6 Search and Decision Processes in Recognition Memory

Richard C. Atkinson
STANFORD UNIVERSITY

James F. Juola
UNIVERSITY OF KANSAS

INTRODUCTION

In this chapter we develop and evaluate a mathematical model for a series of experiments on recognition memory. The model is extremely simple, incorporating only those assumptions necessary for treatment of the phenomena under analysis. It should be noted, however, that the model is a special case of a more general theory of memory (Atkinson & Shiffrin, 1968, 1971); thus its evaluation has implications not only for the experiments examined here, but also for the theory of which it is a special case.

Before discussing the model and the relevant experiments, it will be useful to provide a brief review of the general theory. The theory views memory as a dynamic and interactive system; the main components of the memory system and paths of information flow are diagrammed in Figure 1. Stimuli impinge on the system via the *sensory register*, and the system in turn acts upon its environment through the *response generator*. Within the system itself, a distinction is made between the *memory storage network*, in which information is recorded, and *control processes* that govern the flow and sequencing of information. The memory storage network is composed of the *sensory register*, a *short-term store* (STS), and a *long-term store* (LTS). The sensory register analyzes and transforms the input from the sensory system

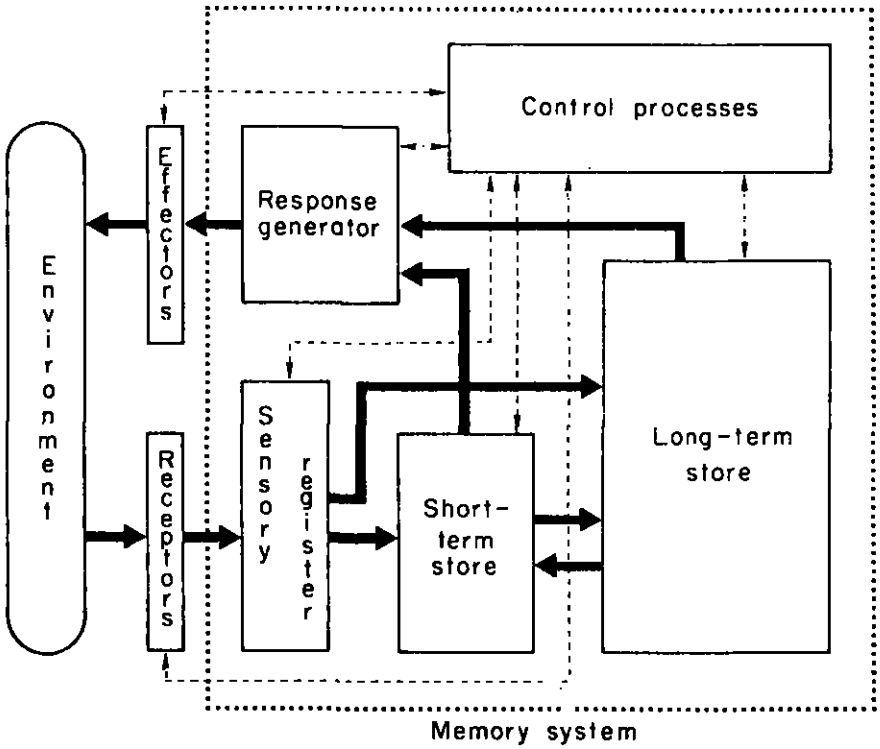


FIGURE 1.
A flowchart of the memory system. Solid lines indicate paths of information transfer. Dashed lines indicate connections that permit comparison of information arrays residing in different parts of the system; they also indicate paths along which control signals may be sent which modulate information transfer, activate rehearsal mechanisms, set decision criteria, alter biases of sensory channels, initiate the response generator, etc.

and briefly retains this information while it is selectively read into one of the memory stores. The STS is a working memory of limited capacity from which information decays fairly rapidly unless it is maintained by control processes such as imagery or rehearsal. The contents of STS may be thought of as the 'current state of consciousness' for the subject. The LTS is a large and essentially permanent memory bank; information once recorded in this store does not decay, but its availability for further processing depends upon the effectiveness of retrieval processes. In the figure, STS and LTS are depicted as two separate boxes, but this is not meant to imply neurologically separate systems; it is quite possible that STS is simply the active phase of neural processes quiescent in LTS. The control processes regulate the transfer of information from one store to another, and the sequencing of operations within each memory store. These processes are labile strategies adopted by the subject in response to environmental and task conditions. They include

selective attention, rehearsal, coding, selection of retrieval cues, and all types of decision strategies.

Although the model developed in this paper is a special case of the theory represented in Figure 1, it also can be interpreted as consistent with a number of other theories.[1] It is possible to theorize about components of the memory process without making commitments on all aspects of a theory of memory. Component problems can be isolated experimentally and local models developed. Work of this sort eventually leads to modification of the general theory, but a close connection between local models and the general theory is not required at every stage of research.

The term 'recognition memory' covers a wide variety of phenomena in which the subject attempts to decide whether or not a given object or event has been experienced previously (Kintsch, 1970a, 1970b; McCormack, 1972). It is a common process in everyday life and one that is readily subject to experimentation. In the recognition task that we have been investigating, the subject must decide whether or not a given test stimulus is a member of a predefined set of target items. For any set S of stimuli, a subset S_1 is defined that is of size d. Stimuli in S_1 will be referred to as *target items;* subset S_0 is the complement of S_1 with respect to S, and its members will be called *distractor items*. The experimental task involves a long series of discrete trials with a stimulus from S presented on each trial. To each presentation the subject makes either an A_1 or A_0 response, indicating that he judges the stimulus to be a target or distractor item, respectively.

The target sets in our experiments involve fairly long lists of words (sometimes as many as 60 words) that are thoroughly memorized by the subject prior to the test session. During the test session individual words are presented, and the subject's task is to respond as rapidly as possible, indicating whether or not the test word is a member of the target set. Errors are infrequent, and the principal data are response latencies (i.e., the time between the onset of the test word and the subject's response). The length of the target list and other features of the experimental procedure prevent the subject from rehearsing the list during the course of the test session, thus requiring that the subject access LTS in order to make a decision about each test word.

In some respects this task is similar to that studied by Sternberg (1966) and others. In the Sternberg task, a small number of items (e.g., 1 to 6 digits) are presented at the start of each trial, making up the target set for the trial. The test item is then presented, and the subject makes an A_1 response if the item is a member of that trial's target set, or an A_0 otherwise. In the Sternberg task the subject does not need to master the target set, for it is small and can be maintained in STS while needed. This type of short-term recognition experi-

[1] See, for example, a collection of papers concerning models of memory edited by Norman (1970).

ment differs then from our long-term studies in terms of the size and mastery of the target set. The data from the two types of studies are similar in many respects, but there are some striking differences. In both types of studies, response latency is an increasing linear function of the size of the target set; however, the slope of the function is about 5 msec per item in the long-term studies, compared with about 35 msec in the short-term studies. Other points of comparison will be considered later.

From a variety of long-term recognition studies we have achieved a better understanding of how information is represented in memory and how it is retrieved and processed in making response decisions. A model based on this work is formally developed in the next section. First, however, a more intuitive account is given.

Consider the case in which the target set consists of a long list of words that the subject has thoroughly memorized prior to the test session. The initial problems are to postulate mechanisms by which this information is used to distinguish target words from distractors. It is assumed that every word in the subject's language has associated with it a particular long-term memory location that we refer to as a node in the lexical store (Miller, 1969; Rubenstein, Garfield, & Millikan, 1970). When a word is presented for test, the sensory input is encoded and mapped onto the appropriate node. This process is essential in identifying or naming the test stimulus as well as in retrieving other information that is associated with the item. Figure 2 shows a representation of a single node in the lexical store (panel A), along with an example of an associative network by which various nodes are interconnected (panel B). Each node is a functional unit representing a single word or concept (such as the relational concepts 'to the left of,' 'above,' or concepts dealing with size and shape). A variety of nodes and their associations in the lexicon is necessary in accounting for language use and other symbolic behavior (Schank, 1972), but for our purposes we need only consider nodes that correspond to potential test words.

At each node is stored an array of codes. The input codes represent the end results of the encoding processes that operate on the auditory, pictorial, or graphemic information in the sensory register. These codes serve as means to access the appropriate node in the lexicon. Internal codes are alternative representations of the stimulus word that can be used to locate the item if it is stored elsewhere in memory. The internal codes can be of various types; they may be abstract pictorial or auditory images, a list of semantic-syntactic markers, predicate relations, etc. Information recorded in memory involves an array of internal codes, and the same object or event may be represented by different codes depending on the memory store involved and related information. Finally, output codes, when entered into the response generator, permit the subject to produce the word in various forms (oral, written, etc.). The property of lexical nodes that allows transformation from one code to

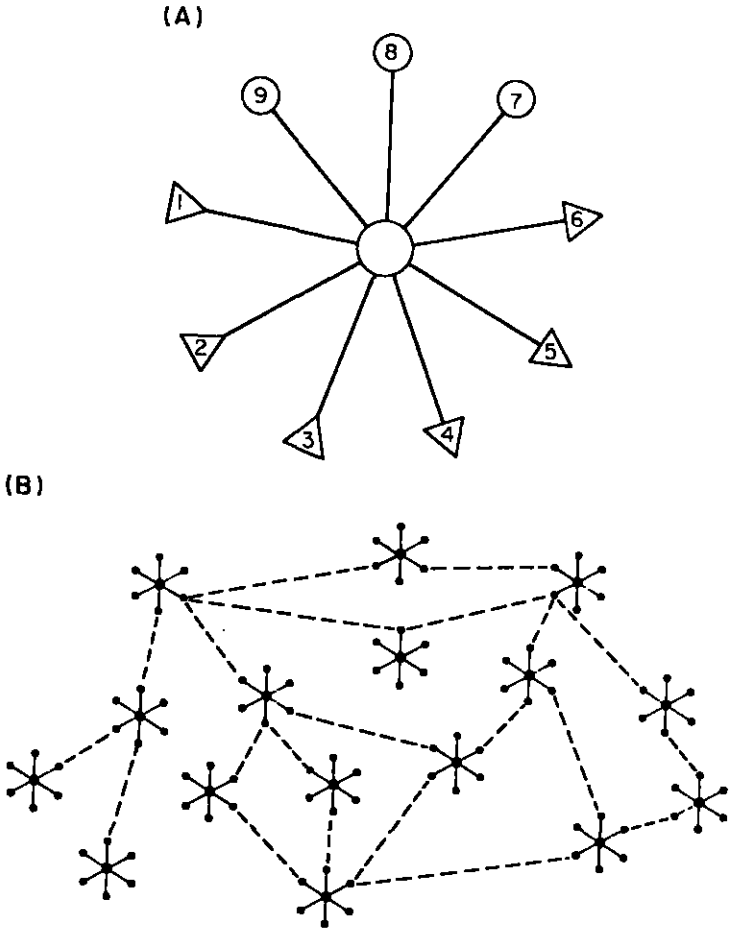


FIGURE 2.
A schematic representation of the lexical store. Panel (A) illustrates a hypothetical node in the lexicon with associated input codes [(1) auditory, (2) pictorial, (3) graphemic], output codes [(4) written, (5) spoken, (6) imaged], and internal codes [(7) acoustical code for STS, (8) imaginal code for LTS, (9) verbal code for LTS]. Panel (B) illustrates a subset of nodes in the lexicon, with dashed lines indicating codes that are shared by more than one lexical node. For example, depending on an individual's experience, the nodes for mare and stallion could share a common internal code; if this code is used (along with others) to represent a particular episode, then information about the horse's sex will not be recorded in memory.

another has proved useful in other theories of memory, most notably in the logogen system of Morton (1969, 1970).

It is possible that information stored at the node representing the test word could lead directly to the decision to make an A_1 or A_0 response. This would be the case if, for example, each node corresponding to a target word has associated with it a marker or list tag that could be retrieved when the item is tested (Anderson & Bower, 1972). We take the alternative view, however, that information contained in the lexical store is relatively isolated from those parts of the memory system that record the occurrence of particular events, experiences, and thought processes. The lexical store contains the set of symbols used in the information-handling process, and the various codes associated with each symbol; these codes are the language in which experiences are recorded, but the actual record is elsewhere in memory. Thus, memorizing a list of words involves extracting appropriate codes from the lexicon and organizing these codes into an array to be recorded in a partition of LTS separate from the lexical store. There is no direct link between a word's node in the lexicon and its representation in the memory structure for the word list; to establish that a word is a member of the memorized list involves extracting an appropriate code from the word's lexical node and scanning it against the list for a possible match.

Thus, LTS is viewed as being partitioned into a lexical store and what we call the event-knowledge store (E/K store). As noted above, the lexical store maintains a set of symbols and codes that can be used by the subject to represent knowledge and the occurrence of particular events. When the subject is confronted with new information, he represents it in the form of an array of internal codes, and if it is to be retained on a long-term basis, that array is recorded in the E/K store.[2] Our representation of words resembles the model proposed by Kintsch (1970b), but differs from his model regarding the representation of a memorized list. Kintsch assumes that acquisition of a list involves increasing the familiarity or strength of an item in the lexical store. Although we agree with Kintsch up to this point, we also propose that the code or codes of a word in the lexical store are copied and placed in the E/K store. The organization of these codes in the E/K store, as suggested by Herrmann (1972), will depend on the particular study procedure used in acquisition (e.g., serial order, an arbitrary pairing of words, or clustering by a common meaning such as category membership). The division of LTS into

[2] In order to simplify the presentation, a sharp distinction has been made between the lexical store and the E/K store. The distinction is satisfactory for the experiments treated in this chapter. However, in general, we view LTS as a graded set of memories; those described here as lexical nodes represent one extreme, while event memories represent the opposite extreme. The lexical store evolves over a lifetime; by analysis of past memories the individual develops new codes that make the storage of future events more economical. Thus one's history of experiences determines the codes available in the lexical system and, in turn, the ability to store different types of information (Atkinson & Wescourt, 1974).

