FACTORS INFLUENCING SPEED AND ACCURACY OF WORD RECOGNITION[1]

Richard C. Atkinson and James F. Juola

Stanford University

Stanford, California 94305

Paper presented at the Fourth International Symposium on Attention and Performance held at the University of Colorado, Boulder, Colorado, August 16-21, 1971.

Table of Contents

ABSTRACT

Seven experiments, designed to investigate the effects of various factors on word recognition, are reported. For each experiment the subject memorized a list of from 16 to 54 words and then was tested with a sequence of single words; each test involved either a target word (member of the list) or distractor word (not on the memorized list). In response to each test word the subject pressed one of two keys, indicating whether the word was a target or a distractor. Response latency was shown to depend upon the number of prior tests on a given word (Experiment 1) and the length of the target list (Experiment 2). Experiments 3 and 6 demonstrated that response latency to a target word can be decreased by repeating the word in the study list or by otherwise making certain words more salient. Experiments 4, 5, and 7 showed that response latency was affected by similarities between target and distractor words and by such word characteristics as frequency, concreteness, and syllable length. The latency and error data were discussed in terms of a model for recognition which assumes that the subject either (1) makes an initial fast response based on the familiarity of the test word or (2) if the familiarity is neither high nor low, delays responding until an extended search of the memorized list is carried out. Quantitative predictions generated by the model compare favorably with the data.

Introduction

This paper describes a series of experiments that were designed to study search and retrieval processes in long-term memory. Specifically, the problem under investigation is how a subject is able to decide whether or not a given test stimulus is a member of a predefined set of target items. For any initial set S of stimuli, a subset S_1 is defined which is of size d. Stimuli in subset S_1 will be referred to as target items; subset S_0 is the complement of S_1, and its members will be called distractor items. The experimental task involves a long series of discrete trials, where on each trial a stimulus is presented from S. To each presentation the subject must make either an A_1 or A_0 response indicating that he judges the stimulus to be either a target or distractor item, respectively. For the experiments reported in this paper, the stimuli are all words presented visually. However, the model that we shall present can be applied to a broader class of stimuli.

In a task of the type described above it is possible to initiate a test sequence without the subject knowing the features that distinguish the target stimuli from distractors. As the discrimination is learned, the probabilities of correct responses and errors would change, eventually reaching stable performance. For our present purposes, however, it was desirable to have the target set extremely well-learned prior to the experimental session. Under these conditions the subject is able to indicate with almost perfect accuracy whether the test stimulus is a target or a distractor, and the principal data are response latencies.

There are many ways to define target and distractor sets so that varying demands are placed on the subject's memory as he makes a decision.

If S_1 is distinguished from S_0 by means of a simple rule, it might be unnecessary to retrieve any information from long-term memory before making a response. For example, if S_1 is the set of all English words beginning with the letter "b" and S_0 is all the remaining English words, the subject can respond to each test word even before the name of the word is retrieved from memory. More complex rules could be constructed so that information stored in long-term memory must be accessed before a decision is made. For example, let S_1 be the set of all four-legged animal names. The subject would not only have to name the word presented on a test, but would have to retrieve some information about the semantic properties of that word before responding.

Alternatively, no rule might suffice to distinguish target stimuli from distractors; e.g., if the target set consists of a list of unrelated words previously memorized by the subject. In this case there are at least two ways that the subject could identify target stimuli. One possibility is that the subject retrieves the contents of the long-term storage location at the address of the tested stimulus. This could include information about whether or not that item has been previously designated a member of S_1 (i.e., the information could contain "list markers" for target words). On the other hand, if the target set is not too large, the subject could store the items in memory as a list structure and then compare the test stimulus with each item on the list. This latter process presumably takes place in experiments where the target set is limited to about six items or less which are placed in short-term memory immediately prior to the onset of a test stimulus (Sternberg, 1966). A similar scan of S_1 could occur if it is permanently stored in long-term memory and accessed at the time of test, as in the case for questions such as, "Does the number '4' occur in your home phone number?"

The task of interest for the present discussion is one in which no rule applies to the distinction between S_0 and S_1 stimuli, and the target set is too large to maintain in short-term memory. This task is comparable to that used by Sternberg (1966), but the memory sets we employ are much larger, and must be maintained in long-term memory rather than being presented shortly before the onset of every test stimulus. The questions to be answered concern the type of memory search that is necessary to make a recognition decision. Different models, incorporating the search process as one of several successive and independent stages, can be tested against latency data to determine the most probable mechanism for recognition.

A Prototype Experiment

Experiment 1 was designed to study the effects on response speed of repeated tests on target and distractor stimuli (Fischler and Juola, 1971). All test stimuli were selected from a common pool consisting of 48 one-syllable nouns (Thorndike-Lorge, 1944, frequency of A or AA). For each of 20 subjects a different set of 24 S_1 and 24 S_0 words were randomly selected from S. The S_1 words were given to the subject as a list to be learned in serial order, approximately 18 hours before the experimental session.

At the start of the test session the subject was allowed to study his target list for a few minutes, and then was given a written serial recall test. All subjects satisfied a pre-experimental criterion by correctly recalling the lists on two successive trials (no subject made any errors).

The subject was then seated in front of a tachistoscope, in which the test words were presented one at a time. To each presentation the subject made either an A_1 or an A_0 response (indicating that the test word was a target or a distractor, respectively) by depressing one of two telegraph keys with his right forefinger. The keys were separated by a central home key on which the subject rested his finger between trials. The assignment of A_1 and A_0 responses to right or left keys was counterbalanced across the subjects.

The test sequence consisted of 120 consecutive trials that were divided into four blocks. For Block I, six target words and six distractors were randomly selected from S_1 and S_0, respectively. For Block II, the 12 Block I words were repeated, and six new targets and six new distractors were also shown. Block III included all the words presented in Block II with 12 new words (six targets and six distractors). Finally, Block IV included all the words of Block III plus the 12 remaining words in S. Thus 12 words were presented in Block I, 24 words in Block II, 36 words in Block III, and 48 words in Block IV. Order of presentation within blocks was randomized.

The subjects were instructed to respond as rapidly as possible to each test word, while being careful to avoid making errors. No feedback was provided for correct responses, but the subjects were informed whenever an error was made. This feedback, however, was unnecessary, because the subjects were almost always immediately aware of the fact that an error had occurred. The trials were self-paced, with the test session lasting approximately 35 min.

The mean error percentages and mean latencies for correct responses for the four trial blocks are presented in Fig. 1. Within each block, errors and latencies are plotted as functions of presentation number for targets and distractors. In each block, the mean latency of A_1 responses is greater than that of A_0 responses when target and distractor words are presented for the first time. When test words are repeated, however, positive response latency decreases whereas negative response latency increases. The strength of this interaction decreases across blocks; i.e., the effects of repetitions on both positive and negative response latencies are not as great in Block IV as they are earlier.[2]

Effects similar to those observed for response latencies can be noted in the error data. The solid bars in the lower part of Fig. 1 are errors to target words, and the open bars are errors to distractors. In each block most errors to targets occurred on initial presentations, whereas most errors to distractors occurred on later presentations.

Mean positive response latency was also plotted as a function of the serial position of the target word in the study list. There was absolutely no trend relating response latency to serial position. This was true for initial and repeated presentations of target words separately as well as for the combined data. Since this result might seem somewhat surprising, it is worth noting that in every experiment we have run using the above paradigm (this includes all the studies discussed in the present paper) there has been no effect of the target word's serial position on response latency. A discussion of these results and those of the following experiment will be delayed until we have considered a model for recognition memory and developed its theoretical implications.

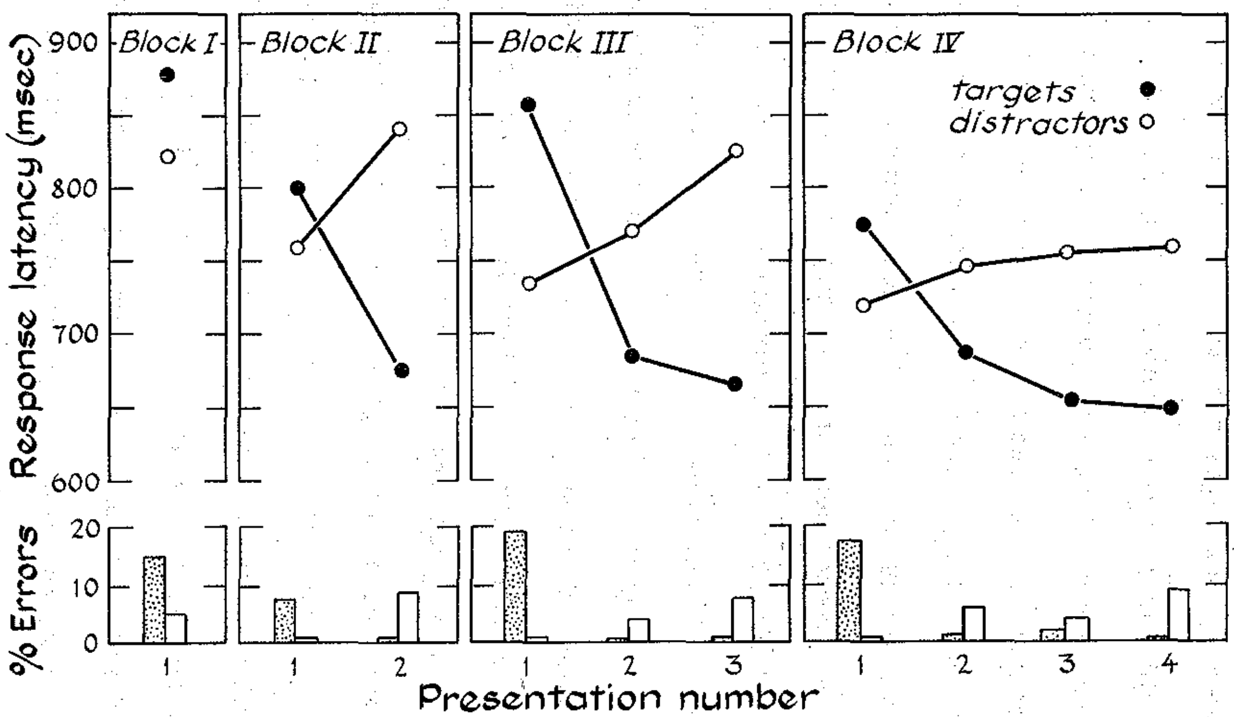

Fig. 1. Mean response latency and error percentages as functions of presentation number for targets and distractors in each of four blocks. Incorrect responses to target words are represented by the shaded bars, and errors to distractors are represented by open bars (Experiment 1).

