

Journal of Experimental Psychology

GENERAL

VOL. 104, No. 2

JUNE 1975

Search Processes for Associative Structures in Long-Term Memory

Richard C. Mohs, Keith T. Wescourt, and Richard C. Atkinson
Stanford University

SUMMARY

Two experiments were conducted using a symbol-element recognition task. In both experiments subjects first learned six separate lists of words. Each list was labeled at the top by a different consonant letter (the symbol), and the number of words (elements) on a list varied from two to six. None of the lists had any words in common. After the lists had been thoroughly mastered, the subject was given a series of test trials involving the presentation of a consonant-word pair. Subjects gave a positive response if the test word was a member of the list designated by the test consonant and a negative response if the test word was from one of the other lists. A negative trial always involved a consonant and a word that were members of the original lists but were not associated with each other during learning. Characteristics of the memory search processes involved in verifying symbol-element relationships were examined by analyzing the relationship between reaction time (RT) and relevant set size variables. In the symbol-element task two set size variables are of particular interest: (a) the symbol set size, t , which specifies the number of words on the list associated with the test symbol, and (b) the element set size, d , which specifies the number of words on the list from which the test word is drawn. On a positive trial both the symbol and element are from the same list; hence, t necessarily equals d . On a negative trial t and d vary independently and may or may not be equal, depending on the particular consonant-word pair.

In Experiment 1 subjects were tested for 4 consecutive days on the same set of lists. On positive trials RT increased in a remarkably linear fashion as set size ($t = d$) varied from two to six elements. The slope of the function

was approximately 160 msec on the 1st day, but decreased to a slope of approximately 50 msec by the 4th day. On negative trials RT increased in a curvilinear fashion with both t and d ; as in the case of positive trials, the size of these increases declined over test days.

In Experiment 2 subjects were tested on only one day; one group of subjects learned the symbol-element relationships as a set of lists (as in Experiment 1), but the other group learned them as word-consonant paired associates. The latter group learned the same symbol-element information, but as individual paired associates rather than as lists of words associated with specific symbols. Both the list-learning and paired-associate groups displayed the same general relationships between RT and the set size variables t and d ; however, the magnitude of the set size effects were greatly reduced in the paired-associate condition. Further, overall response speed was greater for the paired-associate group than for the list-learning group.

The predictions of several memory search models were evaluated in terms of these findings. The model that provided the best account of the results assumes that subjects simultaneously access representations of both the test symbol and test element on each trial. Search processes are then initiated from both representations; whichever search terminates first determines RT for that trial. The model provides a reasonably accurate account of both RT measures and error probabilities. The differences produced by extended testing in Experiment 1 and by manipulating learning strategies in Experiment 2 are explained in terms of the way the symbol-element relationships are represented in memory. ■

■ In studies of long-term recognition memory for words, subjects typically learn a list and subsequently are asked to decide whether test words are old or new; that is, whether they were on the study list or are occurring in the experimental context for the first time. There has been disagreement about how to represent the memory processes that underlie such recognition decisions. One point of view is that the presence of the test stimulus provides a basis for "direct access" to a memory location where information about the stimulus is stored. This information is then evaluated to make the recognition decision (e.g., Kintsch, 1970; Murdock, 1968; Postman, Jenkins, & Postman, 1948). An alternative position is that the study list is encoded as a structure in memory and that recognition involves locating and searching the structure for the presence of the test word's internal representation (e.g., Anderson & Bower, 1974; Shiffrin & Atkinson, 1969; Tulving & Thomson, 1971). These opposing viewpoints have drawn support from inconsistent and sometimes contradictory empirical results concerning the effects of list structure and composition on long-term recognition performance. For example, the frequent failure of recognition studies to find effects of list length and internal list organization (effects that occur consistently in recall tasks) has been used as support for the "direct access" models of recognition (e.g., Kintsch, 1968). Under some conditions, however, these effects have been observed in recognition tasks and taken as evidence in support of memory search models (e.g., Jacoby, 1972; Mandler, 1972; Mandler, Pearlstone & Koopmans, 1969); in particular, effects of organization have been found rather consistently in studies that use reaction time (RT) rather than accuracy as the primary dependent measure (e.g., Homa, 1973; Seamon, 1973).

This research was supported by Grant MH-21747 from the National Institute of Mental Health.

Requests for reprints should be sent to Richard C. Atkinson, Department of Psychology, Stanford University, Stanford, California 94305.

Recently, Atkinson and Juola (1973, 1974) have proposed that performance in recognition tasks depends on both direct access and memory search processes. Their theory recognizes that two questions are implicit in these tasks and that recognition decisions depend on the ability of the memory system to answer either or both questions: (a) Is there an internal representation contained in the memory structure for the study list that matches the representation of the test word? (b) How recently has the test word been processed in memory—was it so long ago that it could not have been on the study list or so recently that the probability is high that it was processed while learning the list? The answer to the first question involves verification of list membership, whereas the answer to the second question depends on inferences based upon knowledge of stimulus recency and task constraints.

In the Atkinson and Juola model, the encoding of stimulus words involves a content-addressable search (similar to direct access) for a location in lexical store, a functional partition of the long-term store (LTS). Each memory location is assumed to have an activity level that is a function of the frequency and recency of past access to it. In the context of a recognition task, a unidimensional familiarity measure can be derived from the activity level of a location in lexical store and sometimes can provide a basis for deciding whether to call a test word old or new. The alternative means for making the recognition decision is to search a stored representation of the study list in the event-knowledge store (EKS), a second partition of LTS, for a match to the internal coding of the test word.

Atkinson and Juola (1973, 1974) apply their model to recognition experiments in which the study lists are well memorized and RT is the principal dependent measure. Under these circumstances, they propose that the familiarity of the test word is used as a basis for decisions whenever possible because such decisions are faster than those involving a search in EKS; decisions are based on EKS search only when familiarity

values fall within a range that would lead to unacceptable error rates. In experimental situations involving study lists that are not well learned and where accuracy is the dependent measure (e.g., Shepard & Teghtsoonian, 1961) it may be assumed that any information available in memory is used in making the recognition decision. In this case familiarity could be used when retrieval from EKS fails because information was not stored or because search processes fail to locate it. The role of familiarity in recognition, therefore, depends on both the parameters of the learning situation and the requirements of the test context.

The experiments described by Atkinson and Juola are primarily concerned with factors affecting the use of familiarity in RT recognition tasks, particularly experimental and extra-experimental relations between words on study lists and distractor words. This article is concerned with how memory structures in EKS are searched in RT recognition tasks; that is, how the list membership of words is ascertained. Because traditional methods of recognition typically involve performance that is based on a mixture of underlying memory processes, we consider here a somewhat different recognition task. We believe that this task has the advantage of eliminating familiarity-based decisions, thereby maximizing reliance upon searches of EKS for successful performance.

The recognition paradigm that we consider can be represented in terms of relations between two sets of stimuli, a set of symbols and a set of elements. A subset of elements is assigned to each symbol with the restriction that both sets are exhausted in the process. The subset of elements associated with symbol S_i will be denoted as $\{e_{i,1}, e_{i,2}, e_{i,3}, \dots, e_{i,n(i)}\}$ where $n(i)$ is the size of the subset associated with symbol S_i . Note that the first subscript on e specifies the symbol to which it is assigned, whereas the second subscript denotes its serial position in the subset of elements.

In the experimental task subjects first learn the assignments between symbols and subsets of elements. They then perform on a series of test trials. On each trial a

symbol and an element are presented; the subject's task is to give a positive response if the element is a member of the subset associated with the symbol and a negative response otherwise. The symbol and element presented on each trial are always members of the original stimulus sets. A negative test trial involves a symbol and element that were learned in the first phase of the experiment but were not associated with each other; rather, the element on a negative trial is a member of a subset associated with a symbol other than the test symbol. In these experiments, then, the stimulus components are equally recent on both positive and negative test trials, eliminating the possibility of familiarity-based decisions.¹ The primary dependent measure is RT, though accuracy may be of interest when the parameters of the experimental situation increase error rates to a level at which differences between conditions can be determined.