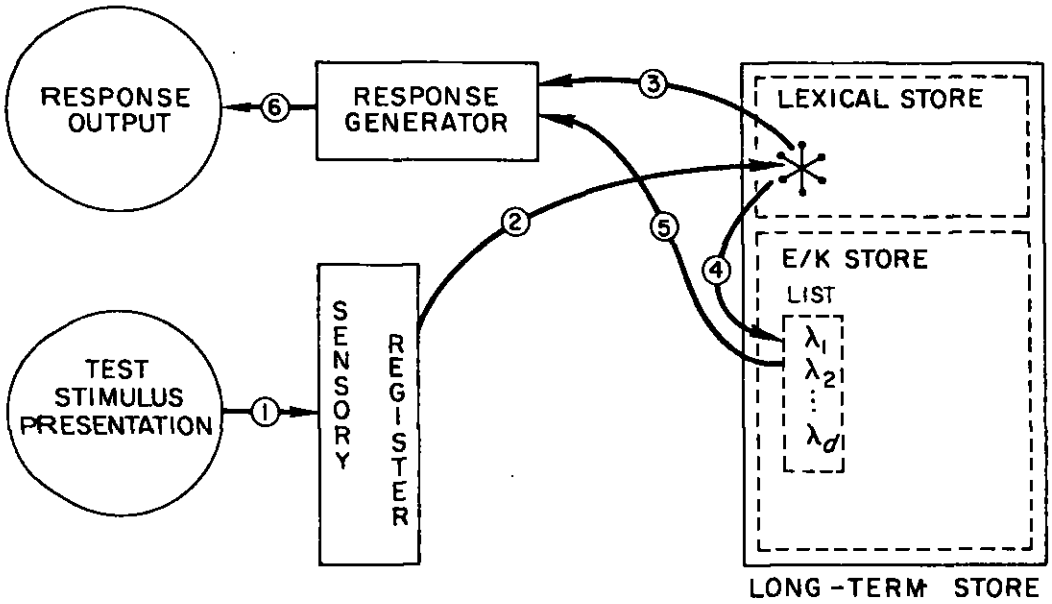


FIGURE 3.
A schematic representation of the search and decision processes in long-term recognition memory. A test stimulus is presented (1) and then encoded and matched to the node in the lexicon (2). The familiarity index associated with the node may lead to an immediate decision (3) and in turn generate a response (6). Otherwise an extended search of the stored target list is initiated (4), which eventually leads to a decision (5) and a subsequent response (6). Path (1), (2), (3), (6) results in a faster response than path (1), (2), (4), (5), (6), and the response that is independent of target-set size.

a lexical store and an E/K store is similar to the distinction made by Tulving (1972) between semantic and episodic memory. In Tulving's taxonomy, the lexical store would be classified as semantic memory. The E/K store, however, might be classified by Tulving as either semantic memory or episodic memory, depending on the type of information in the E/K store. To Tulving, one's memory for a list learned in a psychology experiment constitutes an episodic memory, but the knowledge one learns in a chemistry course (such as the periodic table of elements) constitutes a semantic memory. It is maintained here that both kinds of information are held in the E/K store and are treated by the memory system in essentially the same manner (Atkinson & Wescourt, 1974).

Figure 3 presents a summary of the processes involved in recognition memory for words that are members of a list stored in long-term memory. When the test word is presented, it is encoded into an input code that allows direct access to the appropriate node in the lexical store. Although the node does not contain a tag or marker indicating list membership, it will be as-

sumed that by accessing the node the subject can arrive at an index of the test word's *familiarity*. The familiarity value for any node is a function of the time since that node was last accessed relative to the number of times the node had been accessed in the past. Infrequently occurring words receive a large increase in familiarity after a single test, whereas the test of a frequent word results in only a small increase in its familiarity. The familiarity value for any word is assumed to regress to its base value as a function of time since the last access of the node.[3]

In recognition experiments of the type described above, the familiarity value of a word sometimes can be a fairly reliable indicator of list membership. It will be assumed that, when the subject finds a very high familiarity value at the lexical node of the test word, he outputs an immediate A_1 response; if he finds a very low familiarity value, he outputs an immediate A_0. If the familiarity value is intermediate (neither low nor high), the subject extracts an appropriate code for the test word and scans it against the target list in the E/K store. If the scan yields a match, an A_1 is made; otherwise A_0. The recognition process sketched out above is similar to that proposed by Mandler, Pearlstone, and Koopmans (1969). In the next section, these ideas are quantified and tested against data involving both error probabilities and response latencies.

A MODEL FOR RECOGNITION

Several special cases of the model to be considered here have been presented elsewhere (Atkinson & Juola, 1973; Juola, Fischler, Wood, & Atkinson, 1971; Atkinson, Herrmann, & Wescourt, 1974). These papers may be consulted for further intuitions about the model, as well as for applications to a variety of experimental tasks.

It is assumed that each node in the lexicon has associated with it a familiarity measure that can be regarded as a value on a continuous scale. The familiarity values for target items 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 & Teghtsoonian, 1961), the target set is not well learned and involves stimuli that have received only a single study presentation. Under these conditions the familiarity value of the test

[3] Familiarity as used here is not specific to particular events. It can be viewed as a reverberatory activity that dissipates over time. Whenever a node is accessed, it is set in motion. The amount of reverberation and its time course depend on the prior reverberation of the node and the reverberatory activity at neighboring nodes (Schvaneveldt & Meyer, 1973). When a node is accessed, the system can gauge the current reverberatory level of that node and use the measure as an item of information.

stimulus leads directly to the decision to make an A_1 or A_0 response; that is, the subject has a single criterion along the familiarity continuum that serves as a decision point for making a response. Familiarity values above the criterion lead to an A_1 response, whereas those below the criterion lead to an A_0 response (Banks, 1970; Kintsch, 1967, 1970a, 1970b; Parks, 1966; Shepard, 1967).

The studies that we consider differ from most recognition experiments in that the target stimuli are members of a well-memorized list. In this case, it is assumed that the subject can use the familiarity value to make an A_1 or A_0 response as soon as the appropriate lexical node is accessed, or can delay the response until a search of the E/K store has confirmed the presence or absence of the test item in the target set. These processes are shown in the flowchart of Figure 4. When a test stimulus is presented, the subject accesses the appropriate lexical node and obtains a familiarity value. This value is then used in the decision either to output an immediate A_1 or A_0 response (if the familiarity is very high or very low, respectively) or to execute a search of the E/K store before responding (if it is of an intermediate value).

A schematic representation of the decision process is shown in Figure 5. Here the distributions of familiarity values associated with a distractor item and a target item are plotted along the familiarity continuum (x). If the initial familiarity value is 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 value is between c_0 and c_1, the subject searches the E/K store before responding; this search guarantees that the subject will make a correct response, but it takes time in proportion to the length of the target list.

On the nth presentation of a given item in a test sequence, there is a density function reflecting the probability that the item will generate a particular familiarity value x; the density function will be denoted $\phi_{1,n}(x)$ 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 subscript 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 Figure 6 where $\mu_{1,n}$ and $\mu_{0,n}$ shown in the bottom panel ($n > 1$) have both been shifted to the right of their initial values $\mu_{1,1}$ and $\mu_{0,1}$ shown in the top panel. The effect of shifting up the mean familiarity values is to change the probability that the presentation of an item will result in a search of the E/K store.

We can now write equations for the probabilities that the subject will make a correct response to target and distractor items. As shown in Figure 5, it is assumed that the subject will make an error if the familiarity value of a target word is below c_0, or if the familiarity of a distractor is above c_1. In all

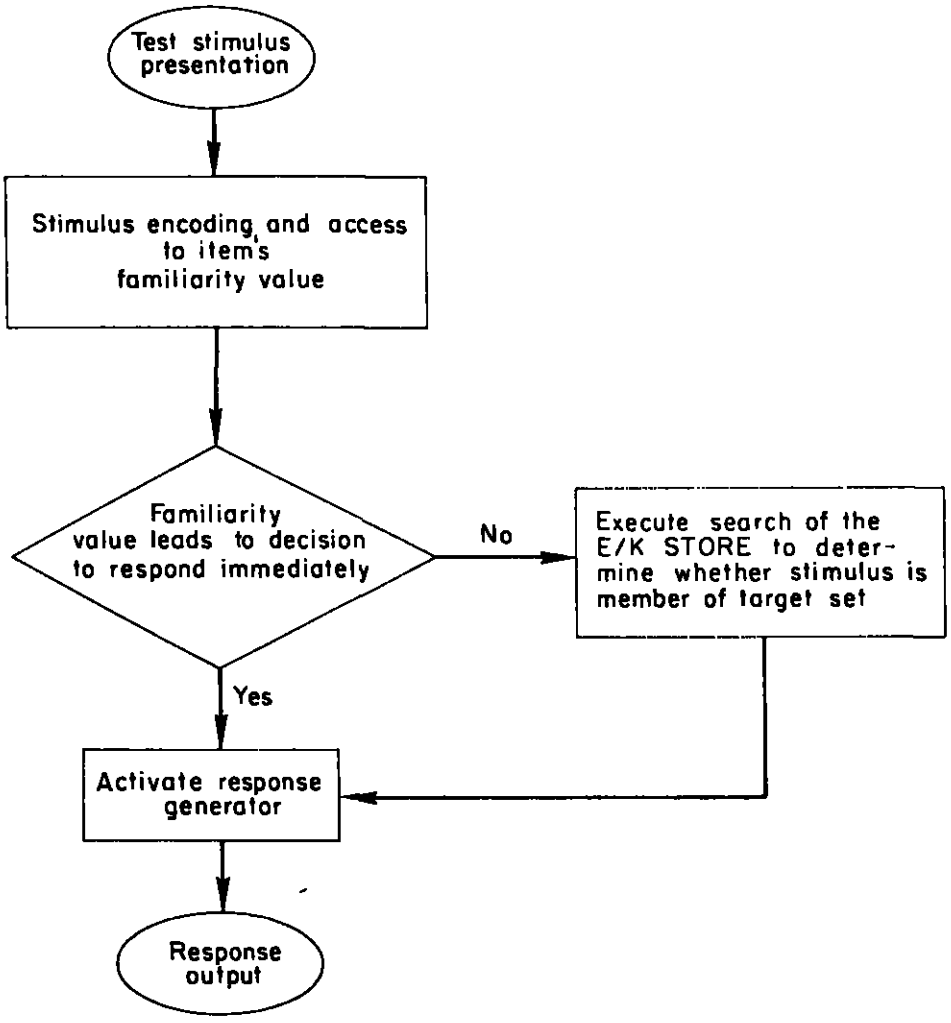


FIGURE 4.
Flowchart representing the memory and decision stages involved in recognition.

other cases, the subject will make a correct response. Thus the probability of a correct response to a target word presented for the nth time is the integral of $\phi_{1,n}(x)$ from c_0 to ∞:

$$P(A_1 \mid S_{1,n}) = \int_{c_0}^{\infty} \phi_{1,n}(x)\, dx = 1 - \Phi_{1,n}(c_0). \tag{1}$$

Similarly, the probability of a correct response to a distractor presented for the nth time is the integral of $\phi_{0,n}(x)$ from $-\infty$ to c_1:

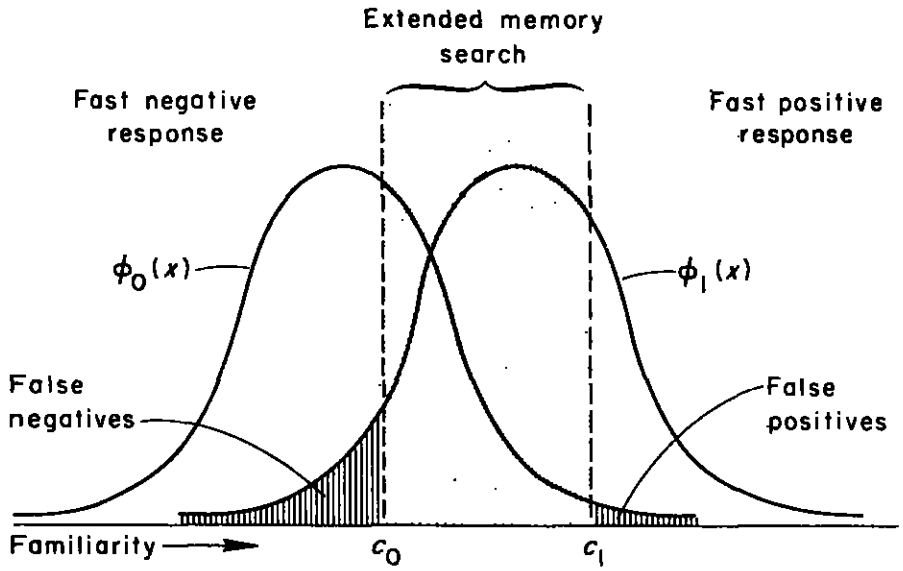


FIGURE 5.
Distributions of familiarity values for distractor items, $\phi_0(x)$, and target items, $\phi_1(x)$.

$$P(A_0 \mid S_{0,n}) = \int_{-\infty}^{c_1} \phi_{0,n}(x)\, dx = \Phi_{0,n}(c_1). \tag{2}$$

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

In deriving response latencies, we assume that the processes involved in encoding the test stimulus, retrieving information about the stimulus from memory, making a decision about which response to choose, and emitting a response can be represented as successive and independent stages. These stages are diagrammed in the flowchart in Figure 7. When the test stimulus is presented, the first stages involve encoding the item, accessing the appropriate node in the lexical store, and retrieving a familiarity value x. The times required to execute these stages are combined and represented by the quantity ℓ in Figure 7. The next stage is to arrive at a recognition decision on the basis of x; the decision time depends on the value of x relative to c_0 and c_1, and is given by the function $r(x)$. If $x < c_0$, a negative decision is made; if $x > c_1$, a positive decision is made. If $c_0 \leq x \leq c_1$, a search of the E/K store is required. The time for this search is assumed to be a function of d, the size of the target set; namely, $\kappa + \theta_i(d)$. In this equation, κ denotes the time to extract an appropriate search code from the lexical node and initiate the scan of the target list; $\theta_i(d)$ is the time to execute the scan and depends upon d and upon whether the test item is a target ($i = 1$) or a distractor ($i = 0$).

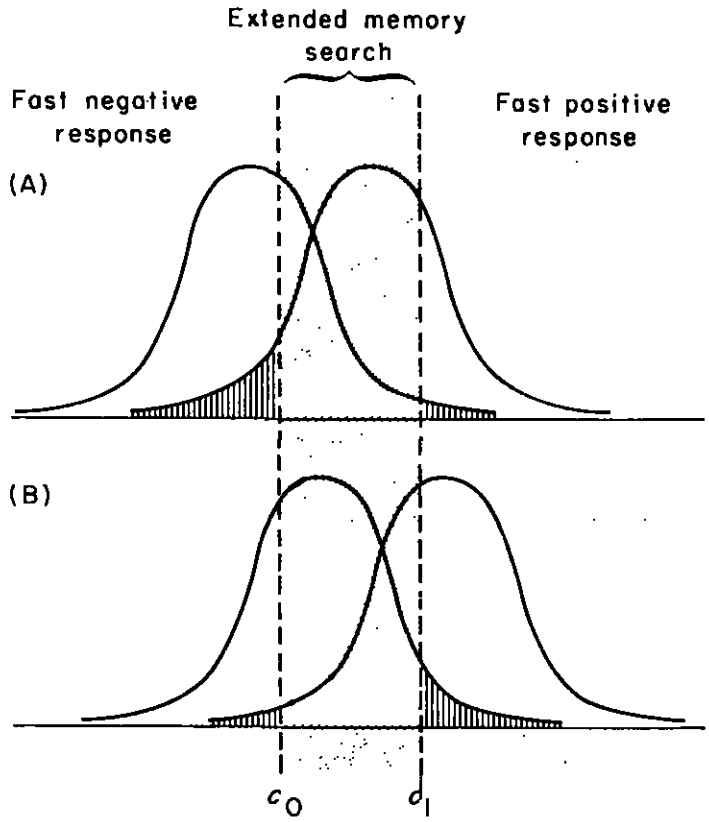


FIGURE 6.
Distributions of familiarity values for distractor items and target items that have not been tested (Panel A), and that have had at least one prior test (Panel B).