Effects of Varying the Length of the Target Set

Experiment 2 was essentially the same as the previous study except that the number of words in the target lists was varied. A population of 48 common, one-syllable nouns was used to generate lists of 16, 24, or 32 words. Subjects were randomly assigned to one of the three list-length conditions, with 24 subjects in each group. All subjects satisfied the serial recall criterion for S_1 words described for Experiment 1.

In order to replicate the design of Experiment 1 as closely as possible, only 16 target words and 16 distractors were tested for all three groups. The distractor words presented were randomly selected from the remaining items in the pool; targets consisted of consecutive strings of 16 words from the S_1 lists (either the first, middle, or last 16 words from the 24 and 32 word sets).[3] Using this procedure, it was possible to present identical test sequences to all three groups of subjects. As in Experiment 1 the test trials were broken into four consecutive blocks; four target words and four distractors were presented for the first time in each block, along with all of the words presented in the previous block.

For each list length the pattern of latency and error data closely matched those presented in Fig. 1 for Experiment 1. To test for the presence of list-length effects, the data from the last two trial blocks (Blocks III and IV) were combined; mean latencies were obtained for A_1 and A_0 responses to test words that were presented for the first times, and for those that had been presented previously. These data are given in Fig. 2. The left panel of Fig. 2 shows mean latencies for initial presentations of target and distractor words, and the right panel shows

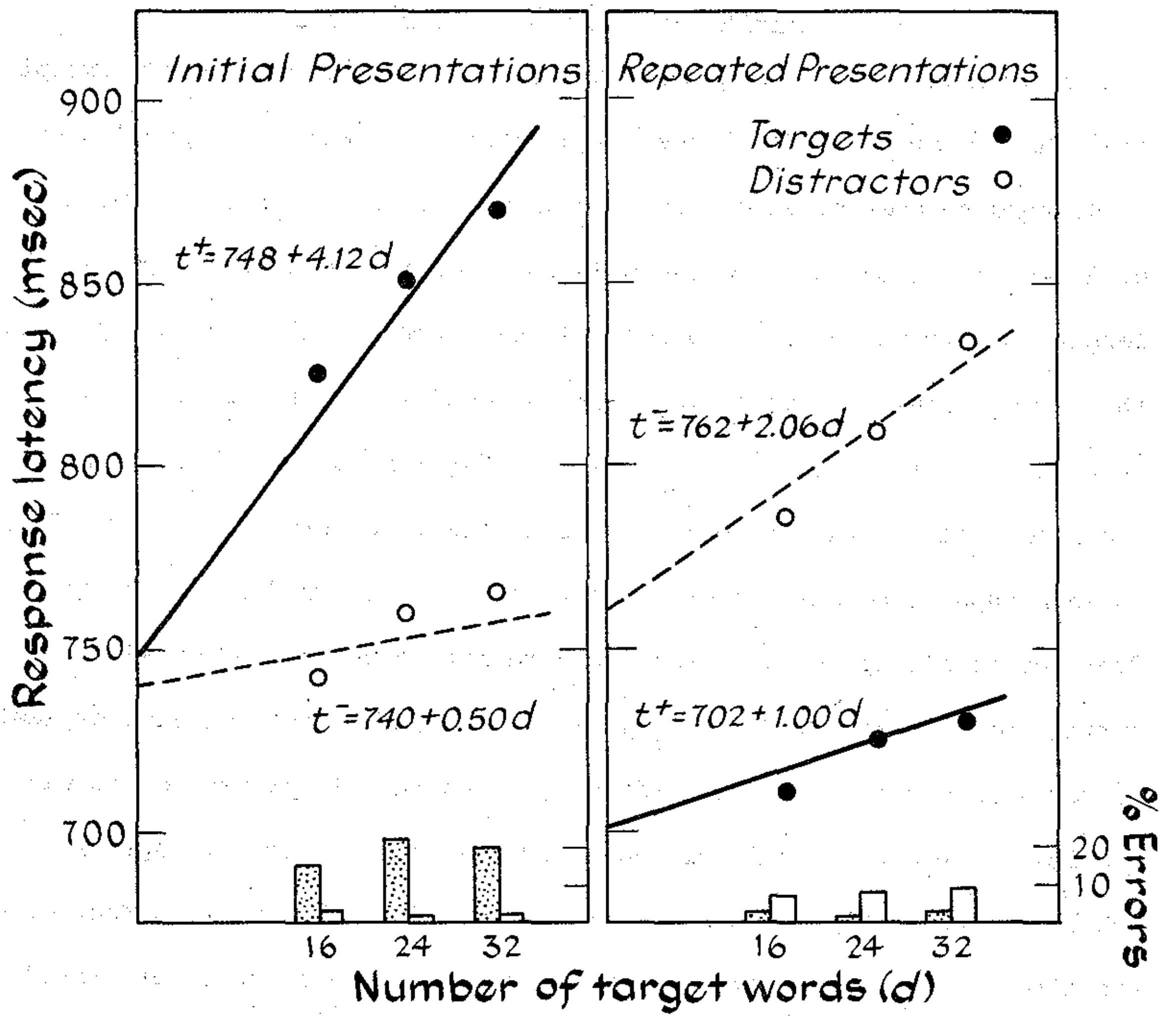

Fig. 2. Mean response latency and error percentages as functions of the length of the target list (d); the data represents a weighted average of response latencies from Blocks III and IV. The left panel presents the data for initial presentations of target and distractor words, and the right panel presents the data for repeated presentations. Incorrect responses to target words are given by the shaded bars, and errors to distractors by open bars (Experiment 2). The linear functions fitted to the data are explained in a later section.

mean latencies for repeated presentations (weighted averages of those words occurring for the second, third, and fourth times). Similarities to the results of Experiment 1 can readily be pointed out: (a) Positive response latency is greater than negative latency on initial presentations, but this order is reversed for repeated tests. (b) Most of the errors to target words occur on initial presentations, whereas most errors to distractors occur on repeated presentations.

The number of target words affected response latency for all types of trials, the effect being strongest for initial presentations of target words and for repeated tests of distractors. The magnitude of the effect on response latency of adding a single word to the target set can be approximated by the slopes of straight-line fits to the data in Fig. 2. The average slope is about 2.0 msec per word for the data of Experiment 2; this value is slightly less than that obtained for a similar experiment (Juola, Fischler, Wood, and Atkinson, 1971), but is much less than the 38 msec per digit obtained for small target sets in Sternberg's short-term memory experiment (Sternberg, 1966).

A Model for Recognition

The model to be considered has been presented elsewhere (Juola, et al., 1971) to account for latency and error data from recognition experiments like those reviewed in this paper. The model is similar to Kintsch's theory for recognition learning (Kintsch, 1967), but the processes associated with the memory states have been changed to account for response latencies as well as hit and false alarm rates.

It is assumed that each test word has associated with it a familiarity measure that can be regarded as a value on a continuous scale.

The familiarity values for targets are assumed to have a mean that is higher than the mean for distractors, although the two distributions may overlap. In many recognition studies (e.g., Shepard and Teghtsoonian, 1961) the target set is not well-learned, but involves stimuli that have received only a single study presentation. Under these conditions the subjective familiarity of the test stimulus leads directly to the decision to make an A_1 or A_0 response; i.e., the subject has a single criterion along the familiarity continuum which serves as a decision point for making a response. Familiarity values that fall above the criterion lead to an A_1 response, whereas those below the criterion lead to an A_0 response (Parks, 1966).

The present studies differ from most previous recognition experiments in that the target stimuli are members of a well-memorized list. In this case, it is assumed that subjects can use their familiarity measure to make an A_1 or A_0 response as soon as the test stimulus is presented, or they can delay their response until a more extensive memory search has confirmed the presence or absence of the test item in the target set. This process is shown in Fig. 3. If the initial familiarity value is either above a high criterion (c_1) or below a low criterion (c_0) the subject outputs a fast A_1 or A_0 response, respectively. If the familiarity associated with the test stimulus is of an intermediate value, the subject will be less confident about which response to choose. Since instructions emphasize correct responding, the subject is likely to make a more extensive search of memory (perhaps including a scan of the target list) in seeking a match for the test stimulus.