Among the factors in this recognition paradigm that could have some bearing on theories of memory search are the degree of overlap in symbol-element assignments (i.e., the number of symbols a given element is associated with), the method of learning the symbol-element associations, and the temporal and spatial characteristics of the test ensemble. However, inferences about memory search processes depend primarily on the effects of the relevant set sizes on each trial. We will let two parameters characterize the RT-set size function. If the test ensemble consists of symbol S_i and element $e_{j,k}$, then: (a) the symbol set size, t , is the number of elements in the set associated with S_i , namely $n(i)$; and (b) the element set size, d , is the number of elements in the subset of which $e_{j,k}$ is a member, namely $n(j)$. On positive trials the symbol set and the element set are the same, so $t = d$ necessarily. On negative trials $t = d$ only if there are at least two symbols, S_a and S_b , with $n(a) = n(b)$. As an

¹ Familiarity-based decisions could only occur in the present paradigm if one were willing to suppose that content-addressable retrieval could be developed for arbitrary symbol-element units in the course of an experiment.

example, consider a situation in which there are three symbols with related element sets of size $n(1) = n(2) = 2$ and $n(3) = 4$:

$$\begin{aligned} S_1 &= \{e_{1,1}, e_{1,2}\}, \\ S_2 &= \{e_{2,1}, e_{2,2}\}, \\ S_3 &= \{e_{3,1}, e_{3,2}, e_{3,3}, e_{3,4}\}. \end{aligned}$$

Instances of positive test ensembles in which $t = d = 2$ are $S_1 \leftrightarrow e_{1,2}$ and $S_2 \leftrightarrow e_{2,1}$; the combination $S_3 \leftrightarrow e_{3,3}$ is a positive trial in which $t = d = 4$. Examples of negative trials are: $S_1 \leftrightarrow e_{3,1}$, where $t = 2$ and $d = 4$; $S_3 \leftrightarrow e_{2,2}$, where $t = 4$ and $d = 2$; and $S_2 \leftrightarrow e_{1,1}$, where $t = 2$ and $d = 2$.

Although the symbol-element task has been described here as a modification of a standard item-recognition paradigm, it also may be viewed as a fact verification experiment. During the study phase of a symbol-element experiment, the subjects memorize a set of facts concerning permissible symbol-element relationships; during the test phase they are asked to verify whether particular combinations of symbols and elements are true or false with regard to the memorized information. Viewed in this way, the symbol-element task is comparable to some of the sentence verification experiments reported by Anderson and Bower (1973). In a typical experiment of this type, the subject first memorizes a list of sentences in which the subject or predicate of any given sentence may be repeated in other sentences on the list. After the sentences have been memorized a series of test sentences are presented. Some are from the memorized list; others are new but are formed by re-pairing the subjects and predicates from sentences in the memorized list. The person being tested is asked to respond as quickly as possible, saying true if the test sentence was from the memorized list and false otherwise. If the subject of a sentence is equated to our symbol and the predicate to our element, then the sentence verification experiments have the same formal structure as the symbol-element task. The sentence verification task can be regarded as a special case of the symbol-element paradigm that employs a different type of stimulus materials; hopefully the

same theoretical analyses are applicable in both situations.

It is tempting to draw another analogy between the symbol-element task and semantic memory experiments. In semantic memory experiments, for example, the subjects must verify propositions such as *A canary is a bird* or *A canary has bones*. Subjects do not memorize a specific set of facts prior to the experiment in these tasks; rather, performance depends upon their ability to retrieve certain naturally acquired information and to draw logical inferences from it (Anderson & Bower, 1973; Atkinson, Herrmann, & Wescourt, 1974; Quillian, 1968; Rumelhart, Lindsay, & Norman, 1972; Smith, Shoben, & Rips, 1974). The uncertain roles of inference and memory retrieval in a semantic memory experiment make comparisons between that task and the symbol-element experiment of questionable value. Certainly, models developed to explain symbol-element experiments will not be general enough to account for semantic memory data.

This paper describes two experiments that were designed to evaluate several possible models for the symbol-element paradigm. In both experiments single consonants were used as symbols and words as elements; there were six symbols, and subsets contained either two, four, or six elements with two subsets of each size. Each element occurred in only one subset. Experiment 1 investigated the effects of t and d on RT and how these effects changed with extended testing and with different test schedules for presenting symbols and elements. Subjects learned the symbol-element relationships as six separate word lists, with each list designated by an arbitrary consonant. They participated in four test sessions, using the same lists throughout. During a test session either symbols or words were tested equally often as a between-subjects factor in order to investigate whether set size effects are sensitive to stimulus frequency. Experiment 2 controlled the method of initial acquisition of symbol-element relationships. In order to minimize the effects of idiosyncratic learning strategies, subjects learned the symbol-

element relationships in an experimenter-controlled situation just prior to testing. One group of subjects learned the materials as six word lists named by consonants, as in Experiment 1. The other group learned them as a single list of word-consonant paired associates.

EXPERIMENT 1

Method

Subjects. The subjects were 18 female students at Stanford University. Each subject was paid \$10 for participating in four test sessions.

Apparatus. The experiment was run on a programmable display system (Imlac Corporation PDS-1) interacting with a PDP-10 time-sharing computer system. The test stimuli appeared in green capital letters (11 characters/in. [4.3 characters/cm]) on a dark gray cathode-ray tube (CRT) screen and were viewed from a distance of approximately 18 in. (46 cm). A typewriter-like keyboard with microswitch keys was located in front of the CRT screen, and subjects made either a yes or a no response on each trial by striking one of two specified keys on the keyboard.

Materials. Test stimuli were constructed from combinations of 24 words and six consonants (G, R, X, M, B, and L). The 24 words, each with six letters and two syllables, were selected so that acoustic, visual, and semantic confusions among the words were minimized. Two sets of six lists (Set 1 and Set 2) were constructed from the 24 words and six consonants. In each set two lists contained 6 words, two lists contained 4 words, and two lists contained 2 words. Each of the six consonants was used as a symbol for one list in each set. For each set of lists, words and consonants were assigned to lists with the constraint that no consonant was used as a symbol for a list of the same size in both sets. Within each set, words were assigned to lists so that the mean frequency of occurrence of words was between 105 and 121 per million for each list (Carroll, Davies, & Richman, 1971). Half the subjects learned, and were tested on, each set of lists.

On each test trial the subject was shown a consonant-word pair and was required to respond yes if the test word was a member of the list designated by the test consonant and no if the test word was from one of the other lists. Trial types were defined in terms of the response required on a trial, the size of the list designated by the test consonant (t), and the size of the list from which the test word was drawn (d). Trials requiring a positive response necessarily had equal values of t and d . Nine trial types requiring a negative response were produced from all possible combinations of the three values of t and d . A list of the twelve trial types is given in the left-hand column of Table 1.

For each set of lists six blocks of 48 trials were constructed for each of the two test schedules. Symbols-equal blocks were constructed so that on both positive and negative trials each of the six test symbols appeared equally often. Across six of these blocks each word from a Size 2 list was presented twice as often as each word from a Size 4 list and three times as often as each word from a Size 6 list. Elements-equal blocks were constructed so that on both positive and negative trials each word (regardless of the size of the list from which it was drawn) appeared equally often. Thus, in an elements-equal trial block each consonant used to designate a list of Size 6 was presented 12 times, each consonant used to designate a list of Size 4 was presented 8 times, and each consonant used to designate a list of Size 2 was presented 4 times.