The final stage is to output a response once the decision has been made, the response time being r_0 for an A_0 response and r_1 for an A_1 response.[4] The quantities ℓ, $r(x)$, κ, $\theta_i(d)$, and r_i are expected values for the times necessary to execute each stage. If assumptions are made about the forms of the distributions associated with these expected values, then expressions for all moments of the latency data can be derived. Their derivation is complicated under some conditions of the model, but under others it simply involves a probabilistic mixture of two distributions; that is, the times resulting from

[4] The successive and independent stages of the process, as represented by the blocks in Figure 7, should be regarded as an approximation to the true state of affairs (Egeth, Marcus, & Bevan, 1972). Psychological and physiological considerations make it doubtful that the phenomena considered here are composed of truly independent stages, but stage models tend to be mathematically tractable, and thus are useful analytic tools.

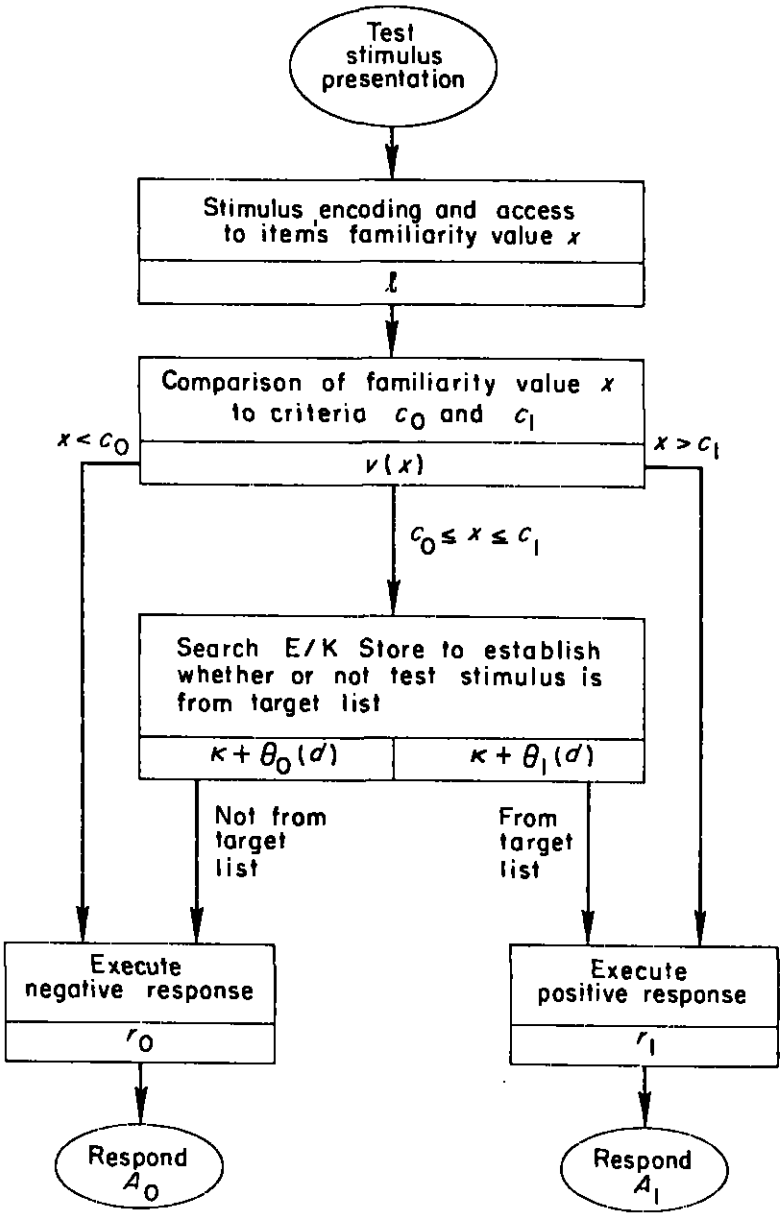


FIGURE 7.
Flowchart representing memory search and decision stages of recognition. The bottom entry in each box represents the time required to complete that stage.

fast responses based on the familiarity value alone and times resulting from slow responses based on the outcome of the extended memory search. In this chapter, however, we only make assumptions about the expected value for each stage, thereby restricting the analysis to mean response data.

We let $t(A_i \mid S_{j,n})$ denote the expected time for an A_i response to the nth presentation of a particular stimulus drawn from set S_j $(i, j = 0, 1)$. Expressions can be derived from these quantities by weighting the times associated with each stage by the probability that the stage occurs during processing. Thus, for example, the time to make an A_1 response to the nth presentation of a given target item (S_1) is simply the time required to execute a response based on the familiarity value alone plus the time to execute a response based on a search of the E/K store, each weighted by their respective probabilities. If x is the familiarity value, then the time for a fast A_1 response is $\ell + v(x) + r_1$; if, however, a search of the E/K store is made, then response time is $\ell + v(x) + \kappa + \theta_1(d) + r_1$. The weighting probabilities must take account of the fact that we are concerned with the time for an A_1 response, conditional on its being correct. The probability of a fast A_1 response, conditional on the fact that it is correct, is the integral of $\phi_{1,n}(x)$ from c_1 to ∞, divided by the probability of a correct A_1 response (the integral of $\phi_{1,n}(x)$ from c_0 to ∞). Similarly, the probability of a slow A_1 response, conditional on the fact that it is correct, is the integral $\phi_{1,n}(x)$ from c_0 to c_1, divided by the integral of $\phi_{1,n}(x)$ from c_0 to ∞. Thus the expected time for an A_1 response to the nth presentation of a particular target item is

$$\left[\int_{c_1}^{\infty} [\ell + v(x) + r_1]\phi_{1,n}(x)\,dx\right]\left[\int_{c_0}^{\infty} \phi_{1,n}(x)\,dx\right]^{-1} + \left[\int_{c_0}^{c_1} [\ell + v(x) + \kappa + \theta_1(d) + r_1]\phi_{1,n}(x)\,dx\right]\left[\int_{c_0}^{\infty} \phi_{1,n}(x)\,dx\right]^{-1}.$$

Note that ℓ and r_1 may be removed from under the integral. Doing this and rearranging terms yields

$$t(A_1 \mid S_{1,n}) = \ell + r_1 + \left[\int_{c_1}^{\infty} v(x)\phi_{1,n}(x)\,dx + \int_{c_0}^{c_1} [\kappa + \theta_1(d) + v(x)]\phi_{1,n}(x)\,dx\right]\left[1 - \Phi_{1,n}(c_0)\right]^{-1}; \quad (3)$$

where again $\Phi(\cdot)$ denotes the distribution function associated with the density function $\phi(x)$. Similarly,

$$t(A_0 \mid S_{0,n}) = \ell + r_0 + \left[\int_{-\infty}^{c_0} v(x)\phi_{0,n}(x)\,dx + \int_{c_0}^{c_1} [\kappa + \theta_0(d) + v(x)]\phi_{0,n}(x)\,dx\right]\left[\Phi_{0,n}(c_1)\right]^{-1}; \quad (4)$$

$$t(A_0 \mid S_{1,n}) = \ell + r_0 + \left[\int_{-\infty}^{c_0} v(x)\phi_{1,n}(x)\,dx\right]\left[\Phi_{1,n}(c_0)\right]^{-1}; \quad (5)$$

$$t(A_1 \mid S_{0,n}) = \ell + r_1 + \left[\int_{c_1}^{\infty} v(x)\phi_{0,n}(x)\,dx\right]\left[1 - \Phi_{0,n}(c_1)\right]^{-1} \tag{6}$$

Equations 3 and 4 are the expected times for correct responses, and Equations 5 and 6 are expected times for incorrect responses, to target and distractor items, respectively.

In fitting the model to data, we assume that $\phi_{i,n}(x)$ is normally distributed with unit variance for all values of i and n. Thus, the presentation of an item causes the distribution to be shifted up without changing its form or variance.[5] No assumptions are made about how $\mu_{i,n}$ changes with n. Several assumptions seem reasonable on an a priori basis; rather than select among them, we bypass the issue by estimating $\mu_{i,n}$ from the data for each value of n. This approach is practical because the range on n is small for the experiments considered here.

It should be remarked that the criteria c_0 and c_1 are set by the subject. In the initial stages of an experiment, they would vary as the subject adjusted to the task, but it is assumed that in time they would stabilize at fixed values. Again, no theory is given of how c_0 and c_1 vary over initial trials, and thus data for the early stages of an experiment are not treated.

Yet another simplifying assumption should be mentioned at this point. Equations 1 and 2 indicate that errors are determined by the values of $\mu_{i,n}$, c_0, and c_1. In the experiments examined in this chapter, there is no evidence to suggest that error rates vary as a function of d, the size of the target list. Thus, in treating data we assume that $\mu_{i,n}$, c_0, and c_1 are independent of d. Experimental procedures can be devised where this assumption would be violated (see Atkinson & Juola, 1973), but for the studies discussed here it is warranted.

What remains to be specified are the functions $r(x)$ and $\theta_i(d)$. It is assumed that $r(x)$ takes the following form:[6]

$$r(x) = \begin{cases} \rho e^{-(x-c_1)\beta}, & \text{for } x > c_1, \\ \rho, & \text{for } c_0 \leq x \leq c_1, \\ \rho e^{-(c_0-x)\beta}, & \text{for } x < c_0. \end{cases} \tag{7}$$

Figure 8 presents a graph of the equation. If the familiarity value x is far above the upper criterion or far below the lower criterion, the decision time approaches zero; for values close to the criteria, the decision time approaches ρ. A special case of interest is when $\beta = 0$; namely,

$$r(x) = \rho. \tag{8}$$

[5] The assumption that only the mean and not the form of the distribution changes is made primarily to simplify the mathematics. Other assumptions, such as those considered by Suppes (1960) for a different problem, seem equally plausible and should be investigated in formulating a more general model of familiarity change.

[6] The $r(x)$ function proposed here is similar to one investigated by Thomas (1971) for a signal-detection task.

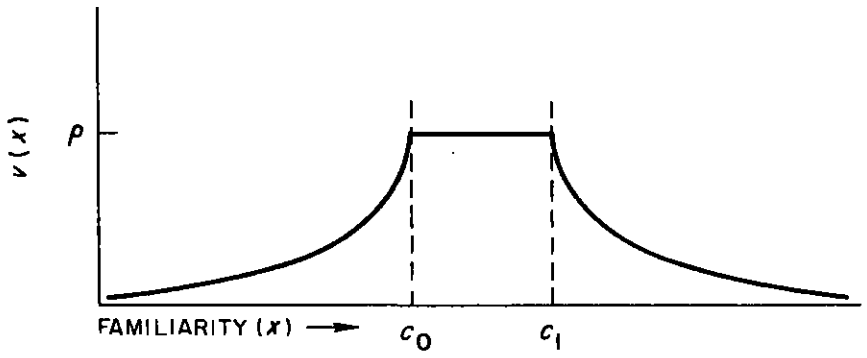


FIGURE 8.
The function $v(x)$.

In this case, the time to evaluate the familiarity value is constant regardless of its relation to c_1 and c_0.

The quantity $\theta_i(d)$ represents the time to search the E/K store, and is assumed to be a linear function of target-set size. For the most general case we assume that search times on positive and negative trials vary independently; that is,

$$\theta_1(d) = \alpha d, \tag{9a}$$

$$\theta_0(d) = \alpha' d. \tag{9b}$$

As a special case of Equation 9, it is possible that the search times are identical for target and distractor items:

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

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 the target items are stored as a list structure, and portions of the list are retrieved and scanned as the subject seeks 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

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

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

The memory-search processes described by Equations 10 and 11 correspond to the exhaustive and self-terminating cases of the serial scanning model proposed by Sternberg (1966, 1969b). While Sternberg's models have proved to be extremely valuable in interpreting data from a variety of memory-search experiments, good fits between the models and data do not 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 Equations 10 and 11 (Atkinson, Holmgren, & Juola, 1969; Murdock, 1971; Townsend, 1971; Shevell & Atkinson, 1974). Thus, the use of the above equations to specify the time to search the E/K store does not commit us to either a serial or parallel interpretation.

EFFECTS OF TARGET-LIST LENGTH AND TEST REPETITIONS

The first experiment we consider was designed primarily to replicate two earlier experiments, as well as to provide a large data base with which to test the model. Juola et al. (1971) demonstrated that recognition time is a straight-line function of the number of items in a large (10 to 26 items) target set: as the number of items in the target set increased, response latency increased linearly for both positive and negative trials. A second experiment (Fischler & Juola, 1971) showed that response latency depends on whether or not the test stimulus has been presented previously. The response latency for a repeated target item was more than 100 msec less than the latency for a target on its first presentation. For a distractor, repetitions increased latency, with reponse time being about 50 msec greater for a repeated distractor than for one receiving its first presentation.

Our study also included repeated tests of target and distractor items, and three target-list lengths were used. Groups of 24 subjects each were given lists of either 16, 24, or 32 words. Each list was constructed by randomly selecting d words from a pool of 48 common, one-syllable nouns. The words remaining in the pool after each list had been selected were used as the distractor set (S_0) to accompany that target set (S_1). Each subject was given a list about 24 hours before the experimental session, and instructed to memorize it in serial order.

At the start of the test session, each subject successfully completed a written serial recall of the target list. 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 A_0 response by depressing one of two telegraph keys with his right forefinger. The experimental procedure was identical to that of Fischler and Juola (1971).

The test sequence consisted of 80 consecutive trials that were divided into four blocks. For Block I, four target words and four distractors were randomly selected from S_1 and S_0, respectively. For Block II, the eight Block I words were repeated, and four new targets and four new distractors were also shown. Block III included all the words presented in Block II with eight new words added (four targets and four distractors). Finally, Block IV

included all the words of Block III and eight new words (the remaining unused target and distractor items). Order of presentation within blocks was randomized.