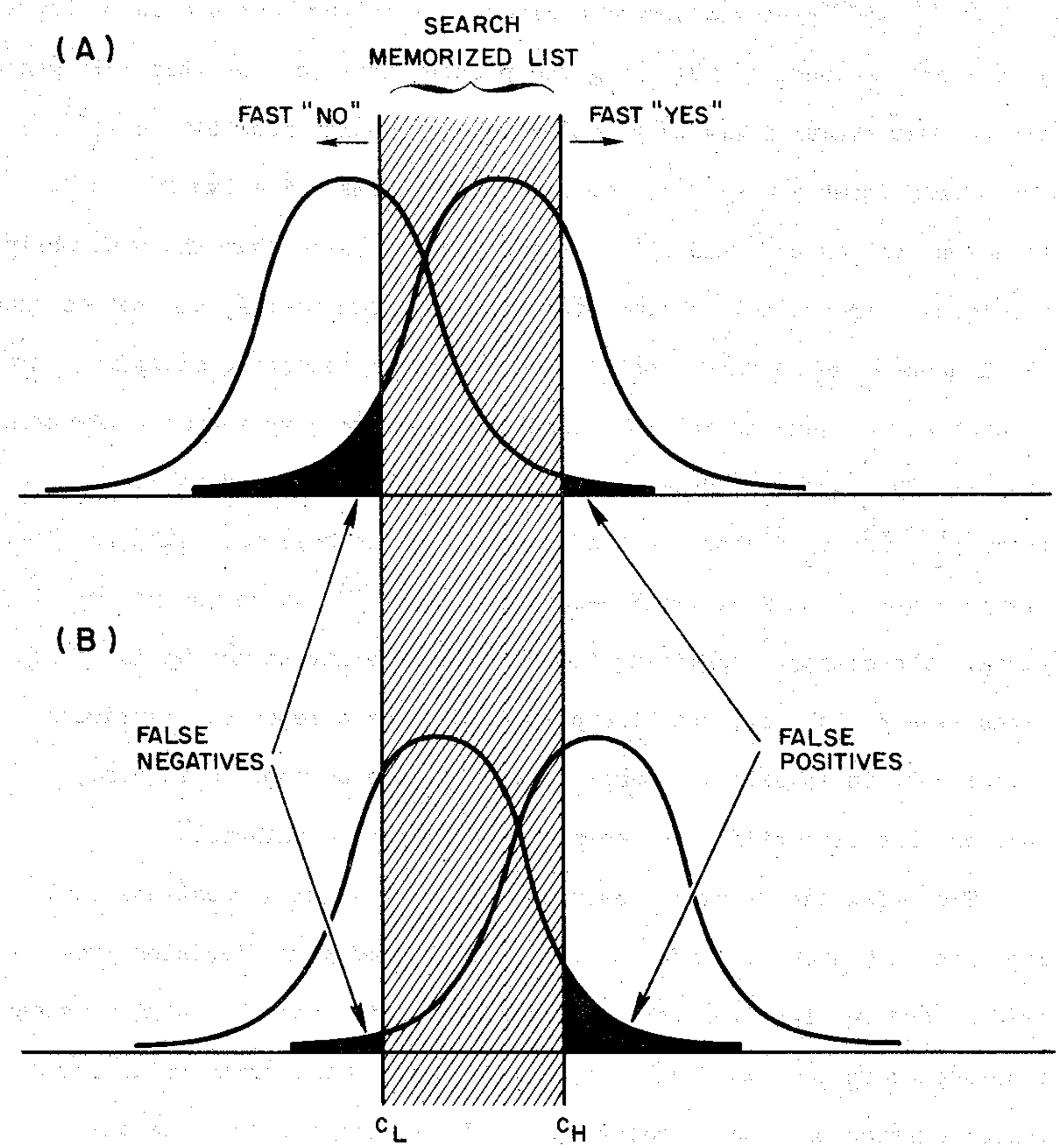

Fig. 3. Distributions of subjective familiarity values for distractor words (left) and target words (right) on the familiarity continuum. Panel A represents the relative locations of the distributions at the start of the session, whereas Panel B shows the increase in the means that occurs for both distributions after the target and distractor words have been tested.

On the n^{th} presentation of a given item in the test sequence, there is a density function reflecting the probability that the item will generate a particular familiarity value x; the density function is $\phi_1^{(n)}(x)$ for target items and $\phi_0^{(n)}(x)$ for distractor items. The two functions have mean values $\mu_1^{(n)}$ and $\mu_0^{(n)}$, respectively. (Note that the superscript n refers to the number of times the item has been tested, and not to the trial number of the experiment.) The effect of repeating specific target or distractor items in the test sequence is assumed to increase the mean familiarity value for these stimuli. This is illustrated in Fig. 3 where $\mu_1^{(n)}$ and $\mu_0^{(n)}$ shown in the bottom panel $(n > 1)$ have both shifted to the right of their initial values $\mu_1^{(1)}$ and $\mu_0^{(1)}$ shown in the top panel. The effect of shifting the mean familiarity values up is to increase the probability that the presentation of a repeated distractor will result in an extended memory search before a response is made, whereas this probability is decreased for repeated targets.[4]

The model can be stated mathematically by writing equations that represent the sums of times for the various memory and decision processes involved in recognition. The probability that the subject makes a correct response is assumed to be 1.0 if the familiarity value for a tested distractor word is below c_1 or if the familiarity value for a target is above c_0; i.e.:

$$Pr^{(n)}(A_1|S_1) = \int_{c_0}^{\infty} \phi_1^{(n)}(x)dx = 1 - \Phi_1^{(n)}(c_0) \qquad (1)$$

$$\Pr^{(n)}(A_0|S_0) = \int_{-\infty}^{c_1} \phi_0^{(n)}(x)dx = \Phi_0^{(n)}(c_1) \qquad (2)$$

Note that $\Phi(\cdot)$ designates the distribution function associated with the density function $\phi(\cdot)$.

In deriving response latencies, we shall assume that the processes involved in encoding the test stimulus, retrieving from memory information about the test stimulus, making a decision about which response to choose on the basis of this information, and emitting a response can be represented as successive and independent stages. These stages are diagrammed in the flow chart in Fig. 4. When the test stimulus is presented, the first stages involve encoding the item and executing a rapid search of long-term memory. This initial search will yield only a limited amount of information, but it will suffice to permit the subject to arrive at an index (x) of the subjective familiarity of the test stimulus. The time required to execute these two stages are combined and represented by the quantity ℓ in Fig. 4. The next stage is to arrive at a recognition decision on the basis of x. If $x < c_0$ a negative decision is made; if $x > c_1$ a positive decision is made. These decision times are functions of the value of x, and are given by the functions $\tau_0(x)$ and $\tau_1(x)$, respectively. If $c_0 \leq x \leq c_1$, an extended search of long-term memory is required, yielding more complete information about the test stimulus. The length of time needed for this search is assumed to be a function of d, the number of stimuli in S_1. The total time for a decision in this case is $\kappa(x) + \theta_i(d)$. In this equation $\kappa(x)$ denotes the time to make the decision to execute an extended search and may depend upon x. The

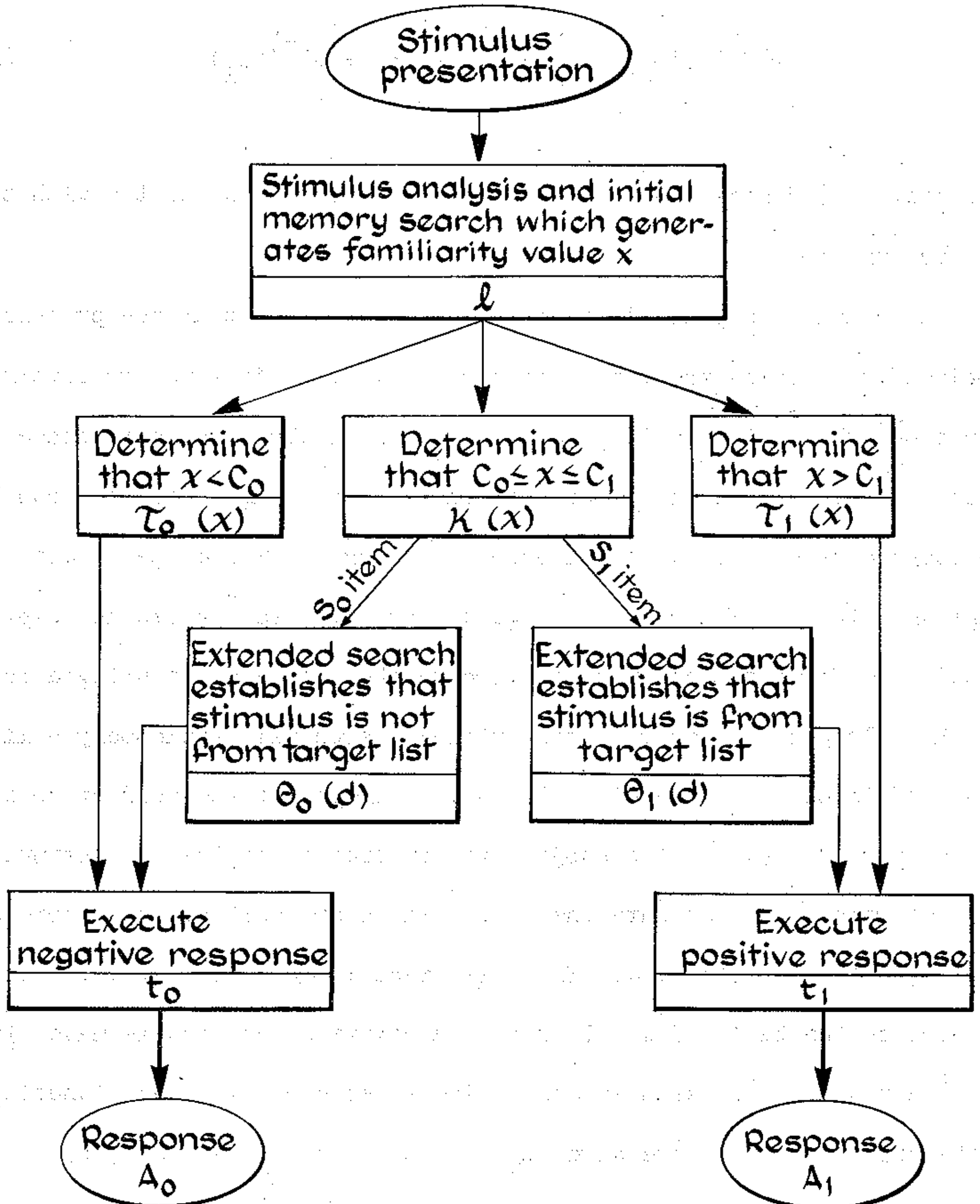

Fig. 4. Flow chart representing the memory and decision stages involved in word recognition. When a stimulus is presented, the subject arrives at a familiarity index x, and on that basis (1) decides to output a fast positive response (if $x > c_1$), or a fast negative response (if $x < c_0$), or (2) to execute a more extensive search of memory before responding (if $c_0 \le x \le c_1$).

function $\theta_i(d)$ is the time to complete the search and depends upon the length of the target list d and upon whether the tested item is a target (i=1) or a distractor (i=0). The final stage of the process is to output a response once the decision has been made, the response time being t_0 for an A_0 response and t_1 for an A_1 response.