For both test schedules equal numbers of positive and negative trials were presented in each block. Across the six symbols-equal blocks, conditional probabilities for trial types were as follows: (a) For positive trials t and d were equal to each other with values of 2, 4, or 6, each with a probability of 1/3; (b) for negative trials t was equal to 2, 4, or 6, each with a probability of 1/3; and (c) for negative trials and a fixed value of t , the value of d was 2, 4, or 6, each with a probability of 1/3. For positive trial types these same conditional probabilities were followed within each block. Thus, for example, each symbols-equal block contained eight positive trials in which $t=d=2$ and eight positive trials in which $t=d=4$. For negative trial types these conditional probabilities could not be followed exactly but were followed as closely as possible within each block. For example, each block contained either two or three negative trials in which $t=2$ and $d=6$, and either two or three trials in which $t=4$ and $d=4$. Conditional probabilities for trial types across the six elements-equal blocks were as follows: (a) For positive trials t and d (equal to each other) were 2, 4, or 6 with probabilities of 1/6, 1/3, and 1/2, respectively; (b) for negative trials t was 2, 4, or 6 with probabilities of 1/6, 1/3, and 1/2; and (c) for negative trials and a fixed value of t , the value of d was 2, 4, or 6 with probabilities of 1/6, 1/3, and 1/2. For positive trial types these conditional probabilities were followed within each block, and for negative trial types they were followed as closely as possible. For example, within each elements-equal block there were four positive trials in which $t=d=2$, eight positive trials in which $t=d=4$, two negative trials in which $t=2$, and $d=6$, and either two or three negative trials in which $t=4$, and $d=4$. For both test schedules the symbol-element pairs used as test stimuli for each trial type were selected so that, across the six blocks, each pair that could serve as a stimulus for a given trial type appeared approximately the same number of times. In addition, each consonant and each word appeared on an equal num-

ber of positive and negative trials across the six trial blocks.

The two test schedules were presented on alternate days for 4 days. On each day a subject received six blocks of test trials. On Day 1 four subjects who learned Set 1 lists and five subjects who learned Set 2 lists received symbols-equal test blocks; elements-equal test blocks were given to five subjects who learned Set 1 lists and four subjects who learned Set 2 lists. On Day 2 each subject received the test schedule that she had not received on Day 1. On Days 3 and 4 subjects received the same trials that they had received on Days 1 and 2, respectively. Trials within each block were randomized separately for each subject on each test day.

Procedure. At least 18 hr before the first test session each subject picked up a set of six 5×8 in. (12×20 cm) index cards. One list appeared on each of these cards and was printed in a column below the consonant used to designate that list. The subject was instructed to memorize the lists so that, when shown one of the six consonants, she could recall the appropriate list in the order given on the card.

The subjects were tested individually. At the start of the first test session each subject was given the six test consonants and required to give a written and then an oral recall of the words designated by each consonant. All subjects were able to give each of the lists in correct serial order with no errors on both recall tests.

After completing both recalls the subject was seated in front of the CRT screen and given instructions about the task. The subject was told that there would be a series of test trials and that the following sequence would occur on each trial: (a) The word *ready* would appear on the screen; (b) the subject would then press the spacebar on the keyboard to start the test sequence; (c) after a delay of 500 msec a fixation point (an asterisk) would appear at the center of the screen and remain there for 850 msec; (d) the fixation point would disappear and the screen would be blank for 150 msec; (e) the test stimulus (a consonant one line above and one character to the right of the fixation point and a word one line below and starting one character to the right of the fixation point) would then appear; (f) the test stimulus would remain visible until the subject made a yes or a no response by pressing either the M or the C key on the keyboard; and (g) a feedback statement indicating whether the response was correct or not would follow the response and remain visible for $2\frac{1}{2}$ sec. The subject was told to respond yes if the test word was on the list designated by the test symbol, and no if the test word was from one of the other lists. The instructions emphasized that the subject was to respond as quickly as possible, while trying to avoid errors. The keys M and C were paired with responses so that each subject made positive responses with her preferred hand.

If at any time the subject pressed a key other than M, C, or the spacebar, the words *illegal key-press* appeared on the CRT screen. Subjects received about 7 trials/min; there were three 45-sec rest periods during the session, one after each 72 trials.

When the experimenter was sure that the subject understood the task a series of 16 practice trials was given. On each of these trials a single digit appeared to the right of the fixation point and the subject was to respond yes if the digit was a 5 and no if it was some other digit. In all other respects the sequence of events followed on these trials was identical to that followed on the regular test trials. After the practice trials the procedure was reviewed and the 288 test trials were given.

The four test sessions were scheduled on consecutive days for each subject. At the start of the second, third, and fourth test sessions, the subject was again required to give an oral recall of each list in correct serial order. The instructions were then reviewed and that day's trials presented. No practice trials were given after the first day.

Results

A preliminary analysis of the data indicated that the pattern of results remained fairly constant across the 4 test days, but that there were large practice effects from Day 1 to Day 3. Since there were only minimal practice effects after Day 3, the data from Days 3 and 4 were pooled and all further analyses were performed only on data from Day 1 and data from Days 3 and 4 combined. Reaction times from trials on which an incorrect response was made were not included in the analysis.

A summary of the data is presented in Table 1. At the left of Table 1 are the values of t and d for the 12 trial types. For each subject the mean and standard deviation of the RTs for each of the 12 trial types were computed. Outlying scores for each subject were eliminated from computations by the following method: (a) For each trial type, the mean was calculated, (b) each score that was more than 2.5 times the mean was deleted, and (c) the mean was recalculated from the remaining scores. Less than 1% of the scores were eliminated in this way. The mean of the subject means and the mean of the subject standard deviations for each trial type are presented in Table 1. Also shown for each trial type are the total number of observations and the

TABLE 1
 MEAN OF SUBJECT MEAN REACTION TIMES, MEAN OF SUBJECT STANDARD DEVIATIONS,
 NUMBER OF OBSERVATIONS, AND PERCENT ERRORS FOR EACH TRIAL TYPE IN
 EXPERIMENT 1

Response and values of <i>t</i> and <i>d</i>	Day 1				Days 3 and 4 (pooled)			
	\overline{RT} (msec)	\overline{SD} (msec)	No. obs.	% Errors	\overline{RT} (msec)	\overline{SD} (msec)	No. obs.	% Errors
Yes								
2 2	1,239	415	648	3.2	832	196	1,296	2.5
4 4	1,574	598	864	7.5	968	271	1,728	4.2
6 6	1,847	728	1,080	9.0	1,025	323	2,160	4.0
No								
2 2	1,605	439	180	5.9	1,084	282	360	9.4
4 2	1,713	582	216	5.2	1,039	293	432	2.3
6 2	1,753	700	252	2.6	1,042	304	504	1.8
2 4	1,528	477	216	3.5	979	251	432	1.7
4 4	1,952	695	288	13.5	1,221	377	567	8.6
6 4	1,854	683	360	5.9	1,173	349	720	3.9
2 6	1,517	529	252	2.3	989	279	504	1.4
4 6	2,049	782	360	7.7	1,128	302	720	3.2
6 6	2,088	805	468	12.9	1,268	367	936	9.3

percent errors averaged over the 18 subjects.

Graphs of the mean RTs from Day 1 and Days 3 and 4 are presented in Figure 1. The means in both panels are plotted as a function of *t*. The four lines in each panel represent predictions generated by a model presented later in the article. The line labeled *Yes* represents the predicted RTs for positive trials; three lines, each labeled with *No* and a value of *d* represent predicted RTs for negative trials. The error rates for positive and negative trials for each value of *t* are represented by bars at the bottom of Figure 1. Adjacent to each of the bars representing an observed error rate is a hatched bar that represents the predicted error rate for that condition.

For both Day 1 and Days 3 and 4, separate analyses of variance were performed on the RT data from positive and negative trials; in each of the four analyses subject means were used as scores. The analyses for trials on which a yes response was made included three factors: (a) the value of *t* (2, 4, or 6) as a within-subjects factor, (b) test schedules (symbols equal vs. elements equal) as a between-subjects factor, and (c) list sets (Set 1 vs. Set 2) as a between-

subjects factor.² For trials on which a no response was made the analyses included these same three factors plus the value of *d* (2, 4, or 6) as a second within-subjects factor.

