasks. Also shown, for reference, is the RT-function from the item-recognition task of Exp. 1.

These parallel lines provide another striking instance of additive effects on mean RT. Here the additive factors are task (context-recognition *versus* recall) and list length; the absence of interaction indicates that these factors influence processing stages selectively, and helps to justify our interpretations of the data. Apparently, the change from recall to recognition does not influence the scanning stage, and, as assumed in Sec. 12, changes in list length do not influence perceptual and response stages.

One final substantive point about these results concerns their implications for the recognition-recall distinction. It is tempting to think that recognition involves less search, in some sense, than recall. These data reveal at least one search process that is as evident in a recognition task (Exp. 8) as in a recall task (Exp. 7).

Summary

I have reviewed informally eight experiments on the retrieval of information from human memory, whose interpretation depended on inferences from the structure of RT data to the organization of mental processes. The experiments have led to the discovery of two kinds of memory search that people use in the retrieval of information from short memorized lists. One is a high-speed exhaustive scanning process, used to determine the *presence* of an item in the list; the other is a slow self-terminating scanning process used to determine the *location* of an item in the list. Among other substantive implications of the experiments are: (1) Apparently one *must* scan a list serially to retrieve information from it, even when it is contained in active memory. There is no evidence in any of these data that one can "think about" more than one thing at a time, and thereby simultaneously compare a set of memorized items to a test item. (2) On the other hand, even a well-learned list can be made more readily available by being maintained in active memory. (3) Despite the possibility that retention may depend on a rehearsal process involving covert speech, visual rather than auditory memory-representations are used for comparison to representations of visual stimuli. (4) The same search process can be involved in both recall and recognition tasks.

Many of the inferences from the data were based on a proposal first made by Donders (1868) that the time between stimulus and response be regarded as the sum of the durations of a series of processing stages. Donders' *subtraction method* depends on this *stage theory*, together with an *assumption of pure insertion* which states that a change in the subject's task can cause the insertion of an additional processing stage without

altering the other stages. It was the questioning of this assumption, and the absence of any objective tests of its validity, that led to the decline of the subtraction method in the late nineteenth century.

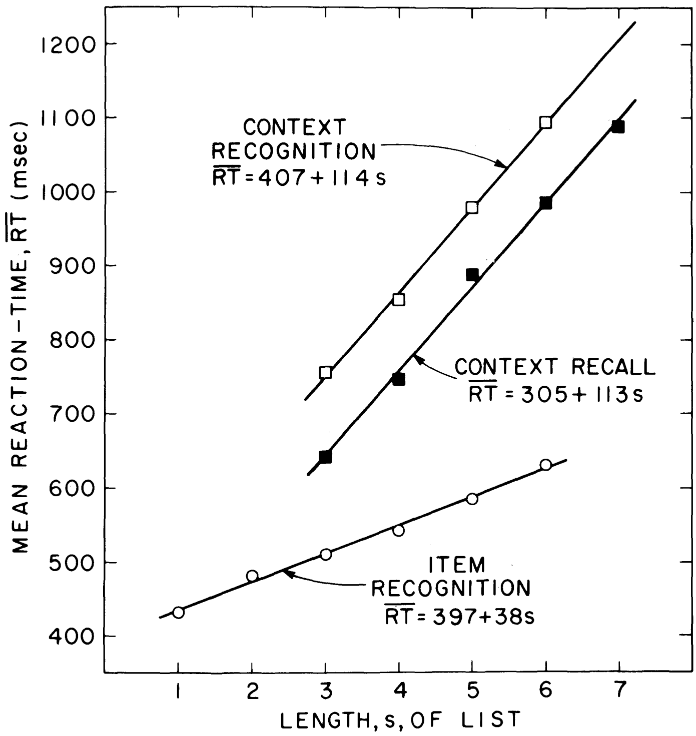


FIG. 27. Comparison of results from context-recognition, context-recall, and item-recognition tasks. Top: Mean RTs from Exp. 8 (Fig. 24), averaged over same-order and reversed-order responses. Middle: Mean RTs averaged over the three conditions in the context-recall task of Exp. 7 (Fig. 22), which used the same subjects as Exp. 8. Bottom: Mean RTs from Exp. 1 (Fig. 4).