With this method of presentation, 16 target words and 16 distractors were presented to each subject. The test words thus included all of S_1 for subjects with lists of 16 words. For the other groups, the 16 target words tested were either the first or last 16 words in the 24-word lists, or they were the first, middle, or last 16 in the 32-word lists. It should be pointed out that the specific part of the target list that was tested during the experimental session had no effect on response times or error rates. Thus, no further distinction on which part of the target list was tested is made between groups of subjects. The lack of any effects due to the list part that was tested is not surprising when it is noted that in several previous experiments (Atkinson & Juola, 1973; Fischler & Juola, 1971; Juola et al., 1971) no effects were observed due to the serial position of the target word, that is, positive response latencies plotted against the target words' serial position yielded a flat function. The overall effect of list length on latency is also uninfluenced by the testing scheme used; the magnitude of the list-length effect observed in this study is the same as in studies where all items of each list are tested (Juola et al., 1971). The procedure used here has the nice feature that the test sequence is the same for all groups, the only difference among groups being the length of the list memorized prior to the test session. The subjects who memorized the longer lists were not told that only part of the list would be used, and in the debriefing session at the end of the experiment no one commented on the fact that some items were not tested.

The mean latencies for correct responses are presented in Figure 9; the data are from the last two trial blocks only (Blocks III and IV). The effects shown in Figure 9 were also obtained in Trial Blocks I and II; however, response times were somewhat greater on these trials, presumably due to practice effects. The data from Blocks III and IV were very similar and will be regarded as representing asymptotic performance. In another paper (Atkinson & Juola, 1973), we used the model to make predictions about all the data, including practice effects, for a similar experiment, but here we are concerned only with the data presented in Figure 9. As shown in Figure 9, means were obtained separately for A_1 and A_0 responses to test words that were presented for the first, second, third, and fourth times ($n = 1, 2, 3$, or 4). Because, within blocks, the presentation number was randomly ordered, the effects shown in Figure 9 are attributable only to the prior number of times the test word had been presented. In general, the results closely replicate the findings of earlier studies. By comparing the mean latencies as the presentation number increases from one to four in Figure 9, it can be seen that the targets and distractors yield opposite effects. Repetitions decrease response latencies for targets, and increase latencies for distractors. The line

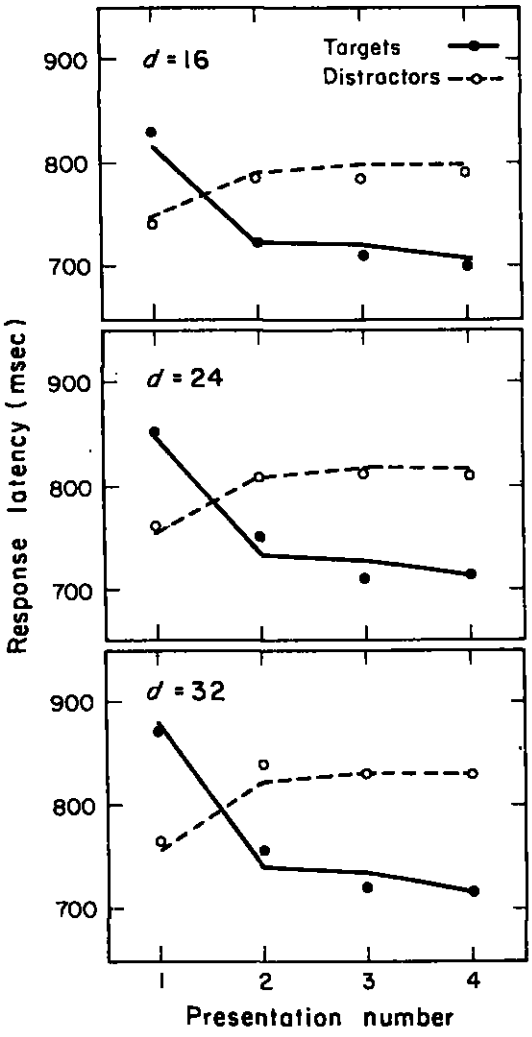


FIGURE 9.
Correct response latencies as functions of presentation number for target and distractors for three list-length (d) conditions: the top panel presents data for $d = 16$, the middle panel for $d = 24$, and the bottom panel for $d = 32$. The broken lines fitted to the data represent theoretical predictions.

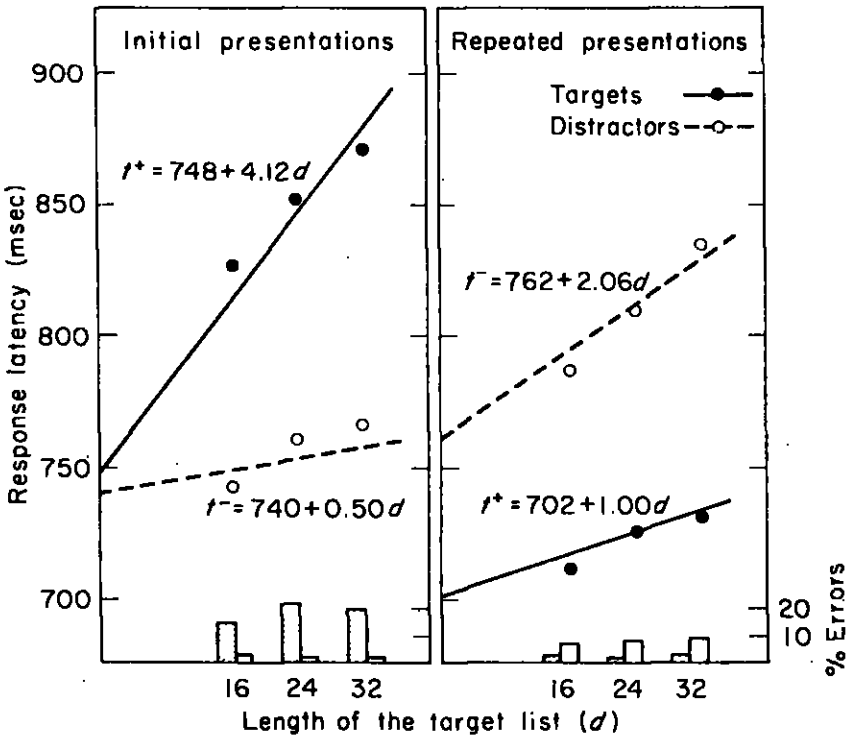


FIGURE 10.
Correct response latencies and error percentages as functions of target-list length; the data represent a weighted average of response latencies from Trial Blocks III and IV. The left panel presents data for initial presentations of target and distractor words, and the right panel presents the data for repeated presentations. Incorrect responses to target words are indicated by the shaded bars, and errors to distractors by the open bars. The straight lines fitted to the data represent theoretical predictions.

segments fitted to the data were generated from the model and are discussed later.

The data from Figure 9 are replotted in Figure 10 so that mean response latencies are presented as functions of target-list length. The left panel includes the data for items receiving their first presentations ($n = 1$), whereas the right panel presents the average data for repeated presentations ($n = 2$, 3, and 4) weighted by the number of observations for each value of n. Again the effects of repetitions are evident; repetitions decrease latency on positive trials by more than 100 msec, whereas repetitions increase negative latencies by about 50 msec, on the average. Similarly, repeated tests decreased errors to target words (shaded bars along the lower axis), and repetitions increased errors to distractors (open bars). The linear functions fitted to the data in Figure 10 are discussed later.

The number of target words affected response latency, with mean latency

being an approximately linear function of target size. By way of contrast, note that error rates do not increase with the number of target words, but are relatively constant across the three list lengths. Further, an examination of error latencies showed that there was no effect of list length on the speed of an incorrect response.

Perhaps most interesting, however, is the interaction between target-set size and the effects of repetitions. For target words, repetitions decrease the size of the list-length effect; that is, the slope of the function relating mean response latency to target-list length is less for repeated targets than for initially presented targets. The opposite is true for distractors; repeating distractors increases the slope of the latency function.

A discussion of these results is postponed until the end of the next section. We first demonstrate how parameters can be estimated and the model fitted to data.

THEORETICAL ANALYSIS OF THE LIST-LENGTH EXPERIMENT

There are several approaches that can be taken to estimate parameters. The method used here is not the most efficient, but it has the merit of being quite simple. It involves using the error probabilities to estimate the $\mu_{i,n}$s. The estimates of the $\mu_{i,n}$s are then substituted into the latency equations and treated as fixed values. The remaining parameters are estimated by selecting them so that the differences between observed and predicted latencies are minimized.[7]

Table 1 presents observed error probabilities for target and distractor items. These probabilities were obtained by averaging over the three list-length conditions, because there were no significant differences in error rates across

TABLE 1
Observed error probabilities for targets and distractors

	$P(A_0 \mid S_{1,n})$	$P(A_1 \mid S_{0,n})$
$n = 1$	0.171	0.005
$n = 2$	0.016	0.039
$n = 3$	0.014	0.049
$n = 4$	0.007	0.049

[7] There are methods that permit simultaneous estimates of all parameters, but practical limitations make them unfeasible except in special cases (see Atkinson & Juola, 1973).

groups. We use these data and Equations 1 and 2 to estimate the $\mu_{i,n}$s. For example, $P(A_0 \mid S_{1,1}) = \Phi_{1,1}(c_0)$ from Equation 1, and the observed value for this probability is 0.171 from Table 1. Consulting a normal probability table, $\mu_{1,1} = c_0 + 0.95$ in order for the error rate to be 0.171. Similarly $\mu_{1,2} = c_0 + 2.14$, $\mu_{1,3} = c_0 + 2.20$, and $\mu_{1,4} = c_0 + 2.46$, using the remaining error data in the first column of Table 1. Proceeding in the same way, using Equation 2 and the error data in the second column of the table, we obtain $\mu_{0,1} = c_1 - 2.58$, $\mu_{0,2} = c_1 - 1.76$, $\mu_{0,3} = c_1 - 1.66$, and $\mu_{0,4} = c_1 - 1.66$. Thus the observed error probabilities fix the estimates of $\mu_{1,n}$ in terms of c_0, whereas $\mu_{0,n}$ is in terms of c_1. It can be shown that the theoretical predictions for error probabilities and latencies do not depend on the absolute values of c_0 and c_1, but only on their difference. Thus, one or the other can be set at an arbitrary value. For simplicity, we let $c_0 = 0$; note that no matter what value is selected for c_1, the error data will be fit perfectly. By setting c_0 equal to zero and by assuming unit variance for the ϕ-distributions, we have in essence defined the zero point and measurement unit for the familiarity scale.

With $c_0 = 0$ and the $\mu_{i,n}$s restricted by the error data, the remaining parameters can be estimated from the latency data. Six special cases of the general model are used to fit the latency data. As indicated in Table 2, the cases differ

TABLE 2
Six models defined in terms of the functions $v(x)$ and $\theta_i(d)$

$v(x)$ \ $\theta_i(d)$	Equation 9	Equation 10	Equation 11
Equation 8	Model I c_1 $(\ell + \rho + r_1)$ r κ α α'	Model II c_1 $(\ell + \rho + r_1)$ r κ α	Model III c_1 $(\ell + \rho + r_1)$ r κ α
Equation 7	Model IV c_1 $(\ell + r_1)$ r κ ρ β α α'	Model V c_1 $(\ell + r_1)$ r κ ρ β α	Model VI c_1 $(\ell + r_1)$ r κ ρ β α

Note: $r = r_0 - r_1$.

in how the functions $r(x)$ and $\theta_i(d)$ are defined. Equations 7 and 8 define two versions of $r(x)$, and Equations 9, 10, and 11 define three versions of $\theta_i(d)$. Listed in Table 2 are the parameters that must be estimated for each case; the parameter r is simply the difference between r_0 and r_1. The parameters grouped in parentheses cannot be individually identified---that is, the predictions of the model depend only on the sum of these parameters, which means that they cannot be estimated separately.[8] Note that the pair of models in each column of Table 2 are equivalent if $\beta = 0$; thus the lower model in a column must predict the data better than the one above it unless β is estimated to be zero. Similarly, Model I reduces to Model II and Model IV to V if $\alpha = \alpha'$; Model I must be better than II and Model IV better than V unless the estimates of α and α' are identical.

Our method of parameter estimation involves the 24 data points in Figure 9. Parameter estimates are selected that minimize the sum of the squared deviations (weighted by the number of observations) between the data points and theoretical predictions. Specifically, we define the root-mean square deviation (RMSD) between observed and predicted values as follows:

$$\text{RMSD} = \left[(1/N)\sum_{i=1}^{24} n_i(t_{p,i} - t_{o,i})^2\right]^{1/2}, \tag{12}$$

where N = the total number of observations; i = an index over the 24 data points shown in Figure 9; n_i = the number of observations determining data point i; $t_{p,i}$ = predicted response latency for data point i; and $t_{o,i}$ = observed response latency for data point i.

For each of the six models, the function defined in Equation 12 is to be minimized with respect to the parameter set given in Table 2. We have not attempted to carry out the minimization analytically, for it appears to be an impossible task; rather a computer was programmed to conduct a systematic search of the parameter space for each model until a minimum was obtained.[9] The minimum RMSDs obtained are shown in Table 3, along with the number of parameters estimated in the computer search for each model. Models III and VI clearly yield the poorest fit and can be eliminated from contention. The fact that Models I and II are about equally good-- as are Models IV and V---indicates that separate estimates of α and α' do not substantially improve the goodness of fit. The conclusion to be drawn from this observation is that the time to search the E/K store is approximately the same for both targets and distractors. Note also that Models I and IV are about equally good, as are Models II and V, suggesting that the more complicated $r(x)$ functions

[8] Proof of this remark is straightforward and is not given here. Note that for Models I, II, and III the parameter ρ is not identifiable but is lumped in the quantity $(t + \rho + r_1)$, whereas for Models IV, V, and VI ρ is identifiable and only $(t + r_1)$ is lumped.

[9] For a discussion of such search procedures, see Wilde (1964).

TABLE 3
Minimum RMSDs obtained in computer search

Model	Minimum RMSD	Number of parameters estimated
I	9.93	6
II	9.94	5
III	10.89	5
IV	9.86	8
V	9.92	7
VI	10.34	7

yield little improvement over the constant function. Add to these observations the fact that Model II with only five parameters produces virtually as good a fit as does Model IV with eight parameters.