The analysis for positive trials on Day 1 indicated that the main effect of *t* was highly significant, $F(2, 28) = 28.46$, $p < .001$; both of the main effects due to between-subjects variables and all of the interactions, however, were nonsignificant ($p > .10$ in all cases). For negative trials on Day 1 the analysis indicated that there were significant main effects due to *t*, $F(2, 28) = 19.65$, $p < .001$, and *d*, $F(2, 28) = 8.53$, $p < .005$, and that the interaction of *t* and *d* was significant, $F(4, 56) = 6.64$, $p < .001$; both of

² In this and all subsequent analyses materials are treated as fixed effects. In general, statistical tests of the type suggested by Clark (1973) are necessary to assess the generalizability of findings to a population of language materials. Because of the complexity of the present experiments, stimulus pairs were not included as a factor in the experimental design, and therefore the most appropriate statistical tests were not performed. However, subsequent experiments using other sets of consonants and words as symbols and elements have produced results similar to those reported here, so it does not seem that generalizability across materials is a critical problem (Appelman & Atkinson, 1975).

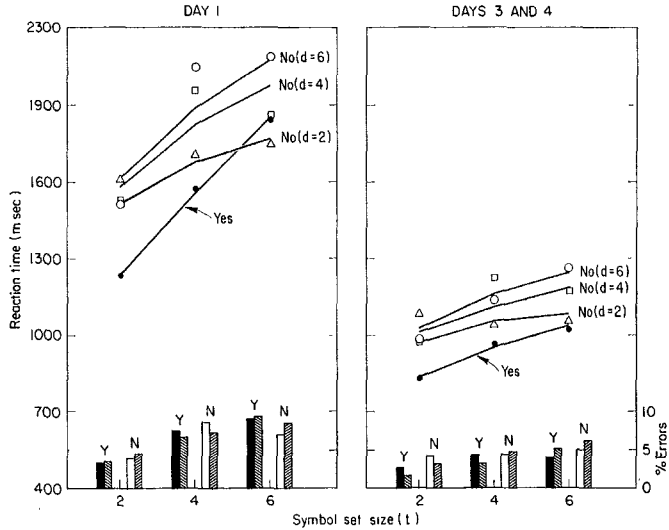


FIGURE 1. Mean reaction times and error rates for Day 1 and Days 3 and 4. (The means for positive trials are represented by filled circles; means for negative trials that had a d value of 2, 4, or 6 are represented by open triangles, open squares, and open circles, respectively. Error rates for positive trials and for negative trials [collapsed across the three values of d] are represented by solid and open bars, respectively. The lines and hatched bars represent predicted values.)

the main effects due to between-subjects factors and all of the other interactions were nonsignificant ($p > .10$ in all cases).

The analyses for Day 1 indicate that mean RT increased with t on positive trials and that on negative trials mean RT was a function of both t and d . From Figure 1 it is clear that the function relating mean RT to t for positive trials was nearly linear but that the three functions relating mean RT to t for negative trials showed marked departures from linearity. In fact, subsequent tests showed that, for positive trials, 99.7% of the variance due to t was accounted for by a linear contrast; for negative trials both the linear, $F(1, 14) = 39.42$, $p < .001$, and quadratic, $F(1, 14) = 8.18$, $p < .025$, components of the main effect due to t were significant. It should also be noted that negative response time was not, in all cases, a monotonically increasing function of t and d . In particular, RT for the $t = d = 2$ condition was slower than either the RT for the $t = 2$, $d = 4$ condition or the $t = 2$, $d = 6$ condition and the RT for the $t = d = 4$

condition was slower than the RT for the $t = 6$, $d = 4$ condition.

The fact that there were no significant effects involving the test schedules indicates that the changes in RT associated with t and d cannot be attributed to a confounding of t and d values with the probability of presenting particular test stimuli. For example, in the symbols-equal test schedule each word from a list of Size 2 appeared three times as often as each word from a list of Size 6, and subjects might have responded more quickly to test words from the short lists for that reason. In the elements-equal test schedule all of the test words appeared equally often, yet this schedule produced results almost identical to those produced by the symbols-equal schedule. The fact that there were no significant differences between subjects who learned different sets of lists suggests that the obtained effects were not related to idiosyncratic properties of the test materials.

The analyses performed on the data from Days 3 and 4 were similar to those per-

formed on the data from Day 1; for Days 3 and 4, however, the test schedules variable divided subjects in the analysis according to the order in which the two schedules were given (symbols equal on Day 3 and elements equal on Day 4 or vice versa). For positive trials on Days 3 and 4 the analysis showed that there was a significant main effect due to t , $F(2, 28) = 22.29$, $p < .001$, and a significant effect due to the interaction of t with the list set that the subjects had learned, $F(2, 28) = 7.02$, $p < .01$; all other main effects and interactions were nonsignificant ($p > .05$ in all cases). Separate examinations of the data from subjects in the two list conditions indicated that the linear function relating mean RT to t had a smaller slope for one group than for the other; in all other respects the data from the two groups were similar. For negative trials the analysis indicated that there were significant main effects due to t , $F(2, 28) = 19.67$, $p < .001$, and d , $F(2, 28) = 5.15$, $p < .025$; in addition, the $t \times d$ interaction, $F(4, 56) = 20.17$, $p < .001$, the $t \times$ Lists

interaction, $F(2, 28) = 3.63$, $p < .05$, and the $d \times$ Lists interaction, $F(2, 28) = 4.88$, $p < .025$, were significant. All of the other main effects and interactions were nonsignificant ($p > .05$ in all cases). Again, separate examinations of the data from subjects in the two list conditions indicated that the pattern of results was similar for the two groups; the effects due to t and d were smaller, however, for subjects who learned one set of lists than for subjects who learned the other set.

In summary, then, the results for Days 3 and 4 were similar to the results for Day 1. As is clear from Figure 1, however, the differences due to t and d were much smaller on Days 3 and 4 than they were on Day 1. In addition, the tendency noted on Day 1 for negative responses to be fast on trials with unequal values of t and d was also evident on Days 3 and 4.

Figure 2 presents the RT data as a function of the serial position of the test word; that is, the data are categorized according to the position that the test word occupied

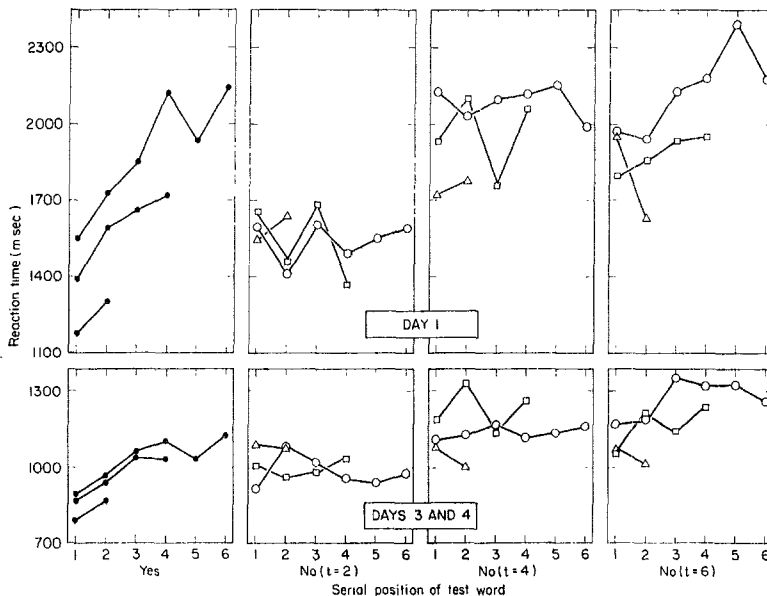


FIGURE 2. Serial position data for Day 1 and Days 3 and 4. (The data from positive trials on which test words were drawn from lists of length 2, 4, or 6 are represented by curves of lengths 2, 4, and 6, respectively. The data from negative trials that had a t value of 2, 4, or 6 are presented in separate panels; the data from trials that had a d value of 2, 4, or 6 are represented by open triangles, open squares and open circles, respectively.)

on the study list. A word at the top of its study list, for example, had Serial Position 1 and was a distance of 0 words from its consonant; the last word in a four-word list had Serial Position 4 and was a distance of 3 words from its consonant.