The present paper advocates retaining the idea of stages of processing. But it shows how the insertion assumption can sometimes be replaced by a weaker *assumption of selective influence,* and how the validity of either assumption for a given experiment can be tested by determining whether the effects of experimental factors on RT are additive. The main ideas are: (1) if separate stages between stimulus and response have been correctly identified, then for each of these stages it may be easier to find a factor that *influences* it without altering other stages than to find one that *inserts* it without altering other stages; and (2) these factors would then have additive effects on mean RT. The discovery of several sets of such additive factors was critical in the interpretation of the experiments described.

REFERENCES

Atkinson, R. C. & Shiffrin, R. M. Human memory: a proposed system and its control processes. In K. W. Spence and J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory.* Vol. 2. New York: Academic Press, 1968. Pp. 89–195.

Baddeley, A. D. The influence of acoustic and semantic similarity on long-term memory for word sequences. *Quart. J. exp. Psychol.*, 1966, *18*, 302–309.

Broadbent, D. E. *Perception and communication.* London: Pergamon Press, 1958.

Cattell, J. McK. The perception time. *Mind*, 1886, *11*, 377–392. Reprinted in *James McKeen Cattell, Man of science.* Lancaster, Pa.: The Science Press, 1947. Pp. 64–79.

Chase, W. G. & Calfee, R. C. Modality and similarity effects in short-term recognition memory. *J. exp. Psychol.*, 1969, *81*, 510–514.

Cohen, R. L. & Johansson, B. S. Some relevant factors in the transfer of material from short-term to long-term memory. *Quart. J. exp. Psychol.*, 1967, *19*, 300–308.

Conrad, R. Acoustic confusions in immediate memory. *Brit. J. Psychol.*, 1964, *55*, 75–84.

Corballis, M. C. Serial order in recognition and recall. *J. exp. Psychol.*, 1967, *74*, 99–105.

Crowder, R. G. Short-term memory for words with a perceptual-motor interpolated activity. *J. verb. Learn. verb. Behav.*, 1967, *6*, 753–761.

Donders, F. C. Over de snelheid van psychische processen. Onderzoekingen gedaan in het Physiologisch Laboratorium der Utrechtsche Hoogeschool, 1868–1869, Tweede reeks, II, 92–120. Transl. by W. G. Koster in W. G. Koster (Ed.), *Attention and performance II. Acta Psychol.*, 1969, *30*, 412–431.

Egeth, H. E. Parallel versus serial processes in multidimensional stimulus discrimination. *Percept. & Psychophys.*, 1966, *1*, 245–252.

Glanzer, M. & Cunitz, A. R. Two storage mechanisms in free recall. *J. verb. Learn. verb. Behav.*, 1966, *5*, 351–360.

Hochberg, J. In the mind's eye. In R. N. Haber (Ed.), *Contemporary theory and research in visual perception.* New York: Holt, Rinehart & Winston, 1968. Pp. 309–331.

Hohle, R. H. Component process latencies in reaction times of children and adults. In L. P. Lipsett and C. C. Spiker (Eds.), *Advances in child development and behavior.* Vol. 3. New York: Academic Press, 1967. Pp. 225–261.

Jastrow, J. *The time-relations of mental phenomena. Fact & theory papers No. VI.* New York: N.D.C. Hodges, 1890.

Johnson, D. M. *The psychology of thought and judgment.* New York: Harper, 1955.

Külpe, O. *Outlines of psychology.* New York: MacMillan, 1895.

Landauer, T. K. Rate of implicit speech. *Percept. mot. Skills*, 1962, *15*, 646.

Landauer, T. K. & Freedman, J. L. Information retrieval from long-term memory: category size and recognition time. *J. verb. Learn. verb. Behav.*, 1968, *7*, 291–295.

McGill, W. J. & Gibbon, J. The general gamma distribution and reaction times. *J. math. Psychol.*, 1965, *2*, 1–18.

McMahon, L. E. Grammatical analysis as part of understanding a sentence. Unpublished doctoral dissertation, Harvard University, 1963.

Miller, G. A. *Psychology, the science of mental life.* New York: Harper & Row, 1962.

Morin, R. E., DeRosa, D. V., & Stultz, V. Recognition memory and reaction time. In A. F. Sanders (Ed.), *Attention and performance. Acta Psychol.*, 1967, *27*, 298–305.

Morin, R. E., DeRosa, D. V., & Ulm, R. Short-term recognition memory for spatially isolated items. *Psychon. Sci.*, 1967, *9*, 617–618.