Equations can be derived for response latencies by weighting the times associated with each stage by the probability that the stage occurs during processing. The expected time to make an A_i response to the n^{th} presentation of a particular stimulus drawn from set S_j (for i,j = 0,1) is as follows:

$$t^{(n)}(A_1|S_1) = \frac{\int_{c_1}^{\infty} \tau_1(x)\phi_1^{(n)}(x)dx + \int_{c_0}^{c_1} [\theta_1(d)+\kappa(x)]\phi_1^{(n)}(x)dx}{1 - \Phi_1^{(n)}(c_0)} + t_1 + \ell \quad (3)$$

$$t^{(n)}(A_0|S_1) = \frac{\int_{-\infty}^{c_0} \tau_0(x)\phi_1^{(n)}(x)dx}{\Phi_1^{(n)}(c_0)} + t_0 + \ell \quad (4)$$

$$t^{(n)}(A_0|S_0) = \frac{\int_{-\infty}^{c_0} \tau_0(x)\phi_0^{(n)}(x)dx + \int_{c_0}^{c_1} [\theta_0(d)+\kappa(x)]\phi_0^{(n)}(x)dx}{\Phi_0^{(n)}(c_1)} + t_0 + \ell \quad (5)$$

$$t^{(n)}(A_1|S_0) = \frac{\int_{c_1}^{\infty} \tau_1(x)\phi_0^{(n)}(x)dx}{1 - \Phi_0^{(n)}(c_1)} + t_1 + \ell \qquad (6)$$

In fitting the model to data from the previous experiments, several special cases will be examined. First, we shall assume that $\phi_i^{(n)}(x)$ is normally distributed with unit variance for all values of i and n. The function $\kappa(x)$ will be assumed to be a constant function of x, with value k. Finally, the function $\tau_i(x)$ will be of the following form:

$$\tau_i(x) = a_i e^{-|c_i - x|b_i} \quad . \qquad (7)$$

The function $\tau_1(x)$ is defined only for $x > c_1$, and $\tau_0(x)$ for $x < c_0$. Equation (7) can be simplified by assuming that both $\tau_1(x)$ and $\tau_0(x)$ have the same value at c_1 and c_0, respectively, (i.e., $a_1 = a_0$), and that they decrease symmetrically as the value of $|c_i-x|$ increases (i.e., $b_1 = b_0$). Two cases of this expression will be given special consideration. First, if $b_1 = b_0 = b = 0.0$, then $\tau_i(x)$ is a constant function of x;

$$\tau_i(x) = a \ . \qquad (8)$$

Second, if $a_1 = a_0 = k$, $\tau_i(x)$ has the same value as k when $x = c_i$,

$$\tau_i(x) = ke^{-|c_i - x|b} \quad . \qquad (9)$$

Finally, the function $\theta_i(d)$ must be specified. This function represents an extended search of long-term memory, and is assumed to be a linear function of the target set size. Two cases we wish to consider differ in the relative length of the memory search for target and distractor items. First, it can be assumed that the search times are identical for both types of items; i.e.,

$$\theta_1(d) = \theta_0(d) = \alpha d \ . \tag{10}$$

Alternatively, it might be that the length of the memory search is shorter on positive trials than on negative trials. This situation would occur if each list item is stored in a separate memory location, and the subject retrieves the contents of each location in seeking a match for the test stimulus. When a match is obtained, the search ends, otherwise all the memory locations are checked. The time for this process is:

$$\theta_1(d) = \alpha[\frac{d+1}{2}] \tag{11a}$$

$$\theta_0(d) = \alpha d \ . \tag{11b}$$

It should be noted that the two memory-search processes described above correspond to the exhaustive and self-terminating cases of the serial scanning model described by Sternberg (1969). While Sternberg's models have proved to be extremely valuable in interpreting data from a wide variety of memory-search experiments, good fits between the models and data do not necessarily require that the underlying psychological process be serial in nature. There are alternative models, including parallel scanning models, that are mathematically equivalent to those proposed by Sternberg and yield the same predictions as Eqs. (10) and (11) (Atkinson, Holmgren, and Juola, 1969). Thus, the use of Eqs. (10)

and (11) to specify the time associated with the extended memory search does not commit us to either a serial or parallel interpretation.

The model as it is now formulated predicts differences in performance as a function of the number of times an item has been tested. However, no mechanism has been incorporated to take into account improvements in performance resulting from extended practice on the task. An inspection of the data in Fig. 1 indicates that practice effects are occurring; for example, in Experiment 1 the first presentation of a distractor item in Block I produces a response latency of 819 msec, whereas the first presentation of a distractor item in Block IV has a latency of 721 msec. The theory can be amended to take into account generalized practice effects by assuming that t_0 and t_1 decrease over trials. For some experiments a meaningful analysis of the data requires an estimate of changes in t_0 and t_1 with practice. For others the problem can be sidestepped by restricting the analysis to the later trial blocks, if it can be assumed that t_0 and t_1 have reached some asymptotic level.

Theoretical Predictions for the List-length Study

The model will now be used to generate predictions for the latency and error data from Experiment 2. To avoid dealing with practice effects in this experiment, we shall confine our analysis to the data from Blocks III and IV where it seems reasonable to assume that performance is asymptotic.

Initially a value must be arbitrarily assigned to either c_0, c_1, $\mu_0^{(1)}$, or $\mu_1^{(1)}$ as a scaling parameter. Once this is done, the other parameters can be estimated from the data. We will let $c_0 = 0.0$. From the error data it is possible to estimate $\mu_1^{(n)}$, since an error to a target word occurs only if its familiarity value lies below c_0; i.e.,

$$Pr^{(n)}(A_0|S_1) = \Phi_1^{(n)}(c_0) \tag{12}$$

The error proportions over the last two trial blocks for the first, second, third and fourth presentation of a target item (averaged across the three list-length groups) were as follows: 0.171, 0.016, 0.014, and 0.007. Using the normal probability distribution it is possible to calculate that $\mu_1^{(1)} = c_0 + 0.950 = 0.950$. Similarly, $\mu_1^{(2)} = 2.14$, $\mu_1^{(3)} = 2.20$, and $\mu_1^{(4)} = 2.46$. The same procedure can be used to arrive at mean familiarity values for distractor words since,

$$Pr^{(n)}(A_1|S_0) = 1 - \Phi_0^{(n)}(c_1) \ . \tag{13}$$

The error proportions for presentations one through four for distractors were 0.005, 0.039, 0.049, and 0.049, respectively. Thus $\mu_0^{(1)} = c_1 - 2.58$, $\mu_0^{(2)} = c_1 - 1.76$, $\mu_0^{(3)} = c_1 - 1.66$, and $\mu_0^{(4)} = c_1 - 1.66$.

With c_0 set equal to zero and $\mu_i^{(n)}$ estimated from the error data, the remaining parameters can be estimated from the latency data. Four models of the theory will be used to generate fits to the data of Experiment 2. The models differ in the functions $\tau_i(x)$ and $\theta_i(d)$ as outlined below:

		$\theta_i(d)$ Eq. (10)	$\theta_i(d)$ Eq. (11)
$\tau_i(x)$	Eq. (8)	Model 1	Model 2
$\tau_i(x)$	Eq. (9)	Model 3	Model 4

The parameters that remain to be estimated are somewhat different for Models 1 and 2 versus 3 and 4. For Models 1 and 2 there are five parameters: c_1, α, $(k + t_0 + \ell)$, $(k + t_1 + \ell)$ and $(a + t_0 + \ell)$. The quantities in parentheses indicate that the component parameters cannot be evaluated separately; only their sum can be estimated. For Models 3 and 4 there are 6 parameters: c_1, α, k, b, $(t_0 + \ell)$ and $(t_1 + \ell)$.

Our method for parameter estimation involves the data presented in Fig. 5; it is simply the weighted average of the data for the third and fourth trial blocks of Experiment 2. Parameter values are selected that minimize the sum of the squared deviations (weighted by the number of observations) between the data points in Fig. 5 and theoretical predictions. A number of problems are involved in minimizing the squared-deviation function analytically, and consequently a computer was programmed to carry out a systematic search of the parameter space until a minimum was obtained accurate to three places. The weighted sum of squared deviations for the models are as follows:

Model 1: 3.81×10^5

Model 2: 4.57×10^5

Model 3: 4.35×10^5

Model 4: 4.68×10^5

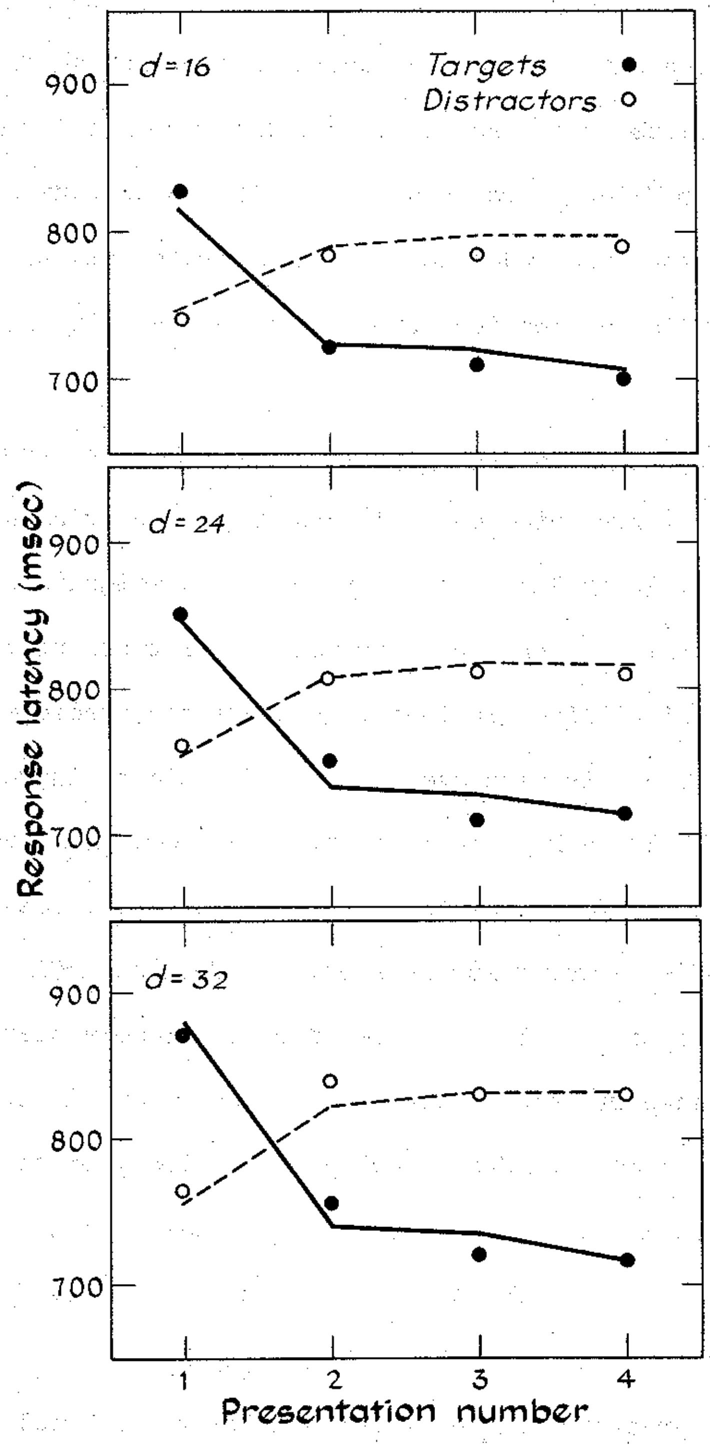

Fig. 5. Mean response latency as a function of presentation number for target and distractor words for three different list length (d) conditions. The top panel presents data for d = 16, the middle panel for d = 24, and the bottom panel for d = 32. The broken lines fitted to the data were generated from Model 1 (Experiment 2).

Model 1 clearly yields the best fit, Model 3 is second; both Models 1 and 3 assume that the extended memory search is represented in Eq. (10). The parameter estimates for these two models are given in Table 1. The predicted values for Model 1 are presented in Fig. 5 as connected lines; it should be noted that the model not only fits these data but (due to the method of parameter estimation) provides a perfect fit to the error data.

The results in Fig. 5 can be replotted by considering those items receiving their first presentation (n=1) and those receiving a repeated presentation (n = 2, 3, or 4); in the latter case a weighted average must be taken. If this is done the data points are those presented in Fig. 2, and the straight lines in that figure are the predicted functions based on Model 1. The fits displayed in Fig. 2 could be improved upon somewhat, but it should be kept in mind that they were obtained using parameter estimates based on a different breakdown of the data.

The latency of an error response should be fast according to the theory, since errors occur only when the secondary memory search is by-passed. The data support this prediction, and accord well with the values generated by Model 1. Specifically, the latency of an error is close to the predicted value of $\ell + t_0 + k = 731$ msec for an S_1 item, and to $\ell + t_1 + k = 687$ msec for an S_0 item. A more detailed account of error latencies is given in Juola, et al., (1971).

A verbal interpretation of the results in terms of Model 1 would proceed as follows: When a target item is presented for the first time, the probability that an extended memory search will occur before a response is made exceeds the probability that a fast positive response

Table 1

Parameter values for the two best-fitting models (Experiment 2)

Model 1	Model 3
$c_1 = 1.02$	$c_1 = .822$
$\alpha = 9.86$ msec	$\alpha = 14.1$ msec
$(k+t_0+\ell) = 868$ msec	$a=k = 144$ msec
$(k+t_1+\ell) = 824$ msec	$b = .320$
$(a+t_0+\ell) = 731$ msec	$(t_0+\ell) = 653$ msec
$(a+t_1+\ell) = 687$ msec*	$(t_1+\ell) = 609$ msec

*Not estimated, but computed from the above three parameters.

will be emitted on the basis of the item's familiarity value alone. The opposite is true for initial presentations of distractors; most trials result in fast negative responses. Thus the mean latency is longer for initial presentations of targets than for initial presentations of distractors ($k > a$), and the list-length effect is greater for targets than for distractors. The effect of repeating tests of words is to increase the familiarity of both targets and distractors. This results in an increased mean latency for responses to distractors (since a greater proportion of trials results in an extended memory search before a response) and a decrease in response latency to targets. The magnitudes of the list-length effects are observed to change concomitantly. Although such variables as number of presentations and number of intervening items between successive presentations affect an item's familiarity value, it is probable that all target items are about equally familiar at the start of the session. Any deviations that exist between the values are most likely due to properties of specific words or idiosyncratic responses on the part of the subject. It is apparent that the target item's familiarity cannot be assumed to depend upon its serial position in the study list, since no serial position effects have been observed.

Effects of Number of Occurrences in Study List

Experiment 3 was designed to test the effects of repeating words in the study lists. Fifteen subjects each memorized a list of 32 words, but some of the words were repeated either once or twice in the list. Specifically, the lists contained eight single words, six words that occurred twice, and four words that occurred three times. The order of the words within the lists was randomized with the constraint that at least four words would occur between successive occurrences of a repeated word. As in the previous experiments, subjects were instructed to learn the list in serial order, and they were tested for serial recall of all 32 items before the recognition tests began. The test session was divided into three consecutive trial blocks of 36 trials each. Within each block, the 18 target words were tested once, along with 18 distractors. Different sets of distractor words were presented in each block.

The results showed that mean latency was significantly shorter for responses to target words that were repeated in the study lists than for responses to those that occurred only once. This effect was obtained in all three blocks. The model that generated the best fit to the data of Experiment 2 (Model 1 of the previous section) was also used to fit the data of Experiment 3. There are at least two ways to account for the effects of repetition of words in the study list within the framework of the model. First, if the extended memory search involves the retrieval of the memorized list and a check for a match with the test stimulus, the expected length of time before a match is found is an inverse function of the number of times the target item occurs in the list. The fact that the best-fitting model assumes an exhaustive memory search, however, makes this analysis seem to be somewhat untenable; the evidence from Experiment 2 suggests that the length of the extended memory search is the same for positive and negative trials. Therefore, proposing a self-terminating scan to account for the repetition effects in Experiment 3 would seem to be theoretically inconsistent.

A second alternative would be to let the expected familiarity value for a target word be an increasing function of the number of times that the word occurred in the target set. It has been previously assumed that all target words initially have about the same familiarity value, for no experiment has shown positive response latency to be a function of the serial position of the target word. An analysis of the error data from Experiment 3 indicates, however, that repetitions in the study list increase the expected familiarities for those target words. Mean error proportions were .051, .034, and .011 for words that occurred one, two, and three times, respectively, in the target lists. Moreover, the mean error proportions for all three types of positive test trials were observed to decline across blocks (which is expected, since all target words were presented in each block). However, the mean error proportion for distractors was about .006 in all three blocks (which is also expected since distractors were not repeated from one block to the next).

The patterns observed in the data for Experiment 3, as shown in Fig. 6, can be fit by the model using the same procedure as in the previous section. The mean familiarity value for an item can be expressed as the distance from the appropriate criterion that will generate the observed error probability. The fit to the data was obtained by using Model 1, and retaining the same parameter values estimated from Experiment 2. The only difference was that in this case additional estimates of t_0 and t_1 had to be made for each trial block of the experiment. Under these conditions the predictions for Model 1 are represented by the curves in Fig. 6. As we see, the model's predictions accord well with the observed values.

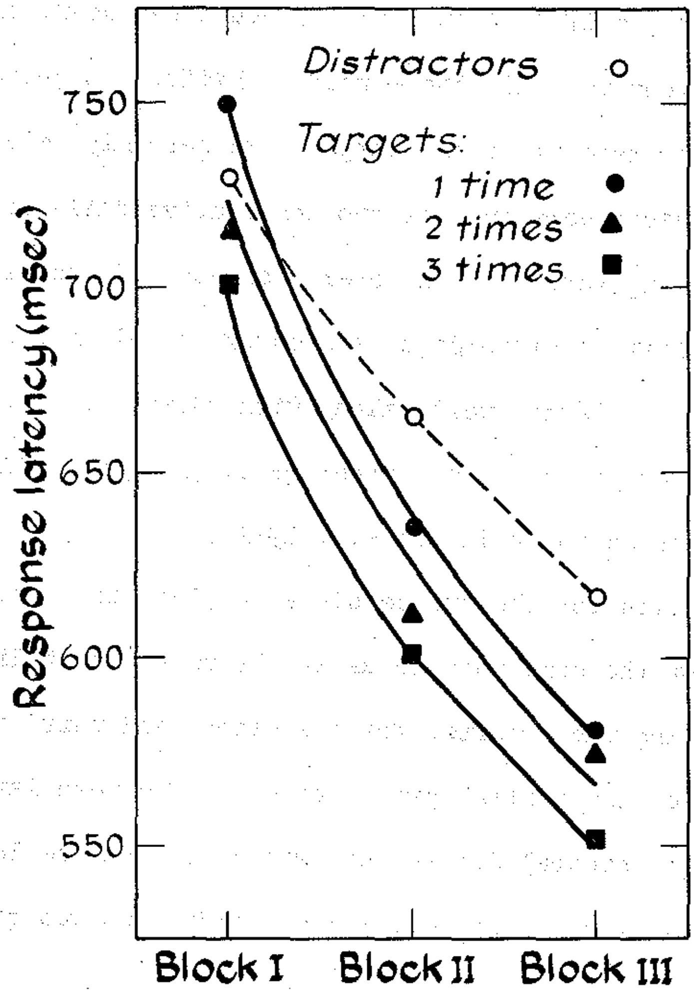

Fig. 6. Mean response latency for distractor words and for target words as a function of the number of times the word occurred in the target list for three consecutive trial blocks. The curved lines fitted to the data were generated from Model 1 (Experiment 3).