The panel in the upper left corner of Figure 2 shows that for positive trials on Day 1, RT was an increasing function of the distance between the test word and its consonant. The functions for positive trials on Days 3 and 4 are similar to those for Day 1, although the functions show only a slight increase in RT for words beyond the third list position.

For negative trials on Day 1, there were no consistent serial position effects except for trials on which the test consonant designated a list of Size 6 ($t = 6$); that is, RT was a function of the distance between the test word and its consonant only on trials with test consonants designating long lists. For these trials the panel in the upper right corner of Figure 2 indicates that RT increased with the distance between a test word and its consonant. The RTs from trials on which the test word came from a list of Size 2 were an exception to this trend; this result suggests that subjects may have used some special strategy to verify relationships involving words from short lists. By Days 3 and 4 there were no clear serial position effects for any of the nine negative trial types. The lower right panel of Figure 2 shows that the increasing functions found on Day 1 for the trials that had $t = 6$ were nearly flat by Days 3 and 4.

EXPERIMENT 2

One problem with the procedure used in Experiment 1 is that subjects were allowed to study the lists in any way they chose as long as they could recall each list in its proper order. Therefore, it is difficult to determine the nature of the information that subjects stored about the symbol-element relationships. Experiment 2 was designed, in part, to remedy this problem. In this experiment subjects learned the symbol-element relationships in an experimenter-controlled situation just prior to testing. Half the subjects learned the materials as

six separate lists each named by consonants, as in Experiment 1. The other half learned them as a set of 24 word-consonant paired associates. A second factor, crossed with acquisition method, was the nature of the test display. Test displays were either a consonant above a word or a word to the left of a consonant, and were therefore either congruent or incongruent with the format experienced by a subject during initial learning.

Method

Subjects and design. The subjects were 24 female students at Stanford University, none of whom served in Experiment 1. They were divided into eight equal groups by three between-subjects variables; each variable had two levels and was crossed with both of the other variables. The eight groups were determined by: (a) whether the subjects learned the materials as lists or as paired associates; (b) whether the subjects were tested with symbol-element pairs presented horizontally or vertically; and (c) the set of materials learned by members of a group (Set 1 or Set 2). Each subject was paid \$2.50 for participating in one test session. The apparatus used was the same as that used in Experiment 1.

Materials. The two sets of materials used in Experiment 1 (Set 1 and Set 2) were used again in this experiment. Six blocks of 24 trials each were constructed according to the symbols-equal test schedule described for Experiment 1. Trials within each block were randomized separately for each subject.

Procedure. The subjects were tested individually. At the start of the test session the subject was given a set of 5 × 8 in. (12 × 20 cm) index cards. For subjects who learned the materials as a set of lists the cards were the same as those given to subjects in Experiment 1; that is, the subjects received six cards, with one word list in a column below a consonant on each card. For subjects who learned the materials as paired associates a word followed by a consonant (e.g., CORNER x) appeared on each of 24 cards. For test purposes, the list symbol (a consonant) was printed on the back of each list card, and the word part of the paired associate was printed on the back of each paired-associate card.

The subject was told to study the cards one at a time in the order in which they appeared in the deck. When the subject had looked at each card once, the cards were turned face down and a recall test was given. The subjects in the list-learning condition were shown the consonant on the back of each card and were asked to recall, in correct serial order, the list designated by that consonant. Subjects in the paired-associate condition were

shown the word on the back of each card and were asked to recall the consonant associated with that word. In each case subjects were informed whenever they made an error. After the recall test the experimenter randomized the cards in the deck and gave them to the subject for another study trial. Study trials and recall tests alternated until the subject completed two successive recall tests without an error. Subjects in both learning conditions took about 20-25 min to learn a set of materials.

The subject was then seated at the CRT screen and given instructions about the task. The instructions were similar to those used in Experiment 1. The subjects in the paired-associate condition, however, were told to respond yes if the consonant-word pair was one that they had learned and to respond no otherwise. The same practice task used in Experiment 1 was then given. After the practice task the instructions were reviewed, and the six blocks of test trials were given.

The sequence of events on each trial was identical to that followed in Experiment 1. For half the subjects, however, the consonant-word test stimulus appeared horizontally on the screen; that is, the test word appeared slightly to the left of the fixation point and the test consonant appeared one character to the right of the fixation point. For the other subjects the test stimulus was vertical, a consonant above a word, as in Experiment 1.

Results

An initial analysis of variance was performed on the data from Experiment 2 with response type (yes vs. no) as a within-

subjects factor, and with learning condition (lists vs. paired associates), test stimulus display type (vertical vs. horizontal), and stimulus materials (Set 1 vs. Set 2) as between-subjects factors. The analysis indicated that the main effect due to differences in the materials learned by subjects and all interactions involving the materials variable were nonsignificant ($p > .10$ in all cases). Since subjects who learned different sets of materials showed no overall differences in performance, data from subjects in the two groups were pooled and all further analyses were performed on the pooled data. The RTs from trials on which an incorrect response was made were excluded from the analysis.

Table 2 presents a summary of the data. The mean of the subject mean RTs and the mean of the subject standard deviations for each trial type were computed in the same way as those reported for Experiment 1. Graphs of the mean RTs in Table 2 are presented in Figure 3. The lines in Figure 3 represent predicted RTs generated by a model to be presented later. The labels used for the lines and bars in Figure 3 are the same as those used in Figure 1.

An analysis of variance using subject means as scores was performed on the data from positive trials. The analysis included

TABLE 2
MEAN OF SUBJECT MEAN REACTION TIMES, MEAN OF SUBJECT STANDARD DEVIATIONS,
NUMBER OF OBSERVATIONS, AND PERCENT ERRORS FOR EACH TRIAL TYPE
IN EXPERIMENT 2

Response and values of <i>l</i> and <i>d</i>	List condition				Paired-associate condition			
	RT (msec)	SD (msec)	No. obs.	% Errors	RT (msec)	SD (msec)	No. obs.	% Errors
Yes								
2 2	1,289	367	288	3.5	1,203	384	288	3.5
4 4	1,763	645	288	4.2	1,375	449	288	6.7
6 6	2,270	922	288	9.4	1,394	470	288	6.6
No								
2 2	1,744	600	96	10.6	1,497	498	96	3.1
4 2	1,832	663	96	5.2	1,500	465	96	1.0
6 2	1,759	660	96	4.3	1,477	485	96	2.1
2 4	1,748	704	96	3.1	1,499	520	96	4.2
4 4	2,117	750	96	8.3	1,526	578	96	3.1
6 4	2,233	881	96	9.5	1,709	579	96	6.4
2 6	1,839	754	96	0.0	1,519	477	96	2.1
4 6	2,325	957	96	3.1	1,652	568	96	4.2
6 6	2,329	721	96	15.6	1,611	669	96	10.4

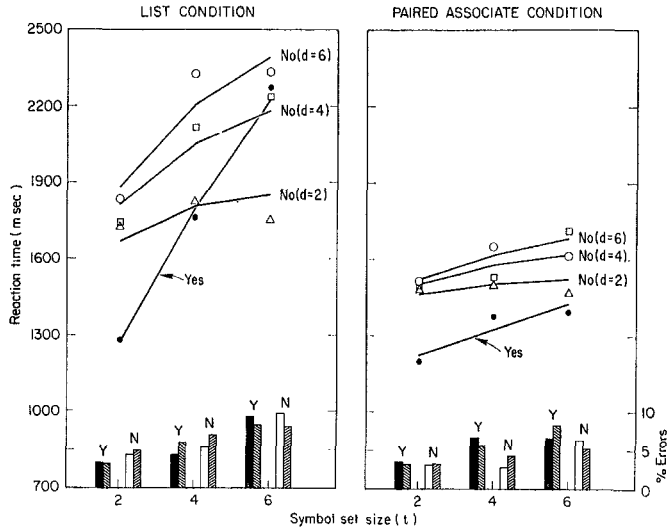


FIGURE 3. Mean reaction times and error rates for the list condition and the paired-associate condition. (The means for positive trials are represented by filled circles; means for negative trials that had a d value of 2, 4, or 6 are represented by open triangles, open squares, and open circles, respectively. Error rates for positive and negative trials [collapsed across the three values of d] are represented by solid and open bars, respectively. The lines and hatched bars represent predicted values.)

the value of t as a within-subjects factor and both learning condition (lists vs. paired associates) and stimulus display type (horizontal vs. vertical) as between-subjects factors. The analysis showed that there were significant effects due to learning conditions, $F(1, 20) = 5.53$, $p < .05$, the value of t , $F(2, 40) = 28.16$, $p < .001$, and the $t \times$ Learning Conditions interaction, $F(2, 40) = 12.98$, $p < .001$; all other effects were nonsignificant ($p > .10$ in all cases). As Figure 3 indicates, positive RT increased with t for both learning conditions. For subjects in the list condition RT increased dramatically as the length of the tested list increased; for subjects in the paired-associate condition, however, the number of words associated with a test consonant had a much smaller effect on RT.