In view of the preceding considerations, Model II is our preferred choice among the six models. Table 4 presents the parameter estimates for Model II;

TABLE 4
Parameter estimates for Model II

Parameter	Estimate
c_1	1.02
$(\ell + \rho + r_1)$	687 msec
r	44 msec
κ	137 msec
α	9.9 msec

Note: $r = r_0 - r_1$.

the predicted response times from this model are shown in Figure 9 as connected lines. The straight lines shown in Figure 10 are the predicted functions based on Model II for initial presentations (left panel) and repeated presentations (right panel). The fits displayed in Figure 10 could be improved upon somewhat, but it should be kept in mind that they were obtained by using parameter estimates based on a different breakdown of the data (i.e., that shown in Fig. 9).[10]

[10] Similarity factors not represented in the model could contribute to the list-length effects displayed in Figures 9 and 10. As the target set increases in size, the probability that any given distractor will be similar to a target item also increases. Visual (or graphemic) similarity could affect the speed with which the appropriate lexical node is accessed, leading to

The latency of an error response should be fast according to the theory, because errors can occur only if the subject responds before the extended memory search is made. The data support this prediction, and they accord well with the values predicted by Model II. Specifically, the latency of an error is close to the predicted value of $\ell + \rho + r_0 = 731$ msec for an S_1 item, and close to $\ell + \rho + r_1 = 687$ msec for an S_0 item. Furthermore, as predicted by the model, the observed error latencies do not appear to be influenced by the length of the target list.

A verbal interpretation of the results in terms of Model II reads as follows. When a target item is presented for the first time, the probability that a search of the E/K store will occur before a response is made exceeds the probability that a fast positive response 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, and the list-length effect is greater for targets than for distractors (because list-length effects depend only upon the search of the E/K store). The effect of repeated tests of words is to increase the familiarities of both targets and distractors. This results in an increased latency for responses to distractors, and a decrease in latency to targets; the magnitudes of the list-length effects are observed to change concomitantly.[11]

APPLICATION OF THE MODEL TO RELATED EXPERIMENTS

Other experiments have been conducted to test various features of the theory. One such study involved target sets in which any specific word was included

confusions of identification. Direct evidence of this possibility is reported by Juola et al. (1971), who showed, among other things, that distractor words graphically *very* similar to target items were responded to more slowly. In this experiment, no estimate can be made of the contribution of similarity to the overall set-size effect. However, results from several long-term recognition studies indicate that both semantic and graphemic similarity cause increased error rates as well as increased response latencies (Atkinson & Juola, 1973; Juola et al., 1971). Because there were no differences in error rates among the three groups, it is unlikely that a significant proportion of the list-length effects is attributable to similarity factors.

[11] Variables other than those represented in the model influence familiarity. Of particular importance is the effect of the number of intervening trials between successive tests on a given item. Lag effects in response latency have been observed, with the magnitude of the effect decreasing with lag for both target and distractor items (Fischler & Juola, 1971; Juola et al., 1971). This phenomenon would be accounted for in the theory by assuming that the familiarity of an item increases immediately after presentation and then gradually declines over trials. To develop this idea mathematically would complicate the model. By design, lags were relatively constant for the data treated here and need not be represented explicitly in the model.

once, twice, or three times in the list memorized prior to the experimental session. If the number of occurrences of a word in the target list affects its familiarity value, then both error rate and latency should be less for multiply represented items than for items appearing only once in the list. If, however, the word's familiarity is unaffected by repetitions in the study list, then the error rate should be the same for all target items; further, any latency effects would have to be due to a faster search of the E/K store for an item multiply represented in the target list compared with one appearing only once. The results showed that error rate and response latency were less for items that occurred two or three times in the list than for items included only once. *Model II was used to generate fits to the data*, assuming that the expected familiarity value of a target word is an increasing function of the number of times it was included in the target list; the search of the E/K store was postulated to take the same time for all items. The model provided an excellent fit to the data (Atkinson & Juola, 1973).

Other experiments have demonstrated the importance of semantic properties of words in determining the familiarity value of an item. Juola et al. (1971) reported that if synonyms of target words were used as distractors both response latencies and error rates increased over the values obtained for semantically unrelated distractors. Another experiment (Atkinson & Juola, 1973) provided target sets arranged into a tree structure to reflect the semantic hierarchy from which the words were taken. During the test session target words were selected either from a 'dense' portion of the hierarchy (one of four nodes on a branch with up to four exemplar words under each node) or from a 'sparse' portion (one of two nodes with only two exemplar words under each node). The data showed that mean latencies for positive responses were less for targets from dense portions of the tree than for targets from the sparsely represented regions. The results from these two experiments indicate that the expected familiarity value of a word can be increased by testing semantically related words.

An experiment by Juola (1973) was designed to test the importance of stimulus-encoding factors in determining an item's familiarity value. The subjects memorized a list of 48 common nouns and then were tested with either words or simple outline drawings of the objects named by the words. Both words and pictures were presented as targets and distractors, and all items were tested twice. Of interest was the nature of the repetition effects when the second test of an item was either identical in form (e.g., 'CAT' followed by 'CAT') or different (e.g., 'CAT' followed by a picture of a cat). Repetition of the same pictorial form resulted in a faster encoding time; repetition (whether in the same or a different form) also increased the familiarity value of the items. The relative importance of these two effects was estimated by comparing mean latencies for repeated targets and distractors for the case in which the exact form of the stimulus was preserved on both

tests with the case in which different forms of the item were presented on successive tests. The results showed that subjects were faster on trials in which repeated items were presented in the same form (word or picture) as they had been shown on the first presentation. This was true for distractors as well as for target items. However, there were no significant differences in the error rates for items that were tested with the same or different stimulus forms on successive presentations. These results indicate that the familiarity value of an item is relatively independent of the form of the stimulus at the time of test. However, the form of the stimulus does have an effect on encoding time.

RECOGNITION MEMORY FOR ITEMS IN SHORT-TERM STORE

The theory presented in the previous sections was originally formulated to deal with recognition experiments involving large target sets stored in LTS. It is possible, however, to extend the model to the case in which the target set consists of a small number of items in STS. The results from experiments using small memory sets have generally shown that response latencies are linear, increasing functions of the number of target items, with roughly equal slopes for positive and negative responses. A model used to account for these findings is the serial scanning process proposed by Sternberg (1966, 1969a, 1969b). According to Sternberg's model, the subject encodes the test stimulus into a form that is comparable to the internal representations of the target items stored in STS. The encoded test item is scanned in serial fashion against each of the memory items, and then a decision is made about whether or not a match was obtained. The model predicts that latency will be a linear function of memory-set size, with both positive and negative responses having the same slope but possibly different intercepts.

Whereas the Sternberg model has proved adequate in explaining the results from many short-term recognition experiments, there are reports in the literature of systematic discrepancies between data and the model's predictions. It is not possible to review these results here (see Nickerson, 1970), but variations from the model have involved departures from linearity in the functions relating response latencies to target-set size, differences in slopes between the functions obtained for positive and negative responses (including cases in which the slope for positive responses is significantly greater than that for negatives), serial position effects in the latencies of positive responses, and trial-to-trial dependencies. These findings have led some authors (Baddeley & Ecob, 1970; Corballis, Kirby, & Miller, 1972) to propose alternative models for short-term recognition memory, suggesting that response decisions might be based solely on the test item's memory strength. Strength models

usually assume that there is a single criterion along the strength continuum; values above this criterion lead to positive responses. In addition, the decision time is assumed to be greater for values near the criterion, and both the criterion itself and the mean strength value of the target items are assumed to decrease as the number of targets increases.

It is our view that the test item's familiarity value (which in some sense is comparable to a strength notion) may play the same role in the short-term case as it does in long-term recognition studies. List-length effects are still to be explained in terms of a scan of the target set, but on occasion this search may be bypassed if the test item's familiarity is very high or very low; as in the long-term case, the probability of bypassing the target-set search will depend upon the reliability of the familiarity measure in generating correct responses.

The probability of bypassing the target-set search should be minimal in experiments using a small pool of items from which targets and distractors are to be drawn on each trial, as in the Sternberg (1966) study, which involved only the digits 0 to 9. The reason is that, during an experimental session, all items in the stimulus pool receive repeated presentations, and the resulting high familiarity values become less and less useful in distinguishing targets from distractors; thus a search of the target set will be made on most trials, resulting in large list-length effects. Support for this view comes from a study by Rothstein and Morin (1972), who reported much larger slopes for the response-time function in a short-term scanning task when the stimuli were selected from a small pool (10 words) than when selected, without replacement, from a very large pool of words. For the small item pool, we assume that repeated presentation increases the familiarity of all items to a uniformly high level, thereby reducing its usefulness as a basis for responding. Thus, the probability of executing a target-set search should be maximal, causing the slope of the response-time functions to take on its maximum value.

Figure 11 presents a flow diagram of the processes involved in recognition memory for items stored in STS. As in the case for target sets stored in the E/K store (Fig. 3), the test item is first encoded and the appropriate node in the lexical store is accessed, leading to the retrieval of a familiarity value for the item. If the familiarity value is very high or very low, the subject outputs a fast response that is independent of memory-set size. For intermediate familiarity values, the subject retrieves an internal code for use in scanning STS. Thus far, the processes proposed for short-term recognition are identical with those of the long-term case. However, the internal code used to search STS may not be the same as that used in the long-term memory search. For example, Klatzky, Juola, and Atkinson (1971) provided evidence that alternative codes for the same test item can be generated and compared with either verbal or spatial representations of target-set items. After retrieval of the appropriate internal code, a search of the target list stored in STS is

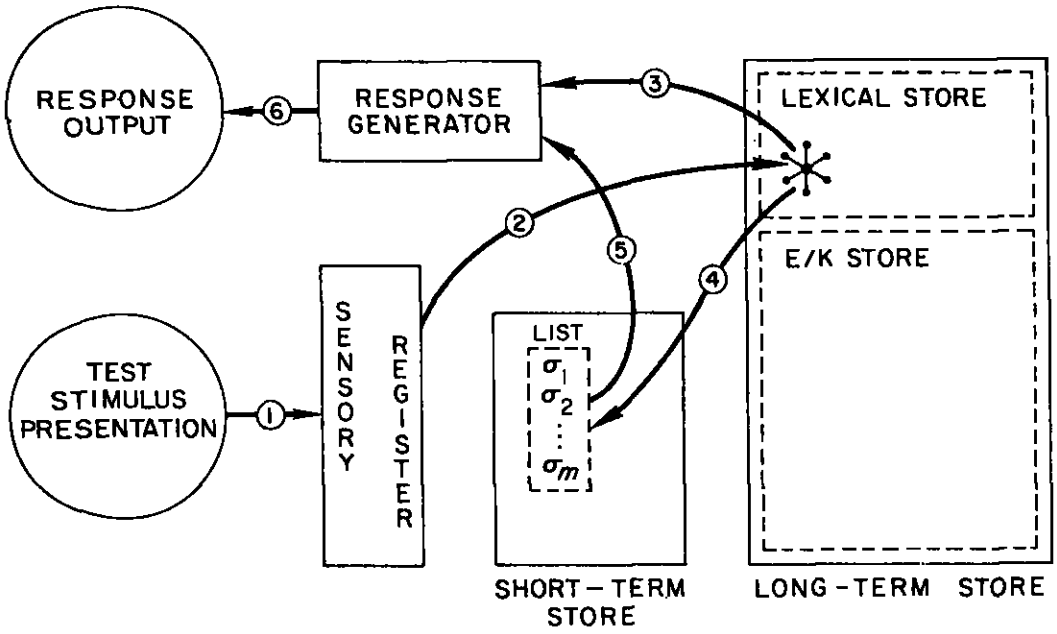


FIGURE 11.
A schematic representation of the search and decision processes in a short-term recognition memory study. A test stimulus is presented (1) and then matched to a node in the lexical store (2). The familiarity value associated with the node may lead to an immediate decision (3) and response output (6). Otherwise, a search code is extracted and scanned against the target list in STS (4), which leads to a decision (5) and subsequent response. Path (1), (2), (3), (6) results in a faster response than Path (1), (2), (4), (5), (6), and the response is independent of the size of the ST set.

executed, and a response based on the outcome of this scan is then made.

An unpublished study conducted by Charles Darley and Phipps Arabie at Stanford University was designed to assess the effects of item familiarity in a short-term memory task. The familiarity values of distractor items were manipulated to determine if this variable would affect the slopes and intercepts of the function relating latency to target-set size. On each of a long series of trials, a target set of from two to five words was presented auditorally, followed by the visual presentation of a single test word. The words used in the target sets were different on every trial of the experiment; that is, a word once used in a target set was never used in any other target set. On half the trials a word from the current target set was presented for test; these trials will be designated P trials to indicate that a 'positive response' is correct. On the other half of the trials, a distractor (a word not in the current target set) was presented for test; these trials will be called N trials because a 'negative response' is correct. The distractor words were of three types: new words never presented before in the experiment (denoted N_1, because the word was presented for the first time); words that had been presented for the first time

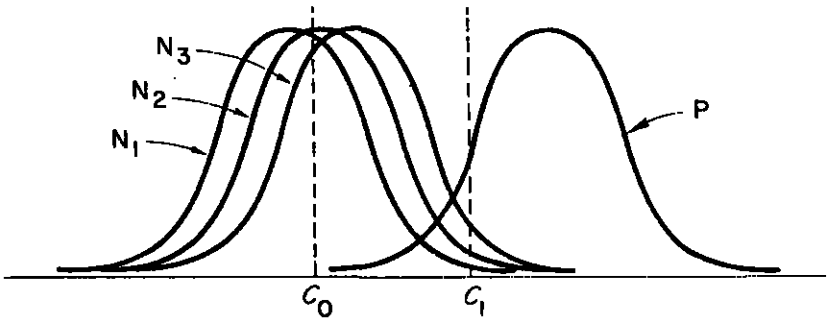


FIGURE 12.
Distributions of familiarity values for the three types of distractor items (N_1, N_2, N_3) and for target items (P).

in the experiment as distractors on the immediately preceding trial (denoted N_2, because the word was now being presented for the second time); and words that had been presented for the first time both as a member of the memory set and as a positive test stimulus on the immediately preceding trial (denoted N_3, because the word was now being presented for the third time). Thus there were four types of test items (N_1, N_2, N_3, P), and we assume that different degrees of familiarity are associated with each.