Effects of Similarity of Distractors to Target Words

For Experiment 4, a list of 16 pairs of nouns was used, eight were synonym pairs and eight were homophone pairs. Fifteen subjects received different study lists made up of one target word randomly selected from each pair. The subjects were run for two consecutive daily sessions of 96 trials each. During both sessions every target word was presented three times each. Distractor words consisted of the eight synonyms and eight homophones of the target words along with additional neutral words. (A more complete description of this experiment, along with lists of the word pairs, is presented in Juola, et al., 1971.)

The mean latencies for the two sessions combined are given in Table 2. The results show the same pattern as in the previous studies, with response latency being much shorter for the second and third tests of target words than for the initial presentations. The mean latency of negative responses to neutral distractors was less than the latency of responses to initial presentations of target words, but was greater than that of subsequent tests.

Negative response latency to synonyms of target words was significantly greater than the latency of responses to neutral distractors. Somewhat surprisingly, the latency for homophones exceeded that of responses to synonyms. After the data had been collected, a closer examination of the homophone pairs revealed that they could be divided into two categories: those that were visually quite similar to each other (only two letters different; e.g., bored and board) and those that were visually dissimilar (more than two letters different; e.g., sense and cents). The mean response latency to homophone distractors

Table 2

Mean response latencies (msec) for the first, second and third presentations of target words, and for distractors that vary in similarity to target words (Experiment 4)

Test word type	Latency
Targets:	
Presentation 1	733
Presentation 2	622
Presentation 3	617
Distractors:	
Homophones	793
Synonyms	731
Neutral words	673

which were classified as being visually similar to their respective target-word pairs was 895 msec, whereas latency to those that were visually dissimilar was 716 msec (not significantly greater than the latency to neutral distractors).

In terms of the proposed model, it appears that the familiarity of any distractor item can be increased by including items that are similar to it in the target set. Specifically, it seems clear that semantic information is used by the subject in determining whether or not a test item is on the target list. Acoustic information apparently is not by itself an important determinant of an item's familiarity, since the relatively long latencies to homophones appear to be due to visual similarities between the words of the homophone pairs. The cause of the effect of visual similarity is not entirely clear. It may be the case that some visual information is used in judging the familiarity of the presented stimulus. It is also possible, and perhaps more likely, that visual similarity between targets and distractor words leads to confusions and errors of identification of the test stimulus before the subject can decide exactly what word is being presented. Thus the effect of visual similarity could be due either to an increase in the distractor item's expected familiarity value μ_0, or it could increase the expected time for the encoding process, thereby raising the length of time for stage ℓ.

Frequency, Concreteness, and Number of Syllables

The words used in Experiment 5 were selected so that they could be separated into two distinct, equal-sized groups on the basis of any of three criteria: frequency in English, abstractness-concreteness, and number of syllables (one or two). Sixty-four five-letter words were used, most of these taken from the Paivio, Yuille, and Madigan (1968) noun list. The frequent words were rated either A or AA according to the Thorndike and Lorge (1944) word count, and the infrequent words were those which occurred fewer than 10 times per million. Similarly, concrete nouns were rated higher than 6.0 and abstract nouns were less than 3.0 on the Paivio-Yuille-Madigan scale. Additional words that were not present in the initial norms but were needed to complete the design were selected from the Thorndike and Lorge word lists. Three independent judges were used to pick those words that seemed to best match the originally selected words on the concreteness dimension.

The critical dimensions arranged the word pool into a 2 X 2 X 2 factorial design with eight words in each cell. For each subject, four words were randomly selected from every cell to make a list of 32 target words. During the test session, all 64 words were shown in two successive random orderings to yield 128 trials.

Mean response latencies were found for each type of word for positive and negative trials separately. Means were then taken across 16 subjects, and the results are presented in Table 3. The differences between levels of the three variables were in the same direction for both positive and negative latencies, with responses to two-syllable words being faster than responses to one-syllable words, those of concrete words being faster than those of abstract words, and those of infrequent words being faster than those of frequent words. However, separate analyses of variance performed on the data for positive and negative responses showed that the only significant differences were between frequent and infrequent target words and between abstract and concrete distractors. The

Table 3

Mean response latencies (msec) as functions of frequency, concreteness, and number of syllables (Experiment 5)

	Target words	Distractor words
High frequency	790	830
Low frequency	762	805
Abstract	785	839
Concrete	767	795
One syllable	782	824
Two syllables	770	810

result for targets is similar to Shepard's (1967) finding that the probability of correct recognition of an "old" word occurring in a long sequence of words is greater if the old word is a relatively uncommon word in English. Presumably, the effect of prior study for an infrequent word is to cause a greater change in its subjective familiarity than that produced for frequent words. When an infrequently occurring word is presented as a test, and it has a relatively high familiarity value, the subject is apparently more likely to output a response on the basis of its familiarity value alone than he is for frequent target words. Two explanations for this process are that the subject may either retrieve a higher familiarity value for infrequent target words than for frequent words, or he may adjust his criterion for a fast positive response so that c_1 is lower when a relatively rare word is tested. This latter argument is equivalent to saying that the subject compares the retrieved familiarity of a test word with its expected familiarity, which is a function of its frequency in English. If a large discrepancy occurs, the subject outputs a fast positive response before any further memory search is initiated.

Previous experiments (e.g., Gorman, 1961) have shown that recognition probability for old words is higher if the words are concrete rather than abstract. A similar effect favoring responses to concrete words was noted in Experiment 5, although it was significant only for distractor items. Since abstract and concrete distractors were balanced for frequency (and, presumably, familiarity), it appears that the concreteness dimension affects ℓ, the encoding and initial information retrieval stage of the model. The time between the stimulus onset and an estimation of the word's familiarity seems to be an increasing function of the abstractness of the distractor word, but a thorough explanation of the mechanism for this process cannot be made on the basis of the present data.

Although the number of syllables in the test word did not have a significant effect on response latency either for target or distractor words, it is interesting to note that response latencies were shorter for two-syllable words in both cases. Since all words were five letters in length it is worth noting that the two-syllable words generally contained more common spelling patterns than did the one-syllable words (e.g., <u>tenet</u> vs. <u>pique</u>). Presumably, words that contain more commonly occurring letter patterns should be easier to encode (decreasing the value of ℓ) than those that have less common patterns. The total response latency to two-syllable words should then be shorter than latency for one-syllable words if all other factors are equal. This explanation is admittedly tentative, and the results presented here were not obtained to provide tests for theories of word perception. However, the arguments follow naturally from the proposed theory and account for the observed latency effects.

Repression Effects on Recognition Latency

Experiment 6 was designed to determine if the procedure used in the previous studies could lend itself to the study of repression effects in recognition memory. Repression is here taken to mean forgetting which is selective to those perceptions and memories which produce anxiety. Supposedly, this type of forgetting is initiated by a mechanism that defends the subject against such anxiety. It should be possible to determine the extent of repression effects on recognition memory by using the paradigm of the previous studies. If certain target words are paired with anxiety-provoking stimuli, and then these words are presented in a recognition task, repression effects should interfere with the initial stages of processing and perhaps later search and decision processes as well. The expected result is that response latency should be greater for target words that have been associated with anxiety-provoking stimuli than for words that have been associated with neutral or positive stimuli.

Each subject memorized a target list of 16 words. The target lists were then divided into three consecutive groups of four words each (the first two and last two words were not included as they were to be tested only in a warm-up block at the start of the test session). One word from each group was assigned to one of four treatment conditions: (1) paired with a positive experience, (2) paired with a neutral experience, (3) paired with a negative experience, or (4) not paired with any experience. The positive and negative experiences were generated by the subjects. They were instructed to write down descriptions of the three most intense occasions when they had experienced feelings similar to those in, first, a list of positive emotionally-descriptive statements, and, second, a list of negative statements. Three neutral experiences were provided by the experimenter. These consisted of descriptions of routine events obtained from newspaper stories. Each of nine list words were randomly paired with one of the experiences. These pairings were achieved by having the subject write down four different associations between the given target word and the appropriate experience as assigned by the experimenter.

The test sequence was divided into a warm-up block of eight trials followed by three trial blocks of 24 trials each. All 12 experimental target words were presented once in each block along with 12 distractors, which were never repeated. The results showed that the four different treatment conditions for target words had a significant effect on positive response latency in the first trial block only. The data for Trial Block I are presented in Table 4. The latencies in the latter two blocks converged for all four conditions, with a mean response latency of 688 msec for Block II and 665 msec for Block III. Latencies for responses to distractor words were 733 msec and 711 msec for the last two blocks.

The data in Table 4 show that response latency was actually shortest for words that had been paired with negative experiences, although the mean latency was not significantly less than the time to respond to words paired with positive experiences. Both of these conditions resulted in latencies significantly below those for words paired with neutral experiences, and the mean latency for unpaired words was significantly greater than that of any other condition.

The results do not support the repression hypothesis of forgetting, namely, that (1) the pairing of negative experiences with target words will elicit anxiety when the words are presented in a recognition test, and (2) this anxiety will block the perception of the word or impede the retrieval of stored information about the word. Either process would result in response latency being greater for negatively-paired words than for words paired with neutral or positive experiences. The data indicate that the mere pairing of a target word to any type of experience results

Table 4

Mean response latencies (msec) to four types of target words and to distractor words (Experiment 6)

Test word type	Latency
Target words:	
Positive experience	755
Negative experience	733
Neutral experience	788
Unpaired	916
Distractor words	782

in a shorter mean latency to that word, especially if the experience is one from the subject's own background.

The interpretation, in terms of the model, is that any effort to associate a target word with a prior experience results in a subsequent higher familiarity value for that word. This effect is evidenced by the fact that responses to target words associated with negative or positive experiences were faster than responses to distractors. This result is unusual; in all of the studies reported here, response latency for target words is greater than response latency for distractors on their initial presentations. It appears that the familiarity value (rather than any positive or negative associations with the word) determines the speed with which the subject can make a recognition decision.