Another analysis of variance using subject means as scores was performed on the data from negative trials. This analysis included learning condition and stimulus display type as between-subjects factors and the values of both t and d as within-subjects factors. The analysis showed that there

were significant main effects due to t , d , and learning conditions, $F(2, 40) = 9.00$, $p < .001$; $F(2, 40) = 8.13$, $p < .005$; and $F(1, 20) = 4.40$, $p < .05$, respectively. Also significant were the $t \times d$ interaction, $F(4, 80) = 4.34$, $p < .005$; the $t \times$ Learning Condition interaction, $F(2, 40) = 3.50$, $p < .05$; and the $t \times d \times$ Display Type interaction, $F(4, 80) = 3.30$, $p < .025$. All other effects were nonsignificant ($p > .05$).

As in Experiment 1, then, negative RT was a function of both t and d . The differences between the two learning conditions can be determined from Figure 3. The figure indicates that subjects gave faster negative responses in the paired-associate condition than in the list condition and that the effects due to t and d were smaller in the paired-associate condition than in the list condition. Figure 3 also indicates that the tendency noted in Experiment 1 for negative responses to be fast on trials with unequal values of t and d was not evident in Experiment 2.

The analyses for both positive and negative responses suggest that the nature of

the stimulus display had almost no effect on RT. Separate examinations of the data from subjects in the two display type conditions showed that the $t \times d \times$ Display Type interaction, which was significant for negative responses, had no obvious interpretation.

DISCUSSION

The results of Experiments 1 and 2 allow the following characterization of performance in the symbol-element recognition task. On positive trials RT increased linearly with symbol set size (t). The slope of the function relating positive RT to t was greatest for sets of elements learned as lists; the slope decreased with extended testing (Experiment 1) and when the symbol-element relationships were learned as paired associates (Experiment 2). On negative trials RT depended on both symbol set size (t) and element set size (d). In both experiments negative RT tended to increase as a function of both t and d ; for subjects who learned the symbol-element relationships outside the laboratory (Experiment 1), however, negative RT on trials that had unequal values of t and d was less than a strictly increasing function of t and d would predict. The effects of t and d on negative RT were greatest for subjects who learned the sets of elements as lists; these effects were reduced by extended testing (Experiment 1) and when the materials were learned as paired associates (Experiment 2). In Experiment 1 performance was not affected by test schedules that varied the probabilities of presenting particular symbol-element pairs. Experiment 2 indicated that performance was not affected by the nature of the test display. The error rates in both experiments were low and positively correlated with RT.

The results of these two experiments argue against the notion that information learned about each symbol-element pair was stored in a separate memory location to which subjects had direct access. Direct access models for recognition memory, such as those outlined in the introduction, do not make differential RT predictions for any of the conditions in the present experiments

(i.e., they predict no effects of symbol set size or element set size). More sophisticated direct access models (for example, a variation of the "access time" model of Murdock, 1974, chap. 4) could be formulated to account for the present results, but such models would be complicated and will not be considered here.

A second class of models for recognition memory makes the assumption that a structure in memory is searched between the onset of the test stimulus and the execution of a response. Although models in this class may be discriminated conceptually, particular models can prove to be formally equivalent when applied to our recognition paradigm. Therefore, a reasonable approach to evaluating models is to reject clearly inconsistent explanations, rather than try to demonstrate that one particular model is "correct."

Before we consider some possible search models it should be noted that the present results differ from those obtained in other recognition tasks for which search models have been proposed. Juola, Fischler, Wood, and Atkinson (1971), for example, report a recognition study in which each subject learned and was tested on a single well-memorized list of 10, 18, or 26 words. In their study RT on positive trials increased less than 6 msec for each additional word on the memorized list. The Atkinson and Juola (1973, 1974) model for recognition memory attributes the effect of list length in the Juola et al. study to the fact that subjects searched a stored representation of the memorized list on a proportion of the test trials. The model assumes that on the remaining trials subjects were able to execute a response without a memory search simply by evaluating the "familiarity" of the test word. The small effect of list length in the Juola et al. single-list experiment can be contrasted with the much larger effect of list size in the present experiments, in which search processes presumably played a more important role; on Day 1 of Experiment 1, for example, positive RT increased about 152 msec for each additional word on the tested list. In studies of short-term recognition memory increases in positive

RT are typically less than 50 msec for each additional item in the positive set (Sternberg, 1969), a much smaller increase than that found in the present experiments. This fact, along with arguments presented below, suggests that the search processes involved in the symbol-element recognition task occur primarily in long-term memory.

A first class of memory search models for the symbol-element recognition task will be referred to as symbol-entry models. These models assume that on each trial the subject locates in memory a representation of the set designated by the test symbol and searches that set for the test element. If the subject entered the search set into short-term memory before checking for the test element, the symbol-element task would be similar to the short-term memory task studied by Sternberg (1969).³ One prediction of a symbol-entry model is that element set size (d) should have no effect on negative RT, since only the elements associated with the test symbol are involved in the search process. Because there were large effects of d on negative RT, symbol-entry models cannot adequately account for performance in the symbol-element recognition task.

A second class of memory search models (complementary to symbol-entry models) is that of element-entry models. These models assume that on each trial the subject locates in memory a representation of the set of information accessed via the test element; by searching that set the subject determines whether the test element was associated with the test symbol during learning. The predictions of element-entry models for the relation between RT and the variables t and d depend on assumptions about the format of stored information and characteristics of the search process. At one extreme, such models may assume either that each element-symbol pair has a sep-

arate location in memory or that the search process is selective and limited to evaluating symbols (but not other elements) in the memory structure; in these cases, RT on both positive and negative trials is expected to be independent of both t and d . Such element-entry models are equivalent to the "direct access" models mentioned previously. However, different predictions are generated by element-entry models that assume that the memory structures located via the test element contain other elements of its subset as well as its symbol and that the search process is sensitive to these elements. For these models, RT will be a function of d (but not of t) on negative trials, since it is the elements in the set containing the test element that affect the search process. Because there were large effects of t on negative RT, element-entry models cannot give an adequate account of performance in the symbol-element recognition task.

We next consider two classes of search models that are somewhat more complicated. These models can predict a linear increase in positive RT with t and also predict that both t and d will affect RT on negative trials. The first of these models will be called probabilistic-entry models; they are derived from symbol-entry and element-entry models and assume that performance reflects a mixture of search strategies across trials. Such models can assume that part of the stored information about symbol-element associations is retrievable only by symbol-entry processes, and the other part only by element-entry processes. Alternatively, it may be assumed that all information is retrievable in both ways but that the test symbol is used as an entry point into memory on some trials and the test element is used on other trials. Probabilistic-entry models make predictions that depend on the representation given to the alternative search processes; unlike either the symbol-entry or the element-entry models, however, they can predict effects of both t and d for negative trials, since these data will reflect a mixture of trials where either symbol-set size or element-set size had an effect.

³ Results of a study by Appelman and Atkinson (1975) indicate that such a strategy is used in the symbol-element task if the subject is cued prior to the start of each trial with the symbol designating the test list for that trial. With symbol cuing, performance closely approximates the results reported by Sternberg (1969).