Figure 12 presents a schematic representation of the four familiarity distributions. The density functions associated with the test word on an N_1, N_2, N_3, or P trial are denoted $\phi(x; N_1)$, $\phi(x; N_2)$, $\phi(x; N_3)$, or $\phi(x; P)$, respectively; as in the previous application, these functions are assumed to be normally distributed with unit variance. Their expected values are denoted μ_{N_1}, μ_{N_2}, μ_{N_3}, and μ_P. The quantity μ_P should be largest because the test word on a P trial is a member of the current trial target set and should be very familiar; μ_{N_1} should be smallest because N_1 words are completely new; and μ_{N_2} and μ_{N_3} should be intermediate because N_2 and N_3 words appeared on the prior trial. The probabilities of errors for the four trial types are determined by the areas of the familiarity distributions above c_1 for distractors, and below c_0 for targets; that is,

$$P(\text{Error} \mid N_i) = \int_{c_1}^{\infty} \phi(x; N_i)\, dx, \quad \text{for } i = 1, 2, 3; \tag{13}$$

$$P(\text{Error} \mid P) = \int_{-\infty}^{c_0} \phi(x; P)\, dx. \tag{14}$$

Let us now derive expressions for reaction times in this situation. For simplicity, only Model II of the preceding section is considered. To obtain equations for response latencies, it is necessary to sum the time for the encoding and familiarity-retrieval process (time ℓ), the time for a fast response decision

based on the familiarity value alone (time ρ) weighted by its probability, the time for a search of the memory list in STS (time $\kappa + \alpha m$, with m defined as the size of the short-term target set) also weighted by its probability, and the time for response output (r_0 and r_1 for negative and positive responses, respectively). Thus, the expected time for a correct response is

$$t(\mathrm{N}_i) = \ell + r_0 + \left[\int_{-\infty}^{c_0} \rho\phi(x;\,\mathrm{N}_i)\,dx + \int_{c_0}^{c_1} (\rho + \kappa + \alpha m)\phi(x;\,\mathrm{N}_i)\,dx\right]\left[\int_{-\infty}^{c_1} \phi(x;\,\mathrm{N}_i)\,dx\right]^{-1} \tag{15}$$

for $i = 1, 2, 3$; and

$$t(\mathrm{P}) = \ell + r_1 + \left[\int_{c_1}^{\infty} \rho\phi(x;\,\mathrm{P})\,dx + \int_{c_0}^{c_1} (\rho + \kappa + \alpha m)\phi(x;\,\mathrm{P})\,dx\right]\left[\int_{c_0}^{\infty} \phi(x;\,\mathrm{P})\,dx\right]^{-1} \tag{16}$$

The expression $t(\mathrm{N}_i)$ gives the time to respond correctly to an N_i item, whereas $t(\mathrm{P})$ gives the time for a correct response to a P item. The preceding expressions can be written more simply if we define

$$s_i' = \left[\int_{c_0}^{c_1} \phi(x;\,\mathrm{N}_i)\,dx\right]\left[\int_{-\infty}^{c_1} \phi(x;\,\mathrm{N}_i)\,dx\right]^{-1}, \quad \text{for } i = 1, 2, 3; \tag{17}$$

$$s = \left[\int_{c_0}^{c_1} \phi(x;\,\mathrm{P})\,dx\right]\left[\int_{c_0}^{\infty} \phi(x;\,\mathrm{P})\,dx\right]^{-1}. \tag{18}$$

Then

$$t(\mathrm{N}_i) = [\ell + \rho + r_0] + s_i'[\kappa + \alpha m], \quad \text{for } i = 1, 2, 3; \tag{19}$$

$$t(\mathrm{P}) = [\ell + \rho + r_1] + s[\kappa + \alpha m]. \tag{20}$$

The quantities s_i' and s are determined by the familiarity distributions and c_1 and c_0 and are not influenced by m. Thus $t(\mathrm{N}_i)$ and $t(\mathrm{P})$ plotted as functions of m yield straight lines with slopes $\alpha s_i'$ and αs, respectively.

The latency data from the experiment are presented in Figure 13. Note that latency increases with memory-set size and is ordered such that P is fastest, and N_1, N_2, and N_3 are progressively slower. To fit the model to these data, we proceed in the same way as we did for the long-term experiment. The observed probabilities of an error on N_1, N_2, and N_3 trials were 0.008, 0.018, and 0.058, respectively. Using these error probabilities and Equation 13 yields the following relations: $\mu_{\mathrm{N}_1} = c_1 - 2.41$; $\mu_{\mathrm{N}_2} = c_1 - 2.10$; $\mu_{\mathrm{N}_3} = c_1 - 1.56$. The probability of an error on a P trial was 0.028; using Equation 13 yields $\mu_{\mathrm{P}} = c_0 + 1.91$. Setting c_0 equal to zero leaves the following five parameters to be estimated from the latency data: c_1, $(\ell + \rho + r_1)$, r, κ, α, where r is again defined as $r_0 - r_1$. An RMSD function equivalent to the one

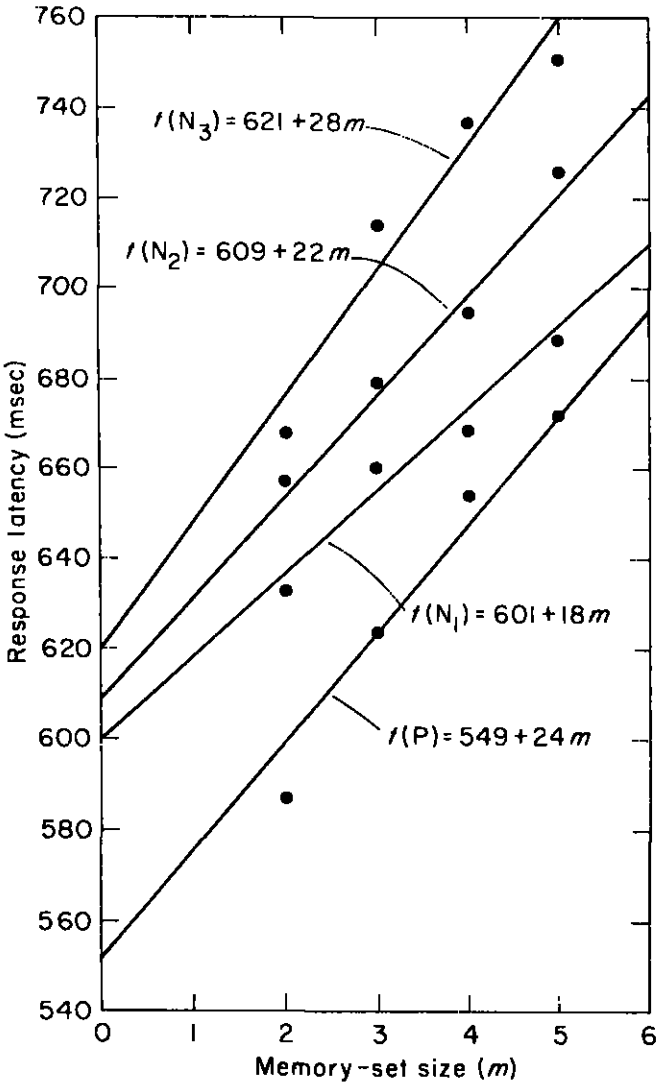


FIGURE 13.
Correct response latencies as a function of the size of the memory set. The straight lines fitted to the data represent theoretical predictions.

presented in Equation 12 was specified for the 16 data points in Figure 13, and a computer was programmed to search the parameter space for a minimum.

Table 5 presents the parameter estimates, and the theoretical predictions are graphed as straight lines in Figure 13. In carrying out these fits, 9 parameters were estimated from the data; however, there are 4 error probabilities and 16 latency measures to account for. Thus 9 of 20 degrees of freedom

TABLE 5
Parameter estimates for the short-term memory study

Parameter	Estimate
c_1	2.52
$(\ell + \rho + r_1)$	499 msec
r	64 msec
κ	70 msec
α	33.9 msec

Note: $r = r_0 - r_1$.

were used in the estimation process, leaving 11 against which to evaluate the goodness of fit.

The results in Figure 13 indicate that the familiarity value of the distractor item has a large effect, with the slopes and intercepts of the negative functions increasing with their expected familiarity values. These effects are captured by the model, which generally does a satisfactory job of fitting the data. The predicted slope of the $t(P)$ function is 24 msec, whereas the predicted slopes for $t(N_1)$, $t(N_2)$, and $t(N_3)$ go from 18 msec to 22 msec to 28 msec, respectively. If the subject ignored the familiarity measure and made a search of the memory list on every trial, then all four functions would have a slope equal to α, which was estimated to be 33.9 msec.[12]

The results shown in Figure 13 support the proposition that familiarity effects play a role in short-term memory scanning experiments. Further, these effects can be accounted for with the same model that was developed for long-term recognition studies. However, examination of the parameter estimates for the short- and long-term cases indicates that the time constants for the two processes are not the same (see Tables 4 and 5). For example, the time to initiate the extended search, κ, is 70 msec in the short-term study compared with 137 msec in the long-term study. In contrast, the search rate, α, is 33.9 msec in the short-term case and only 9.9 msec in the long-term case. Thus, the search is initiated more rapidly in the short-term case, but the search rate is faster in the long-term case. We will not pursue these comparisons here, but will return to them later.

In the next section the model is generalized to an experiment in which target items were stored in either STS or LTS, or in both. For this case, the theory must be elaborated to account for such possibilities as sequential or simultaneous search of the two memory stores and changes in the decision

[12] Similar fits were carried out using Model I, which involved estimating both α and α'. The estimate of α' was somewhat below that of α, but the goodness of fit was only slightly improved over that obtained for Model II, using one less parameter.

criteria, depending on whether the test item is potentially a member of a list stored in LTS, in STS, or in both.

AN EXPERIMENT INVOLVING BOTH LONG- AND SHORT-TERM TARGET SETS

Experiments by Wescourt and Atkinson (1972) and Mohs, Wescourt, and Atkinson (1973) were designed to compare results for the cases in which the subject maintained target sets in LTS, in STS, or in both. Figure 14 presents a flow diagram for the case in which the test stimulus could be a member of a target set in either store. When the test stimulus is presented, it is encoded and the appropriate lexical node is accessed. If the familiarity value associated with that node is above the high criterion or below the low criterion, a fast response is emitted. If familiarity is of an intermediate value, the subject executes an extended search of the two memory stores. Again, it is likely that the internal representations of items in STS and the E/K store are different;

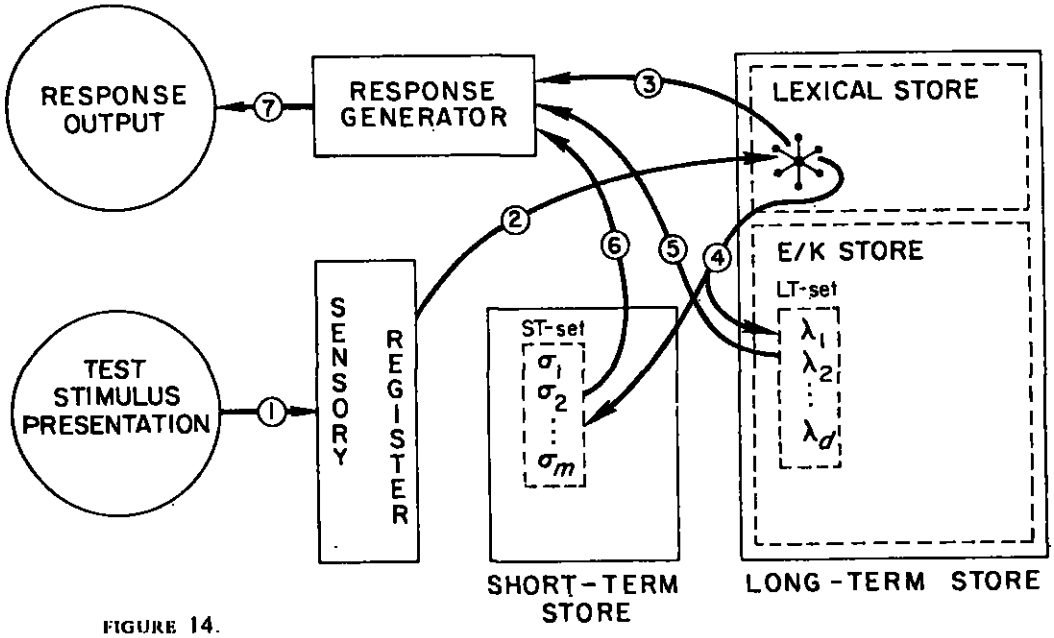


FIGURE 14.
A schematic representation for the case where part of the target set is in STS and part is in LTS. A test item is presented (1) and then matched to its node in the lexical store (2). The familiarity index of the node may lead to an immediate decision (3) and response output (7). Otherwise, an ST code and an LT code are extracted for the lexical node, and then used to search STS and LTS (4). A decision about the test item is eventually made based on the search of LTS (5) or of STS (6), and a response is output (7).

thus, different codes of the test item must be extracted from the test item's lexical node before this search can begin. The search continues until a match is obtained or until both sets are searched without finding a match, and then the appropriate response is made.

In the study considered here, two types of trial blocks were used. For one type, designated the S Block, the target set consisted of only short-term items (ST set). For the other, the M Block, the target set involved a 'mix' of both an ST set and an LT set. The ST set is distinguished from the LT set in two ways:

(1) The ST set was presented on each trial before the onset of the test stimulus; it always involved a new set of words never before used in the experiment. On the other hand, the LT set was thoroughly memorized the day before the first test session and used throughout the experiment.
(2) The ST set contained a small number of words (1 to 4), which could readily be maintained in short-term memory without taxing its capacity. The LT set consisted of a list of 30 words (memorized in serial order) stored in long-term memory.

The subjects were tested in three consecutive daily sessions (the data from the first day are not included in the results reported here). Each session was divided into M and S Blocks. On each trial of an M Block, 0 to 4 words (ST set) were presented prior to the onset of the test word. On positive trials, the test word was selected from either the LT set or the ST set if the ST set was nonempty (load condition); or the test word was selected from the LT set if there were no ST items (no-load condition). On negative trials, the test word was not in either the ST or LT set and had never been used before in the experiment. On each trial of the S Block, an ST set of from 1 to 4 words was presented prior to the onset of the test stimulus; on positive trials a word from the ST set was presented for test, and on negative trials a word never used before was presented.

Trials in the S Block are like those in a short-term memory-scanning experiment and are referred to as S trials. The no-load trials of the M Block correspond to those in a long-term recognition task such as the one reported earlier in this chapter; because tests involve only the long-term target set, these trials are called L trials. The load trials of the M Block require the subject to evaluate a test word against both an ST set and the LT set, and they are called M trials. Thus, S trials involve a pure test of short-term memory, L trials a pure test of long-term memory, and M trials involve a mix of both short- and long-term memories.[13] Figure 15 illustrates the various trial types.

[13] Studies of this sort have been reported by Forrin and Morin (1969) and Doll (1971). However, they have employed very small LT sets, and there is the possibility that the subject could enter the entire LT set into short-term memory on some or all of the trials. Thus a complete separation of the long- and short-term searches might not have been achieved.