Recognition Latency for Words in a Semantic Hierarchy

Experiment 7 was designed to test for the effects of imposing an organizational scheme on the words of the target set. Specifically, all target words were taken from the semantic hierarchy shown in Table 5. The hierarchy of the present study is an expansion of two hierarchies used in a study by Bower, Clark, Winzenz, and Lesgold (1969). Of the 86 words in Table 5, each subject received a target set of 54. All the target sets included both words in Level 1 and the four words in Level 2. Twelve words were included from Level 3, these being either all four or only two of the exemplars of the Level 2 words. Similarly, either four or two exemplars were included under each of the Level 3 words. An example of one target set, in the form it was presented to the subject, is shown in Fig. 7. Different target sets were made such that every word within any level was used as a target equally often. Distractor words

Table 5

Hierarchical organization of 86 nouns (Experiment 7)

Organism							
Plant				Animal			
Vegetable	Flower	Tree	Fruit	Mammal	Insect	Bird	Fish
Carrot	Rose	Oak	Apple	Dog	Ant	Robin	Bass
Bean	Tulip	Elm	Orange	Cat	Mosquito	Eagle	Trout
Corn	Carnation	Pine	Pear	Cow	Fly	Sparrow	Shark
Pea	Daisy	Maple	Banana	Horse	Bee	Cardinal	Herring

Instrument							
Musical				Precision			
Brass	Woodwind	Percussion	String	Drafting	Optical	Surgical	Navigation
Tuba	Clarinet	Drum	Violin	Compass	Telescope	Scalpel	Radar
Cornet	Oboe	Cymbals	Banjo	Ruler	Microscope	Forceps	Gyroscope
Trumpet	Flute	Bells	Cello	Protractor	Monocle	Stethoscope	Sextant
Horn	Saxophone	Triangle	Guitar	Pen	Spyglass	Pincers	Altimeter

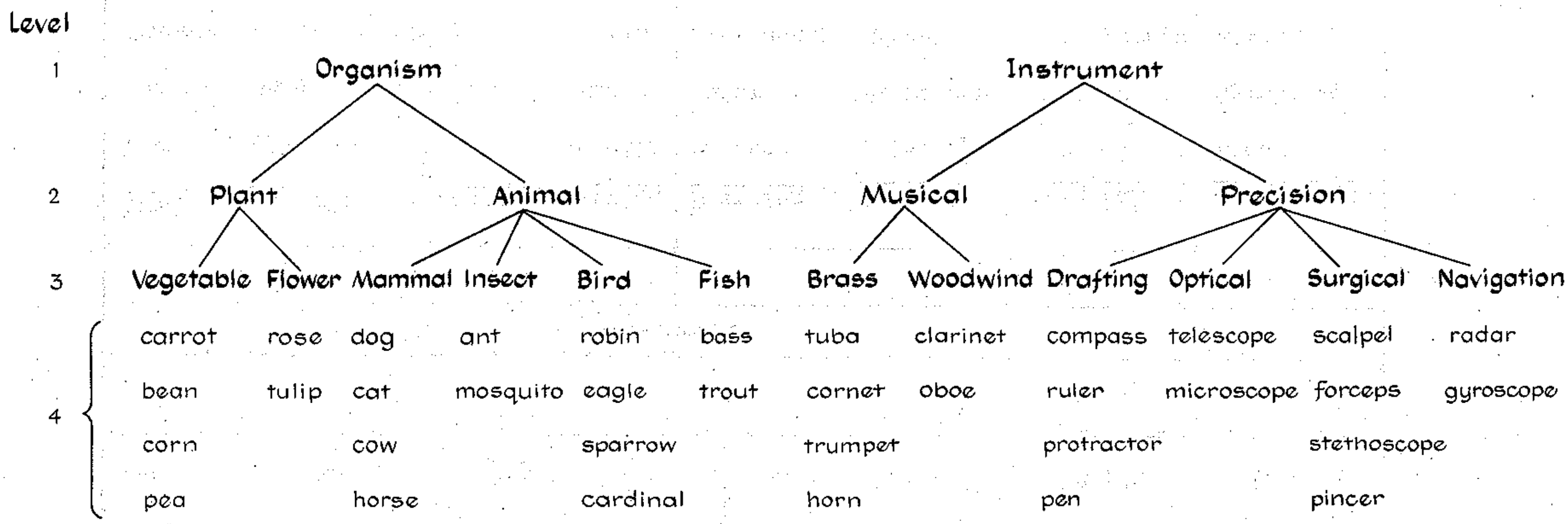

Fig. 7. An example of a semantic hierarchy containing the words presented to the subject as a target list. The level designation on the left side of the figure did not appear on the subject's copy (Experiment 7).

were chosen randomly from the Thorndike-Lorge list, and were matched with the words in Table 5 in length and frequency.

The subjects memorized the target set the day prior to the experiment; they were tested before the experimental session by filling in a hierarchy with blank lines drawn in where the target words appeared on the study sheet. No subjects made any errors on the recall task. During the test session all 54 target words were shown once, and 54 different distractors were also presented.

The data for the two halves of the target set were combined, and the mean latencies were found for each level. There was no effect of level within the hierarchy on positive response latency. The only significant effects were between words that were members of subsets of different sizes within Level 3 and Level 4. These subsets are denoted Level 3 (4 nodes) for words that are one of the four exemplars of a Level 2 node (e.g., mammal in Fig. 7), Level 3 (2 nodes) for words that are one of the two exemplars of a Level 2 node (e.g., woodwind), Level 4 (List 4-4) for words that were one of four items listed under one of a group of four nodes in Level 3 (e.g., scalpel), Level 4 (List 4-2) for words that were one of four items listed under one of a pair of nodes in Level 3 (e.g., carrot), Level 4 (List 2-4) for words that were one of two items listed under one of four nodes in Level 3 (e.g., ant), and Level 4 (List 2-2) for words that were one of two items listed under one of a pair of nodes in Level 3 (e.g., clarinet). The mean response latencies for all types of trials are shown in Table 6.

The results do not support a memory search that follows the structure of the semantic hierarchy of Fig. 7. Even if the target words are

Table 6

Mean response latency (msec) for words as functions of their locations in a semantic hierarchy (Experiment 7)

Test word type	Latency	
Target item:		
Levels 1 and 2	728	
Level 3 (4 nodes)	722	733
Level 3 (2 nodes)	755	
Level 4 (List 4-4)	693	716
Level 4 (List 4-2)	717	
Level 4 (List 2-4)	724	
Level 4 (List 2-2)	738	
Distractor items	708	

organized hierarchically in the subject's memory, it does not appear to be the case that a search through such an organization is necessary before a recognition decision can be made. Studies which have shown effects of hierarchical organization have typically employed a task which requires the subject to recall information about more than one word (Collins and Quillian, 1969; Meyer, 1970), whereas the present task can be performed adequately if only information about the test word is retrieved.

The significant results that were obtained are more difficult to interpret than the lack of effects due to the level in a hierarchy. It seems to be incompatible with earlier results that the subjects can respond more quickly to an item that is a member of a large subset than to one that belongs to a smaller subset. A tentative explanation for this result can be made on the basis of the data from Experiment 4. It was demonstrated that semantic similarity between a distractor word and a target can increase the familiarity of the distractor, resulting in a slower negative response time. In the present experiment, items that are most highly related semantically are those that share the same relative positions in the hierarchy. If the study of one word serves to increase the familiarity value of related words, then one would expect those words with the greatest number of similar words in the target set to have the highest mean familiarity. From this argument the prediction that positive response latency should be shortest to words that belong to a relative large subset can be made if the subset contains words that are semantically related. It is clear that any words at the same level in the hierarchy are closely related semantically, and the prediction that response latencies should be shorter to a word which is one of a subset of four items at a given level than to one which is one of a subset of two is upheld by the data in Table 6. The explanation offered here is like the one proposed by Schaeffer and Wallace (1969) to account for judgments of word meanings. As in the present study, semantic similarity between test words facilitated the decision that the words belonged to the same category.

Summary and Conclusions

A simplified version of the model proposed earlier is shown in Fig. 8. It is assumed that when a stimulus word is presented for a recognition test, the subject performs an initial, rapid access of the information stored about the test item. This information provides the subject with a familiarity rating for the word. Response decisions based on the familiarity of the stimulus alone can be made very quickly, but they result in a relatively high error rate. If the results of the initial memory search do not provide the subject with enough information to respond with confidence (i.e., if the familiarity value is neither very high nor very low) a secondary, extended memory search is performed before a response is emitted. This latter search virtually guarantees that the subject will arrive at the correct decision, but with a consequent increase in response latency. By adjusting the criteria for emitting responses based on familiarity alone, the subject can achieve a stable level of performance, matching the speed and accuracy of responses to the demand characteristics of the experiment.