The second of these models will be called simultaneous-entry models. Similar to probabilistic-entry models, these assume that both the test symbol and the test element serve as entry points into memory. However, they further assume that both entry points are used on each trial. Search processes are initiated simultaneously at both entry points and the search that finishes first determines the RT for that trial. Again, like the probabilistic-entry models, the predictions of simultaneous-entry models depend on particular representations given to the symbol-entry and element-entry search processes. In general, however, these models predict effects of both t and d on negative RT.

We have fitted our data with both a probabilistic-entry model and a simultaneous-entry model. For simplicity we will present only the predictions generated by the simultaneous-entry model. In this regard it should be noted that, although neither model accounted for all aspects of the results, the fit of the best simultaneous-entry model was clearly superior to the fit of the best probabilistic-entry model. In addition, Anderson and Bower (1973, chap. 12) have had considerable success in applying a model similar to our simultaneous-entry model to RT data from a number of recognition memory experiments.

Parameters of the simultaneous-entry model were estimated separately for Day 1 and Days 3 and 4 of Experiment 1, and separately for the list-learning condition and the paired-associate learning condition of Experiment 2. For each of the four sets of data the model was used to generate 12 predicted RTs, 1 for each of the trial types displayed in Figures 1 and 3. The model also generated predicted error rates for each of the observed error probabilities presented in Figures 1 and 3; these predicted values are presented in the figures.⁴ A mathe-

matical formulation of the model together with a description of the method used to fit the model to the data are presented in the appendix; the parameter estimates are also presented in the appendix.

The simultaneous-entry model assumes that when a test display is presented, search processes are initiated simultaneously from two locations in memory. One search begins at a location where a representation of the test symbol is stored; this process determines whether a representation of the test element is in the memory structure associated with the test symbol. When the test element is in the memory structure associated with the test symbol, this process has an expected completion time of αt ; if the test element is not in the structure, the expected completion time is $\alpha' t$. A second search process begins at a location in memory where a representation of the test element is stored; this process determines whether the test symbol is associated with the memory structure to which the test element belongs. The expected completion time for this process (on both positive and negative trials) is βd . Starting simultaneously with the two search processes is a guessing process that has an expected completion time γ on both positive and negative trials. The two search processes and the guessing process are all assumed to have completion times that are exponentially distributed.⁵ The first of the three processes to finish activates a response. If either of the search processes finishes first, the subject makes a correct response; if the guessing process finishes first, the subject guesses yes with probability g and guesses no otherwise. The sum of the times necessary to encode the test stimulus, initiate the search processes, and execute a response

no consistent tendency to be either larger or smaller than correct RTs.

⁵The assumption of exponential distributions leads to mathematically tractable expressions that predict the data reasonably well. On theoretical grounds, a somewhat more realistic assumption is that the completion times are gamma distributed (Anderson, 1974); unfortunately, there is no tractable characterization of the fastest of an arbitrary number of gammas.

⁴For each trial type the model predicts that the distribution of error reaction times (RTs) will be the same as the distribution of correct RTs. Since the error rates in these experiments were too low to produce reliable estimates of error RTs, predictions for error times were not evaluated; however, the obtained error RTs showed

is r_1 when a yes response is given, and r_0 when a no response is given.

The model as outlined above incorporates no specific assumptions about the way information is represented in memory; that is, the model says nothing about the number or type of associations that are examined during any particular search, as, for example, is the case for the models considered by Anderson and Bower (1973). Because no particular representation is incorporated in the model, the search parameters (α , α' , and β) cannot be given a specific interpretation such as rates at which associations are examined. An assumption that the model does make is that the completion time for a search beginning at a symbol increases linearly as the size of the set designated by that symbol increases, and that the completion time for the search beginning at an element increases linearly with the size of the set to which that element belongs.

The model's predicted RTs and error rates for Experiment 1 are presented in Figure 1. The error rates predicted by the model increase with symbol set size for both positive and negative trials; the error rates obtained in Experiment 1 showed a similar trend. From Figure 1 it is clear that the predicted RTs for positive trials are nearly identical to the obtained positive response times. For negative trials, however, there are some rather large discrepancies between predicted and observed RTs. The difficulty is that the model predicts a monotonic increase in negative RT both with t and with d when, in fact, negative RT did not increase monotonically with these variables. What this nonmonotonicity in the data indicates is not entirely clear. One hypothesis is that subjects were able to use information about the sizes of the lists associated with test symbols and elements to make some fast decisions. Thus, on some trials with a symbol and an element from lists of different sizes, subjects may have made a negative response without searching for an associative path between the test symbol and the test element. According to this hypothesis the simultaneous search processes were initiated consistently only on positive trials and on negative trials with equal values of

t and d ; on some of the trials with unequal values of t and d , subjects gave a negative response immediately after noting that the test symbol and the test element were from lists of different sizes. Although we have not tested this "size estimation" hypothesis directly, it does account qualitatively for the fact that negative RTs were particularly fast on trials with unequal values of t and d .

The simultaneous-entry model can account for the serial position effects shown in Figure 2; for it to do so, however, some additional assumptions must be made about the way symbol-element relationships were represented in memory. The data from positive trials on Day 1 suggest that, at the start of testing, the representation of the symbols and elements in memory was similar to the display used by the subjects to memorize the lists. By this we mean that representations of the words from each list were associated in serial order, and that a representation of the list symbol was linked to the first list word. With a representation of this type the simultaneous-entry model correctly predicts that positive RT will increase with the distance of the test element from the top of its list. For negative trials, the model predicts a serial position effect only when the response is initiated by the search process starting from the test element; whenever the search process starting at the test symbol finishes first, negative RT will be independent of the serial position of the test element. Thus the model correctly predicts that the largest serial position effects on negative trials will occur when the test symbol designates a long list; it is on those trials that the search process starting at the test element has the greatest chance of finishing first.

The data from Days 3 and 4 suggest that after extended testing the symbol-element relationships were no longer represented in memory merely as a set of ordered lists. In addition, it appears that by Days 3 and 4 most of the test elements were associated directly with the appropriate symbol. This hypothesis accounts both for the fact that the effects of t and d decreased with extended testing and for the fact that by Days

3 and 4 the serial position effects were much smaller than they were on Day 1.

The greatest weakness of the simultaneous-entry model, as presented here, is that it does not characterize completely the way learned relationships are represented in memory. The notion that subjects in Experiment 1 started with a set of ordered lists in memory and gradually added a set of direct associations between elements and symbols is clearly an incomplete picture of the information that subjects had stored about the symbols and elements. It is certainly true that each subject in Experiment 1 had stored a great deal of contextual information about the experimental situation along with the symbol-element relationships. In addition, the length of time subjects had to study their lists gave them ample opportunity to develop various strategies for performing in the test sessions. Although effects due to alternative learning strategies are difficult to eliminate, the controlled learning procedure used in Experiment 2 was designed to minimize these effects; in particular, this procedure was designed to control the way in which the symbols and elements were associated in memory.

The predictions of the simultaneous-entry model for Experiment 2 are presented in Figure 3. The error rates in Experiment 2 tended to increase with symbol set size, as the model predicts. As in Experiment 1, the predicted RTs for positive trials are very close to the obtained positive response times. The predictions for negative trials in Experiment 2 are much closer to the obtained RTs than were the predictions in Experiment 1. As Figure 3 shows, the ordering of the negative RTs in Experiment 2 was in nearly all cases the same as the ordering predicted by the model. In Experiment 1 the unusually fast RTs on negative trials in which t was not equal to d were attributed to the fact that subjects did not always search for an associative path between the test symbol and the test element. The results of Experiment 2 suggest that when the learning of the test materials was controlled subjects were not able to bypass the search processes.