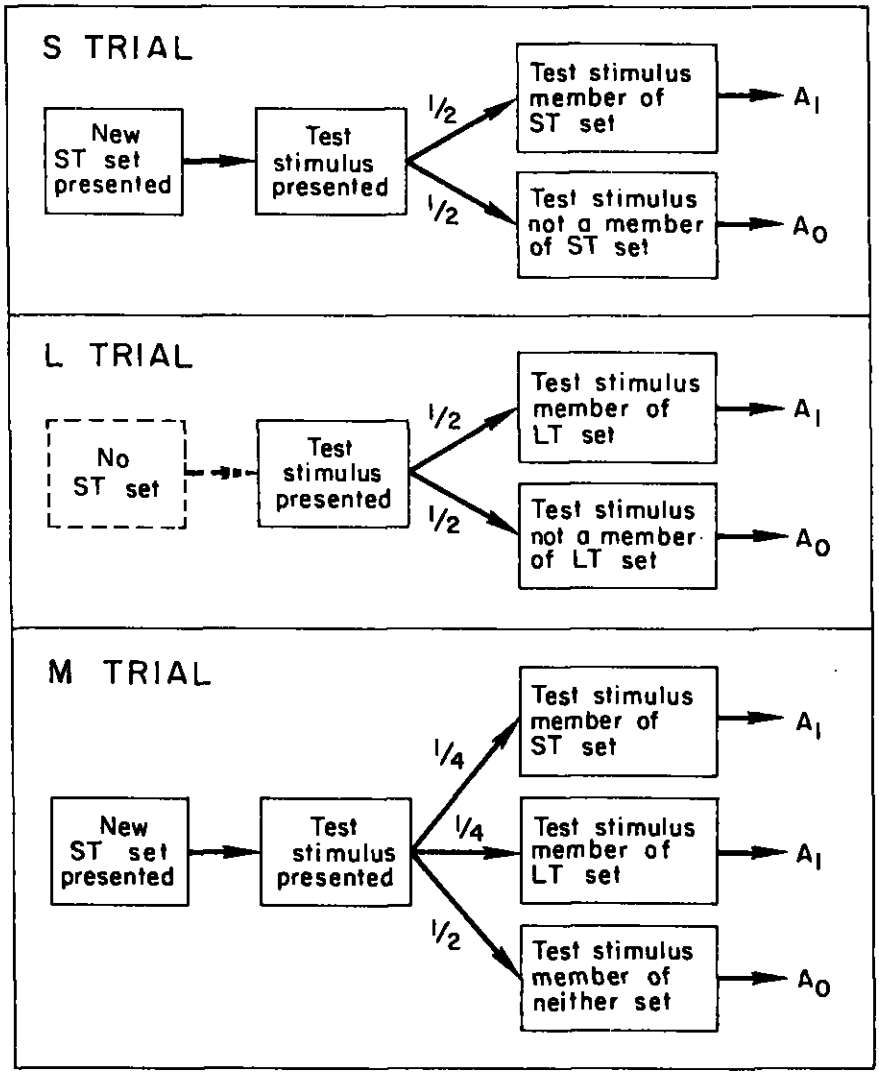


FIGURE 15.
Diagram representing the three types of trials. In all blocks, distractors involve words never presented before in the experiment.

Figure 16 presents the mean latencies of correct responses for the various trial types. The straight lines fitted to the data represent theoretical predictions and are discussed later. In discussing these results, it is useful to adopt the notation defined in Table 6. In all cases these measures refer to the latency of a correct response. The subscript on t indicates the trial type (S, L, or M); the P in parentheses indicates that a positive response was correct (i.e., a

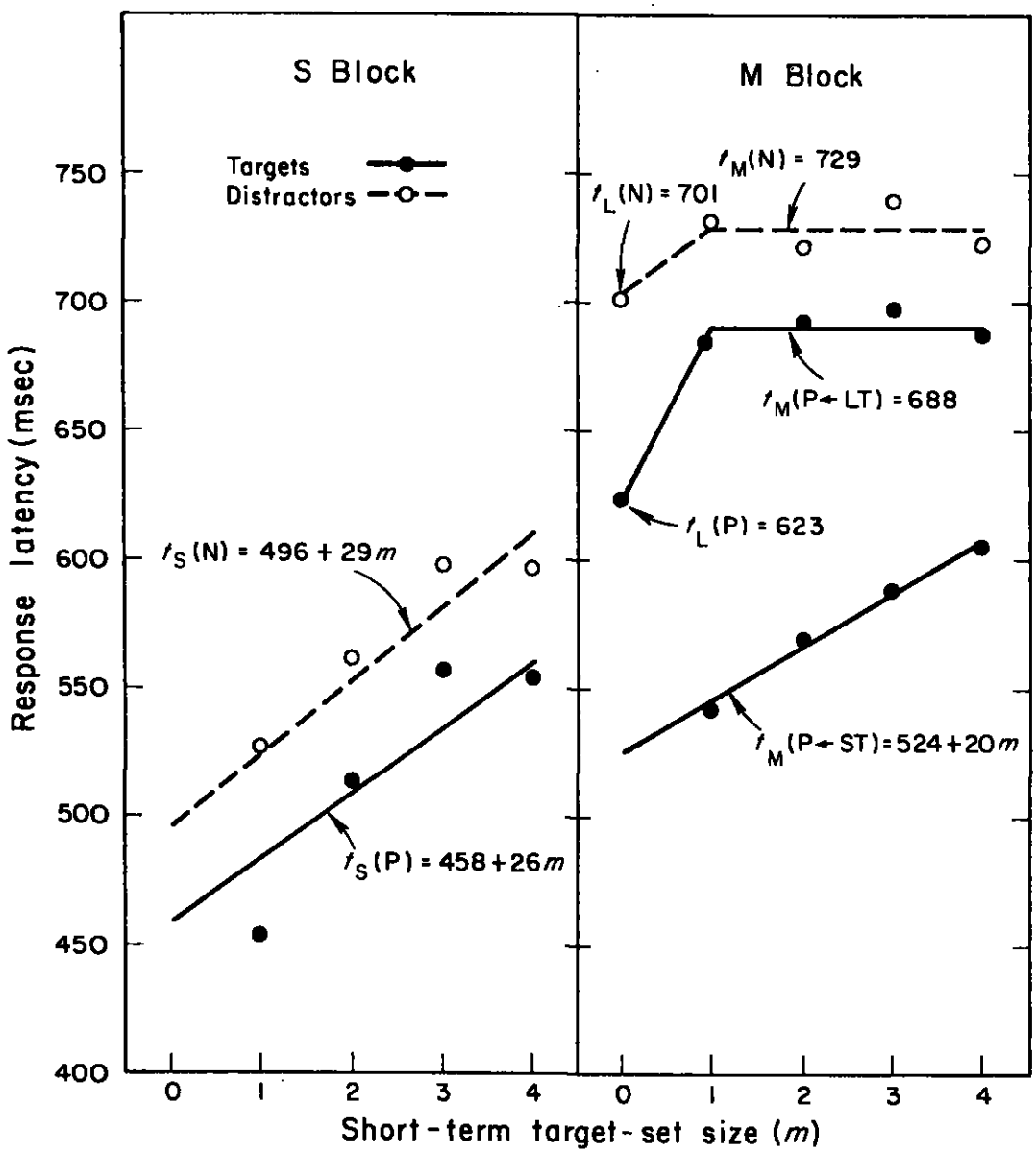


FIGURE 16.
Mean response latencies as functions of ST-set size (m) for the S Block (left panel) and the M Block (right panel). The linear functions fitted to the data are explained in the text.

TABLE 6
Definition of notation

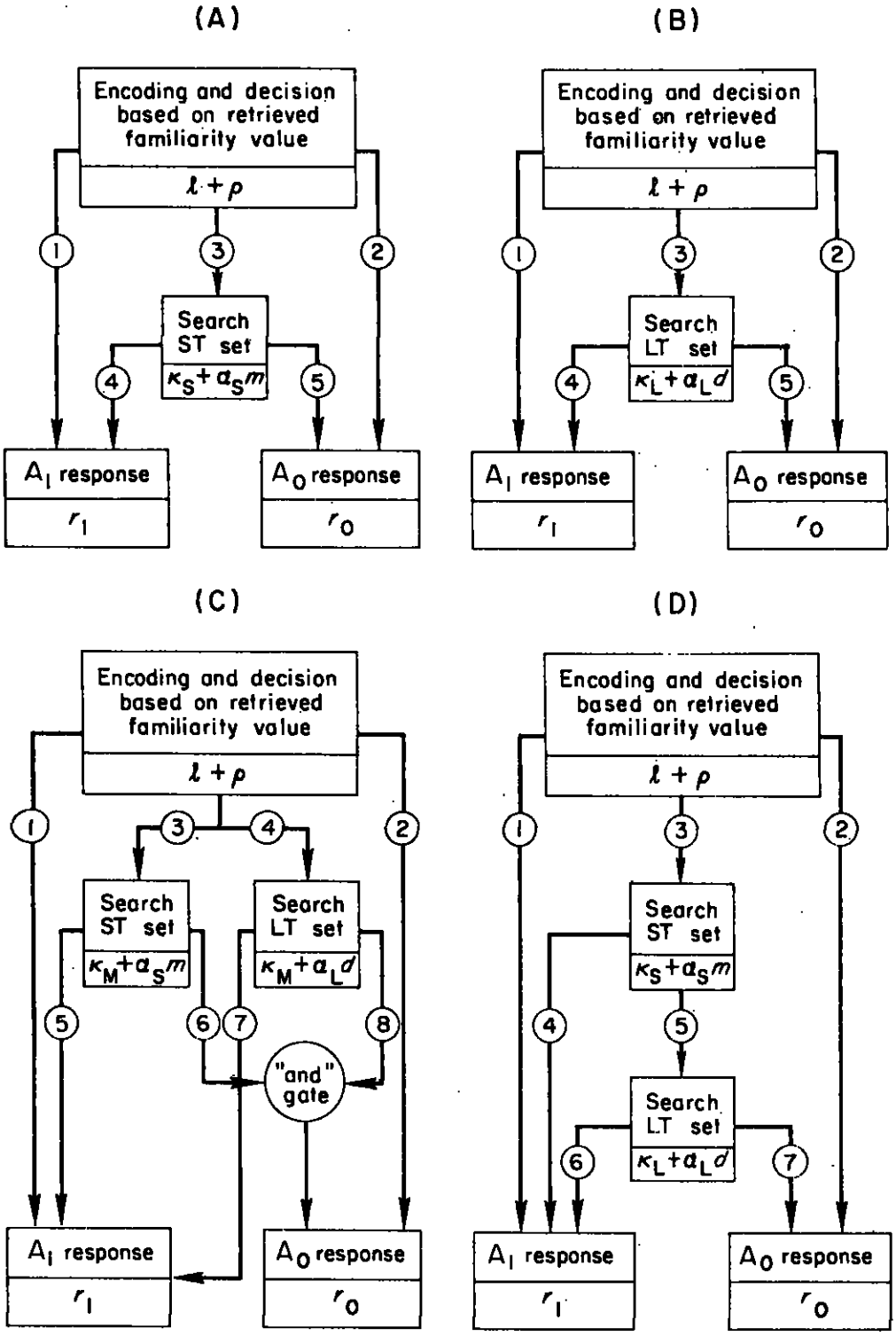
Notation	Definition
$t_S(P)$	Time for a positive response on an S trial
$t_S(N)$	Time for a negative response on an S trial
$t_L(P)$	Time for a positive response on an L trial
$t_L(N)$	Time for a negative response on an L trial
$t_M(P \leftarrow ST)$	Time for a positive response to a test item from the ST set on an M trial
$t_M(P \leftarrow LT)$	Time for a positive response to a test item from the LT set on an M trial
$t_M(N)$	Time for a negative response on an M trial.

target item was presented for test), whereas N indicates that a negative response was correct (i.e., a distractor was presented for test).

Inspection of Figure 16 shows that the observed values for $t_S(P)$, $t_S(N)$, and $t_M(P \leftarrow ST)$ are all increasing functions of m, the size of the ST set. In contrast, neither $t_M(P \leftarrow LT)$ nor $t_M(N)$ appears to be systematically influenced by the size of the ST set. The presence or absence of an ST set in the M Block, however, does have an effect, as is evident by comparing responses on L trials with comparable ones on M trials. Specifically, note that the four observed values for $t_M(P \leftarrow LT)$ are well above $t_L(P)$, and that the four $t_M(N)$ values are above $t_L(N)$.

The model to be tested against these data assumes that the extended searches are executed separately in STS and in the E/K store. The questions to be asked involve the notion of whether the two memory stores are searched sequentially or simultaneously. Figure 17 presents several flowcharts that represent the differences between serial and parallel searches of STS and the E/K store. The diagram in Figure 17(A) represents the sequence of events on an S trial and corresponds to the short-term recognition model presented

FIGURE 17. (*facing page*)
Flowcharts representing models for processing strategies in searching the memory stores. The model for S trials is shown in Panel A; arrows (1) and (2) represent fast responses based on familiarity alone, whereas (4) and (5) represent responses after a search of STS has occurred. The model for L trials is shown in Panel B and has the same interpretation as Panel A except that the search involves the E/K store. Two alternative models for M trials are presented in the bottom two panels. Panel C presents a parallel search. As before (1) and (2) indicate fast responses based on familiarity; (3) and (4) indicate that the searches of STS and the E/K store are done simultaneously. If the test item is found in the ST set (5) or in the LT set (7), a positive response is made; if the item is not found in the ST set (6) the subject has to wait for a similar outcome from the search of the LT set (8) before a negative response can be made. In Panel D, a sequential search model is presented for M trials. The arrows (1) and (2) represent fast responses based on familiarity. When a search is required, the ST set is examined first (3). If a match is found, a positive response is made (4); if not, the LT set is searched (5). When the LT-set search is complete, either a positive (6) or negative response (7) is output.


(A)
Encoding and decision based on retrieved familiarity value
$l + \rho$
1
3
2
Search ST set
$\kappa_S + \alpha_S m$
4
5
A_1 response
r_1
A_0 response
r_0
(B)
Encoding and decision based on retrieved familiarity value
$l + \rho$
1
3
2
Search LT set
$\kappa_L + \alpha_L d$
4
5
A_1 response
r_1
A_0 response
r_0
(C)
Encoding and decision based on retrieved familiarity value
$l + \rho$
1
3
4
2
Search ST set
$\kappa_M + \alpha_S m$
Search LT set
$\kappa_M + \alpha_L d$
5
6
7
8
"and" gate
A_1 response
r_1
A_0 response
r_0
(D)
Encoding and decision based on retrieved familiarity value
$l + \rho$
1
3
2
Search ST set
$\kappa_S + \alpha_S m$
4
5
Search LT set
$\kappa_L + \alpha_L d$
6
7
A_1 response
r_1
A_0 response
r_0

in the preceding section. It assumes that initially the subject makes a familiarity estimate of the test item, and on this basis he outputs a fast positive or negative response if its value is above the high criterion or below the low criterion, respectively. Otherwise, the subject delays his response until a search of STS has been made, the length of this search being a linear function of m (the size of the ST set). Figure 17(B) represents the stages involved on an L trial. Again, the subject can output a fast negative or positive response based on familiarity alone. Otherwise he initiates a search of the E/K store before responding; the time for this search is a linear function of d (the size of the LT set).[14]

For M trials there are at least two search strategies that suggest themselves. First, it is possible that the subject might search both STS and the E/K store simultaneously, outputting a response when the test item is found or when both stores have been searched exhaustively without finding the target. This strategy is represented in Figure 17(C). Alternatively, it is possible that the two memory stores are searched sequentially. Because response time is less to a test item from the ST set than to one from the LT set, we assume that STS is searched first, as shown in Figure 17(D). For both of these M-trial strategies, a fast response will be emitted before a search of either store is made if the retrieved familiarity value is above the high criterion or below the low criterion.