The model provides a tentative explanation for the results of several recognition-memory experiments. The memory and decision stages are indicative of possible mechanisms involved in recognition; we do not,

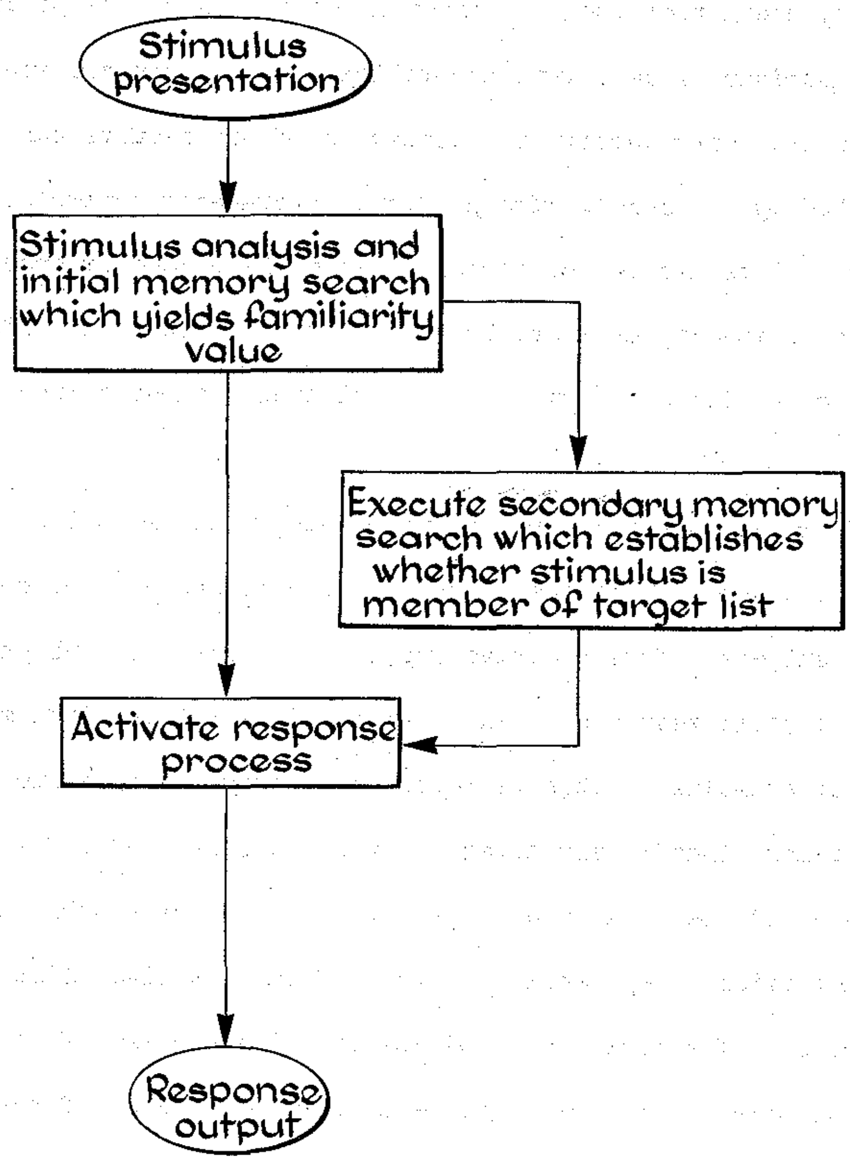

Fig. 8. Simplified flow chart representing the memory and decision stages involved in word recognition.

however, claim that they are exact descriptions of the processes involved. The comparison of data with theoretical predictions are reported mainly to show that many quantitative features of our results can be adequately described by the model. The particular parameter estimates reported are not to be interpreted too closely, since the parameter space was fairly shallow in the region where best fits were obtained; considerable play could be permitted in some of the estimates without seriously affecting the goodness-of-fit.

There are several encouraging points, however, which suggest that memory and decision states of the model correspond to processing stages of the subject. Introspective reports indicate that subjects may indeed output a rapid response based on initially retrieved information about the test stimulus. Subjects report that they are sometimes able to respond almost immediately after the word is presented before "knowing for sure" if the item is a target or not. The same subjects report that on other trials they recall portions of the memorized list before giving a response. The fact that subjects are always aware of their own errors also supports the general outline of the model; even if the initial familiarity of an item produces a decision to respond immediately, the extended memory search continues, and results in the subject confirming that he has or has not made the correct response. These introspective reports lend support to the general theoretical representation, and go beyond the goodness-of-fit demonstrations. No other model that we have yet considered combines these features into a workable alternative.

The data from our experiments provide clues regarding which features of the test stimuli affect the various memory and decision processes.

While the sources of these effects are difficult to isolate, some conclusions can be drawn. It appears that visual properties of the stimulus affect the ease with which it can be encoded and thus the speed of the initial memory search. Words that contain infrequent spelling patterns or are visually quite similar to one another can result in longer response latencies. Other factors seem to influence the familiarity of the tested word. These include the lag since the word was last activated in memory (either through exposure to certain words immediately before the test session, or through repeated tests during the session) and the presentation of items related to the test word.

While we can list manipulations that affect the familiarity of words, the definition of familiarity needs to be more precise. Familiarity apparently depends on the time since the last previous exposure or retrieval of the word. A more general formulation would involve the description of a word as a stimulus and how it is represented in memory. If a word is represented as a complex of features (Anisfeld and Knapp, 1968), each feature (semantic, acoustic, visual, etc.) might be independently time-tagged when activated in memory. This activation may occur when the word itself is tested, or when words similar on one or more dimensions are presented. Familiarity could then be based upon a feature count, with familiarity increasing with the number of tagged features (Kintsch, 1970).

A comparison between the data of the present experiments and those of short-term recognition studies is revealing. We have completed an experiment (Juola and Atkinson, 1971) that uses the same paradigm as Sternberg's (1966) study. The subjects were from the same pool as those who participated in the experiments reported here, and the same apparatus was used. Additionally, the stimulus words were the same as those used in Experiments 1 through 3. In the Juola-Atkinson experiment, each trial began with the auditory presentation of a target set of from one to four words, followed by the visual presentation of a single test word. All other features of the experiment were the same as for the studies reported in this paper. The results showed response latency to be a linear function of the target set size, with the following best-fitting straight line: 617 + 26d (d=1 to 4). In accord with Sternberg's findings, a much larger target-set effect was found in the short-term study (the slope parameter was about 38 msec in the Sternberg (1966) experiment). While the overall response speed was slower than that reported by Sternberg, it was still considerably below that of any of the long-term experiments reported in this paper.

Some conclusions can be drawn about similarities and differences between these two classes of experiments. Presumably, since new items are constantly being presented and a given stimulus may change from being a target on one trial to being a distractor on another, subjects do not rely on the test stimulus' familiarity in the short-term experiments. With the test set already in short-term memory it is possible to initiate the "extended" search (in the sense of the previous discussion) without first retrieving information from long-term memory. Since this type of search is target-set dependent, the greater set-size effects observed in short-term recognition studies are expected.

The overall response latencies appear to be greater in the long-term experiments than in the short-term recognition studies. This effect is presumably due to the fact that information must be retrieved from long-term memory in the former type of experiment before a decision can be made. If an extended memory search is necessitated, additional information must be retrieved resulting in an even longer response latency. It is interesting to note that for the models tested in the present paper, the best representation of the extended search process had the same mathematical form as the one proposed by Sternberg to account for the short-term studies (i.e., linear functions with equal slope for both targets and distractors). While the overall data from these two classes of experiments show marked differences, some features of the results suggest that there may be common processes involved in short-term and long-term recognition.

References

Anisfeld, M., and Knapp, M. Association, synonymity, and directionality in false recognition. Journal of Experimental Psychology, 1968, 77, 171-179.

Atkinson, R. C., Holmgren, J. E., and Juola, J. F. Processing time as influenced by the number of elements in a visual display. Perception and Psychophysics, 1969, 6, 321-326.

Bower, G. H., Clark, M., Winzenz, D., and Lesgold, A. Hierarchical retrieval schemes in recall of categorized word lists. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 323-343.

Collins, A. M., and Quillian, M. R. Retrieval time from semantic memory. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 240-247.

Fischler, I., and Juola, J. F. Effects of repeated tests on recognition time for information in long-term memory. Journal of Experimental Psychology, 1971, in press.

Gorman, A. M. Recognition memory for nouns as a function of abstractness and frequency. Journal of Experimental Psychology, 1961, 61, 23-29.

Juola, J. F., and Atkinson, R. C. Memory scanning for words versus categories. Journal of Verbal Learning and Verbal Behavior, 1971, in press.

Juola, J. F., Fischler, I., Wood, C. T., and Atkinson, R. C. Recognition time for information stored in long-term memory. Perception and Psychophysics, 1971, 10, 8-14.

Kintsch, W. Memory and decision aspects of recognition learning. Psychological Review, 1967, 74, 496-504.

Kintsch, W. Models for free recall and recognition. In D. A. Norman (Ed.), Models of Human Memory, New York: Academic Press, 1970.

Meyer, D. E. On the representation and retrieval of stored semantic information. Cognitive Psychology, 1970, 1, 242-300.

Paivio A., Yuille, J. C., and Madigan, S. A. Concreteness, imagery, and meaningfulness values for 925 nouns. Journal of Experimental Psychology Monograph Supplement, 1968, 76, No. 1, Part 2, 1-25.

Parks, T. E. Signal detectability theory of recognition memory performance. Psychological Review, 1966, 73, 44-58.

Schaeffer, B., and Wallace, R. Semantic similarity and the comparison of word meanings. Journal of Experimental Psychology, 1969, 82, 343-346.

Shepard, R. N. Recognition memory for words, sentences, and pictures. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 156-163.

Shepard, R. N., and Teghtsoonian, M. Retention of information under conditions approaching a steady state. Journal of Experimental Psychology, 1961, 62, 302-309.

Sternberg, S. High-speed scanning in human memory. Science, 1966, 153, 652-654.

Sternberg, S. Memory scanning: Mental processes revealed by reaction-time experiments. American Scientist, 1969, 57, 421-457.

Thorndike, E. L., and Lorge, I. The teacher's word book of 30,000 words. New York: Bureau of Publications, Teachers College, Columbia University, 1944.

Footnotes

[1] Preparation of this paper was supported by the National Science Foundation, Grant No. NSFGJ-443. The authors gratefully acknowledge the assistance of Christine Wood for collecting and analyzing the data for five of the experiments reported here. Experiment 1 was run by Ira Fischler and Experiment 6 by Michael Ryan. Thanks are also due to Dexter Fletcher for assistance in writing the computer program.

[2] The reduction in the strength of the interaction is apparently due to a lag effect. Juola, et al. (1971) demonstrated that as the number of items between successive presentations of any item increases, the effect of repetition decreases. In Experiment 1, the number of trials increases by 12 from one block to the next. Thus, the lag between successive presentations of the same test word also increases.

[3] There was no effect on response latency of the segment (beginning, middle, or end) of the target list that was tested. Therefore no further distinctions will be made between groups on this basis.

[4] In this paper we do not consider lag effects; i.e., the number of trials intervening between one presentation of an item and its next presentation. However, such effects do exist and can be significant under some experimental conditions. To account for lag effects, we would assume that μ_i increases immediately upon the presentation of an item, but over an extended period of time gradually drifts back to its initial value. For a discussion of this problem and data see Juola, et al., (1971).