The differences between the data from the list condition and the paired-associate condition in Experiment 2 were similar to the differences between Day 1 and Days 3 and 4 in Experiment 1. The differences between Day 1 and Days 3 and 4 in Experiment 1 were attributed primarily to a new set of associations that were added during testing; we suggested that on Day 1 the symbols and elements were represented in memory as a set of ordered lists and that by Days 3 and 4 most of the elements were associated directly with the appropriate symbol. The learning conditions used in Experiment 2 were designed specifically to produce these two kinds of associative structures in memory. By requiring serial recall of each list, the list-learning procedure encouraged the formation of ordered list structures in memory, whereas the paired-associate learning procedure encouraged the formation of direct associations between each element and the appropriate symbol. The results of Experiment 2, then, indicate that the changes produced in Experiment 1 by extended testing can be approximated by inducing subjects during learning to store either a set of list structures or a set of direct associations in memory.

REFERENCES

- Anderson, J. R. Retrieval of propositional information from long-term memory. *Cognitive Psychology*, 1974, 4, 451-474.
- Anderson, J. R. & Bower, G. H. *Human associative memory*. Washington: V. H. Winston, 1973.
- Anderson, J. R. & Bower, G. H. A propositional theory of recognition memory. *Memory & Cognition*, 1974, 2, 406-413.
- Appelman, I. B. & Atkinson, R. C. Search of list structures stored in long-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1975, 14, 82-88.
- Atkinson, R. C., Herrmann, D. J., & Wescourt, K. T. Search processes in recognition memory. In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola symposium*. Potomac, Md.: Erlbaum, 1974.
- Atkinson, R. C. & Juola, J. F. Factors influencing speed and accuracy of word recognition. In S. Kornblum (Ed.), *Attention and performance IV*. New York: Academic Press, 1973.
- Atkinson, R. C. & Juola, J. F. Search and decision processes in recognition memory. In D. H. Krantz, R. C. Atkinson & P. Suppes (Eds.),

- Contemporary developments in mathematical psychology*. San Francisco: Freeman, 1974.
- Carroll, J. B., Davies, P., & Richman, B. *Word frequency book*. New York: American Heritage Press, 1971.
- Clark, H. H. The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 335-359.
- Homa, D. Organization and long-term memory search. *Memory & Cognition*, 1973, 1, 369-379.
- Jacoby, L. L. Effects of organization on recognition memory. *Journal of Experimental Psychology*, 1972, 92, 325-331.
- Juola, J. F., Fischler, I., Wood, C. T., & Atkinson, R. C. Recognition time for information stored in long-term memory. *Perception & Psychophysics*, 1971, 10, 8-14.
- Kintsch, W. Recognition and free recall of organized lists. *Journal of Experimental Psychology*, 1968, 78, 481-487.
- Kintsch, W. Models for free recall and recognition. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970.
- Mandler, G. Organization and recognition. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972.
- Mandler, G., Pearlstone, R., & Koopmans, H. S. Effects of organization and semantic similarity on recall and recognition. *Journal of Verbal Learning and Verbal Behavior*, 1969, 8, 410-423.
- Murdock, B. B., Jr. Modality effects in short-term memory: Storage or retrieval? *Journal of Experimental Psychology*, 1968, 77, 79-86.
- Murdock, B. B., Jr. *Human memory: Theory and data*. Potomac, Md.: Erlbaum, 1974.
- Postman, L., Jenkins, W. O., & Postman, D. L. An experimental comparison of active recall and recognition. *American Journal of Psychology*, 1948, 61, 511-519.
- Quillian, M. R. Semantic memory. In M. Minsky (Ed.), *Semantic information processing*. Cambridge, Mass.: M.I.T. Press, 1968.
- Rumelhart, D. E., Lindsay, P. H., & Norman, D. A. A process model for long-term memory. In E. Tulving & W. Donaldson (Eds.), *Organization of Memory*. New York: Academic Press, 1972.
- Seamon, J. G. Retrieval processes for organized long-term storage. *Journal of Experimental Psychology*, 1973, 97, 170-176.
- Shepard, R. N. & Teghtsoonian, M. Retention of information under conditions approaching a steady state. *Journal of Experimental Psychology*, 1961, 62, 302-309.
- Shiffrin, R. M. & Atkinson, R. C. Storage and retrieval processes in long-term memory. *Psychological Review*, 1969, 76, 179-193.
- Smith, E. E., Shoben, E. J., & Rips, L. J. Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, 1974, 81, 214-242.
- Sternberg, S. Memory scanning: Mental processes revealed by reaction-time experiments. *American Scientist*, 1969, 57, 421-457.
- Tulving, E. & Thomson, D. M. Retrieval processes in recognition memory: Effects of associative context. *Journal of Experimental Psychology*, 1971, 87, 116-124.

APPENDIX

The parameters of the simultaneous-entry search model are described in the text. Because the completion times for the two search processes and the guessing process are assumed to be exponentially distributed, we can readily write expressions for the desired RTs and error rates. The expected time to make a correct positive response on trials with symbol set size t is:

$$RT(\text{correct yes}; t) = \frac{\alpha\beta\gamma t}{(\alpha + \beta)\gamma + \alpha\beta t} + r_1.$$

The expected error probability on these trials is:

$$P(\text{error}|\text{yes}; t) = \frac{(1 - g)\alpha\beta t}{(\alpha + \beta)\gamma + \alpha\beta t}.$$

The expected time to make a correct negative response on trials with symbol set size t and

element set size d is:

$RT(\text{correct no}; t, d)$

$$= \frac{\alpha'\beta\gamma t d}{(\alpha' t + \beta d)\gamma + \alpha'\beta t d} + r_0.$$

The expected error probability on these trials is:

$$P(\text{error}|\text{no}; t, d) = \frac{g\alpha'\beta t d}{(\alpha' t + \beta d)\gamma + \alpha'\beta t d}.$$

Twelve RTs and six error rates were used to estimate parameters; namely, the 12 RTs and six error probabilities displayed in each of the panels of Figures 1 and 3. The model predicts error rates separately for each of the nine negative trial types; to obtain predictions for negative trials collapsed across values of d , predictions for the three values of d were

averaged at each value of t . Parameter estimates were selected that minimized the sum of squared deviations between both predicted and observed RTs and predicted and observed error rates. Specifically, the following loss function was defined:

$$\sum_{i=1}^{12} \left\{ \frac{m_i - \mu_i}{s_i / \sqrt{n_i - 1}} \right\}^2 + \sum_{j=1}^6 \left\{ \frac{[i_i - N_j p_j]^2}{N_j p_j} + \frac{[f_j' - N_j(1 - p_j)]^2}{N_j(1 - p_j)} \right\},$$

where i is an index over the 12 RTs and j is an index over the six trial types for which error rates are predicted. In addition:

- m_i = the observed RT for data point i
- μ_i = the predicted RT for data point i
- s_i = the mean standard deviations over subjects for data point i
- n_i = the number of RT observations for each subject determining data point i
- f_j = the number of incorrect responses associated with the j th trial type
- f_j' = the number of correct responses associated with the j th trial type
- $N_j = f_j + f_j'$
- p_j = the predicted error probability for the j th entry

To determine a minimum value for the loss

TABLE A1
PARAMETER ESTIMATES AND VALUES OF THE LOSS FUNCTION FOR EXPERIMENTS 1 AND 2

Parameter	Experiment 1		Experiment 2	
	Day 1	Days 3 & 4	List condition	Paired-associate condition
α	338	79	1,411	132
α'	282	126	487	136
β	554	224	433	123
γ	4,771	2,931	7,077	2,175
r_1	854	727	678	1,099
r_0	1,179	827	1,243	1,344
g	.553	.519	.625	.446
Loss function	55.1	148.0	11.5	12.2

Note. Estimates of all parameters are in milliseconds except for g , which is a probability.

function, a computer was programmed to conduct a systematic search of the parameter space. The parameter estimates that minimized the loss function, together with the minimum values of the loss function, are presented in Table A1. The minimized loss function is χ^2 distributed with 18 degrees of freedom minus 1 for each estimated parameter; in this case 7 parameters were estimated, so there are 11 degrees of freedom. The parameters in Table A1 were used to generate the predictions shown in Figures 1 and 3.

(Received August 12, 1974)