Neisser, U. *Cognitive psychology.* New York: Appleton-Century-Crofts, 1967.

Newell, A. & Simon, H. A. Computers in psychology. In R. D. Luce, R. R. Bush and E. Galanter (Eds.), *Handbook of mathematical psychology.* Vol. 1. New York: Wiley, 1963. Pp. 361–428.

Nickerson, R. S. Categorization time with categories defined by disjunctions and conjunctions of stimulus attributes. *J. exp. Psychol.*, 1967, *73*, 211–219.

Posner, M. I. Short-term memory systems in human information processing. In A. F. Sanders (Ed.), *Attention and performance. Acta Psychol.*, 1967, *27*, 267–284.

Posner, M. I., Boies, S. J., Eichelman, W. H., & Taylor, R. L. Retention of

visual and name codes of single letters. *J. exp. Psychol. Monogr.*, 1969, *79*, No. 1, Part 2, 1–16.
POSNER, M. I. & MITCHELL, R. F. Chronometric analysis of classification. *Psychol. Rev.*, 1967, *74*, 392–409.
POSNER, M. I. & ROSSMAN, E. Effect of size and location of informational transforms upon short-term retention. *J. exp. Psychol.*, 1965, *70*, 496–505.
RESTLE, F. & DAVIS, J. H. Success and speed of problem solving by individuals and groups. *Psychol. Rev.*, 1962, *69*, 520–536.
ROCK, I. A neglected aspect of the problem of recall: The Höffding function. In J. M. Scher (Ed.), *Theories of the mind.* New York: Free Press, 1962. Pp. 645–659.
SANDERS, A. F. Rehearsal and recall in immediate memory. *Ergonomics*, 1961, *4*, 25–34.
SLAMECKA, N. J. Serial learning and order information. *J. exp. Psychol.*, 1967, *74*, 62–66.
SMITH, E. E. Effects of familiarity on stimulus recognition and categorization. *J. exp. Psychol.*, 1967, *74*, 324–332.
SMITH, E. E. Choice reaction time: an analysis of the major theoretical positions. *Psychol. Bull.*, 1968, *69*, 77–110.
SPERLING, G. The information available in brief visual presentations. *Psychol. Monogr.*, 1960, *75* (11, Whole No. 498).
SPERLING, G. A model for visual memory tasks. *Hum. Factors*, 1963, *5*, 19–31.
STERNBERG, S. Stochastic learning theory. In R. D. Luce, R. R. Bush and E. Galanter (Eds.), *Handbook of mathematical psychology.* Vol. 2. New York: Wiley, 1963. Pp. 1–120.
STERNBERG, S. Estimating the distribution of additive reaction-time components. Paper presented at the meeting of the Psychometric Society, Niagara Falls, Ont., October 1964.
STERNBERG, S. High-speed scanning in human memory. *Science*, 1966, *153*, 652–654.
STERNBERG, S. Retrieval of contextual information from memory. *Psychon. Sci.*, 1967, *8*, 55–6. (a)
STERNBERG, S. Two operations in character-recognition: Some evidence from reaction-time measurements. *Percept. & Psychophys.*, 1967, *2*, 45–53. (b)
STERNBERG, S. The discovery of processing stages: Extensions of Donders' method. In W. G. Koster (Ed.), *Attention and performance II. Acta Psychol.*, 1969, *30*, 276–315.
SUPPES, P. & GROEN, G. Some counting models for first-grade performance data on simple addition facts. Technical Report 90, Institute for Mathematical Studies in the Social Sciences, Stanford University, 1966.
TAYLOR, D. H. Latency components in two-choice responding. *J. exp. Psychol.*, 1966, *72*, 481–488.
WAUGH, N. C. & NORMAN, D. A. Primary memory. *Psychol. Rev.*, 1965, *72*, 89–104.
WICKELGREN, W. A. Auditory or articulatory coding in verbal short-term memory. *Psychol. Rev.*, 1969, *76*, 232–235.
WOODWORTH, R. S. *Experimental psychology.* New York: Holt, 1938.
WUNDT, W. *Grundzüge der physiologischen Psychologie*, Vol. II, 2nd ed. Leipzig: Engelmann, 1880.
YONAS, A. The acquisition of information-processing strategies in a time-dependent task. Unpublished doctoral dissertation, Cornell University, 1969.