Examination of the data in Figure 16 indicates that the sequential model of Figure 17(D) can be rejected. In this model, the search of the E/K store cannot begin until the STS scan has been completed. Because the length of the STS search depends on the size of the ST set, the beginning of the search of the E/K store and, in turn, t_M(P ← LT) and t_M(N) should increase as the ST set increases. The data in Figure 16 indicate that this is not the case; both t_M(P ← LT) and t_M(N) appear to be independent of ST-set size. However, these data are compatible with a parallel search model of the type shown in Figure 17(C), if it is assumed that the rate of search in the E/K store is independent of the number of ST items. In order to make a detailed analysis of the models shown in Figure 17, we must derive theoretical equations and fit them to the data.

THEORETICAL PREDICTIONS FOR THE STS-LTS INTERACTION STUDY

The decision stage of the general model, as represented in Figure 5, must be adapted to account for the experimental conditions of the experiment. It is necessary to allow for differences in the decision process, depending on

[14] Throughout this chapter, d is used to denote the size of a long-term target set and m to denote the size of a short-term target set.

whether the test item is potentially located in STS only, in the E/K store only, or in both. These differences may be included in the model either by allowing the means of the familiarity distributions to vary as a function of the trial type, or by allowing the decision criteria to change. For the present analysis, we assume that the means of the familiarity distributions are constant over all conditions. This seems to be the most parsimonious assumption; familiarity should be a property of the test stimulus, but the subject could be expected to adjust his decision criteria differently depending on whether it is an S trial, an L trial, or an M trial. Three familiarity distributions are specified: one associated with a test item drawn from the ST set; another for a test item from the LT set; and the third for a distractor item. These distributions are assumed to be unit-normal, with cumulative distribution functions $\Phi_S(\cdot)$, $\Phi_L(\cdot)$, and $\Phi_D(\cdot)$, respectively. The means of the distributions are designated μ_S, μ_L, and μ_D, and they are assumed to be fixed for the data analyzed in this chapter. The reasons for fixing the means are the following: distractor items and ST items appear only once during the experiment, and thus repetition effects on familiarity are not a factor; for the LT items, we treat data only after these items have had several prior tests, and their familiarity should be close to an asymptotic level.

Figure 18 presents a diagram of the familiarity distributions as they apply on S, L, and M trials. Note that the mean for each distribution is placed at the same point on the familiarity scale, no matter what type of trial is involved. Differences in the decision process arise because the subject can set his criteria at different values in anticipation of an S, L, or M trial. This possibility is indicated in Figure 18. The low and high criterion values are denoted as $c_{0,S}$ and $c_{1,S}$ for S trials; as $c_{0,L}$ and $c_{1,L}$ for L trials; and as $c_{0,M}$ and $c_{1,M}$ for M trials. How the subject sets the criteria depends on the trade-off he is willing to accept between speed and accuracy; the nature of the trade-off, of course, varies as a function of the trial type.

Notation comparable to that in Table 6 is used to denote error probabilities. For example, $E_S(P)$ denotes the probability of an error on an S trial for which the correct response was positive. This probability is the tail of the ST distribution to the left of $c_{0,S}$ in Figure 18. Table 7 presents theoretical expressions for the various types of errors.

As before, it is possible to derive equations for response latencies by weighting each stage of the process by the probability that it occurs, and then summing over stages. On every trial the test stimulus must be encoded and the appropriate node in the lexical store accessed; time for this stage is ℓ and is assumed to be the same for all trial types. Next, the subject must make a decision based on the retrieved familiarity value; using Model II, we assume that this decision time is ρ and also is independent of the trial type. If a fast positive or negative response is called for, based on the familiarity value, it will be executed with time r_1 or r_0, respectively.

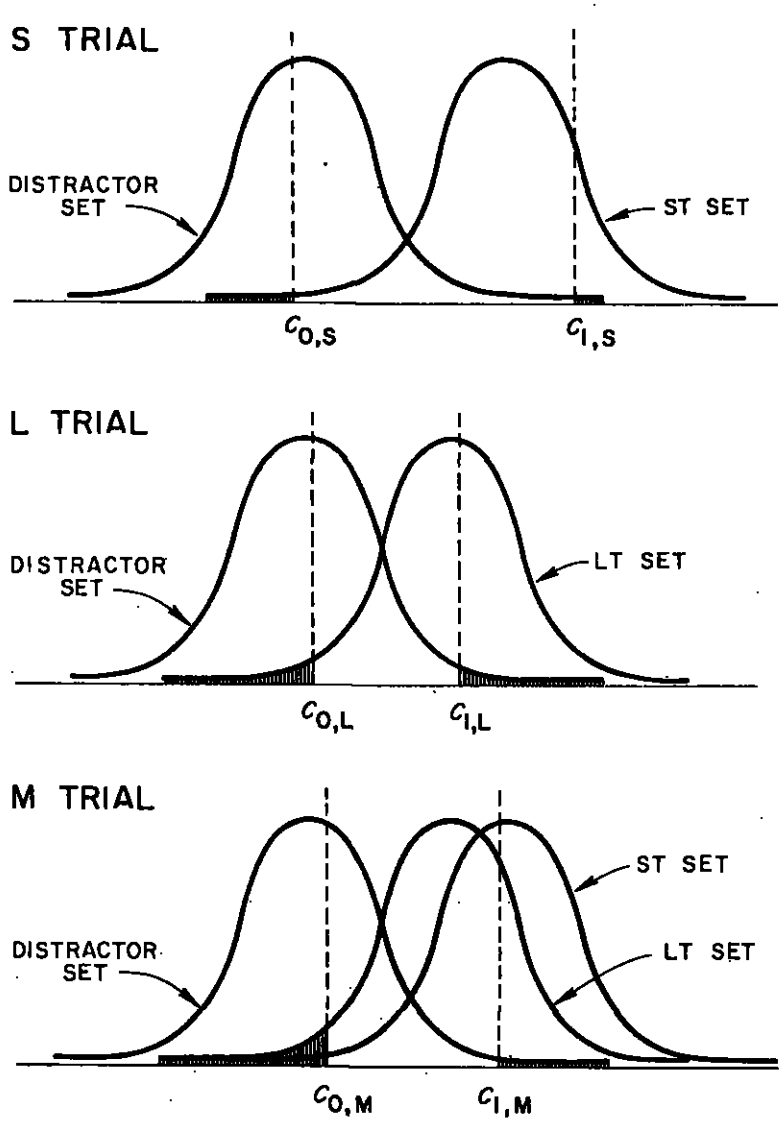


FIGURE 18.
Distributions of familiarity values for the three trial types.

When the familiarity value falls between the two criterion values, a search of the stored target list or lists is required. The nature of this search depends on the trial type because different internal codes may be used and different memory stores scanned. Three classes are to be considered.

S trials. An ST code is extracted from the test item's lexical node and then is scanned against the target set in STS; the time to extract the code

TABLE 7
Theoretical expressions for the probabilities of seven types of errors

S trials	L trials	M trials
$E_S(P) = \Phi_S(c_{0,S})$	$E_L(P) = \Phi_L(c_{0,L})$	$E_M(P \leftarrow ST) = \Phi_S(c_{0,M})$
$E_S(N) = 1 - \Phi_D(c_{1,S})$	$E_L(N) = 1 - \Phi_D(c_{1,L})$	$E_M(P \leftarrow LT) = \Phi_L(c_{0,M})$
		$E_M(N) = 1 - \Phi_D(c_{1,M})$

is denoted as κ_S, and then time $m \cdot \alpha_S$ is required to scan the m items in the ST set.[15]

L trials. An LT code is extracted from the lexical node, which takes time κ_L, and then is scanned against the d items in the LT set, which takes time $d \cdot \alpha_L$ (d in the experiment is 30).

M trials. Both an ST code and an LT code are extracted from the node, and each is scanned against the appropriate list. The extraction of the two codes takes time κ_M, and the respective scans take times $m \cdot \alpha_S$ and $d \cdot \alpha_L$. (Thus, a positive response to an ST or LT item takes time $m \cdot \alpha_S$ or $d \cdot \alpha_L$, respectively; a negative response takes time $d \cdot \alpha_L$ because both lists must be scanned, and the time is determined by the slowest scan which always involves the LT set.)

Whichever of these three cases applies, a positive or negative response—once a decision has been made—requires time r_1 or r_0, respectively.[16]

In terms of these assumptions, we can derive expressions for the latency of a correct response for each of the trial types. The derivation is similar to that for Equations 15 and 16, and only the results are presented:

$$t_S(P) = (\ell + \rho + r_1) + s_S(\kappa_S + m\alpha_S); \quad (21a)$$

$$t_S(N) = (\ell + \rho + r_0) + s'_S(\kappa_S + m\alpha_S); \quad (21b)$$

$$t_L(P) = (\ell + \rho + r_1) + s_L(\kappa_L + d\alpha_L); \quad (22a)$$

[15] The parameters κ and α are used here in the same way as in earlier accounts of the theory. The subscript indicates that κ depends on the code(s) to be extracted, and α on the memory store to be scanned.

[16] It is assumed that α_L is independent of the size of the ST set, and that any differences in scanning the LT set on L trials and on M trials is due to κ_L and κ_M, respectively. Independent support for this assumption comes from a study that replicated the M-Block trial sequence, except for the fact that all targets were drawn from the LT set. Subjects had to maintain a set of items in STS (that varied from 0 to 4 words); however, they knew that the test would involve either an LT item or a distractor. Under these conditions the latency of a positive response to an LT item and the latency of a negative response to a distractor were both constant as the ST-set size varied from 0 to 4 (i.e., no change in latency occurred when an ST set was or was not present). In this experiment the scan of the LT set was determined by α_L and κ_L on all trials; the parameter κ_M was not required because only the LT code had to be extracted from the lexical node on both L trials and M trials.

$$t_L(N) = (\ell + \rho + r_0) + s'_L(\kappa_L + d\alpha_L); \quad (22b)$$

$$t_M(P \longleftarrow ST) = (\ell + \rho + r_1) + s_{M,S}(\kappa_M + m\alpha_S); \quad (23a)$$

$$t_M(P \longleftarrow LT) = (\ell + \rho + r_1) + s_{M,L}(\kappa_M + d\alpha_L); \quad (23b)$$

$$t_M(N) = (\ell + \rho + r_0) + s'_M(\kappa_M + d\alpha_L). \quad (23c)$$

The s functions in these equations represent the probability of an extended search conditional on the occurrence of a correct response; they are comparable to those in Equations 17 and 18 and are given in Table 8.

TABLE 8
Probability of an extended memory search conditional on a correct response

s function	Theoretical expressions
s_S	$[\Phi_S(c_{1,S}) - \Phi_S(c_{0,S})][1 - \Phi_S(c_{0,S})]^{-1}$
s'_S	$[\Phi_D(c_{1,S}) - \Phi_D(c_{0,S})][\Phi_D(c_{1,S})]^{-1}$
s_L	$[\Phi_L(c_{1,L}) - \Phi_L(c_{0,L})][1 - \Phi_L(c_{0,L})]^{-1}$
s'_L	$[\Phi_D(c_{1,L}) - \Phi_D(c_{0,L})][\Phi_D(c_{1,L})]^{-1}$
$s_{M,S}$	$[\Phi_S(c_{1,M}) - \Phi_S(c_{0,M})][1 - \Phi_S(c_{0,M})]^{-1}$
$s_{M,L}$	$[\Phi_L(c_{1,M}) - \Phi_L(c_{0,M})][1 - \Phi_L(c_{0,M})]^{-1}$
s'_M	$[\Phi_D(c_{1,M}) - \Phi_D(c_{0,M})][\Phi_D(c_{1,M})]^{-1}$

In fitting the model to the data, we used a procedure somewhat different from the one employed in the previous experiments. An RMSD function comparable to that given in Equation 12 was defined, but it was composed of two components that were weighted and summed. The first component involved deviations between the 7 observed and predicted error probabilities of Table 7, and the second component involved deviations between the 22 observed and predicted latencies given by Equations 21 through 23. Parameter estimates were then obtained by using a computer to search the parameter space and obtain values that minimized the RMSD function; in the search μ_D was arbitrarily set at zero. The parameter estimates are given in Table 9. Fifteen parameters were estimated from the data, but there are 7 error probabilities and 22 latency measures to be predicted; thus 15 of 29 degrees of freedom were used in parameter estimation, leaving 14 against which to judge the goodness of fits.

The theoretical fits for the latency data are presented as straight lines in Figure 16. The most deviant point is that for $t_S(P)$ when $m = 1$. This particular discrepancy is not unexpected in view of previous research (Juola & Atkinson, 1971); it appears that for a memory set of one item (in the pure

TABLE 9
Parameter estimates

Latency measures	Familiarity measures and decision criteria
$(t + \rho + r_1) = 408$ msec	$\mu_D = 0$
$r = 30$ msec	$\mu_L = 1.53$
$\kappa_S = 69$ msec	$\mu_S = 1.51$
$\kappa_L = 140$ msec	$c_{0,S} = -0.99$
$\kappa_M = 207$ msec	$c_{1,S} = 2.13$
$\alpha_S = 35.0$ msec	$c_{0,L} = -0.33$
$\alpha_L = 9.8$ msec	$c_{1,L} = 1.56$
	$c_{0,M} = -0.25$
	$c_{1,M} = 1.72$

Note: $r = r_0 - r_1$.

short-term case) a decision can be based on a direct comparison between a sensory image of the memory item and the sensory input for the test item. Thus, a different process is operative on these particular trials, leading to unusually fast response times. Otherwise, the fits displayed in Figure 16 are quite good, given the linear character of the predictions.[17] Also, the parameter estimates are ordered in the expected way. The estimate of κ_S is less than κ_L, as would be expected by comparing the κ values for the long-term and short-term recognition experiments given in Tables 4 and 5; κ_M is the largest of the group and should be since it involves extracting both an ST and LT code. There is close agreement between the estimate of κ_S in this study (69 msec) and in the short-term study (70 msec); similarly, the estimate of κ_L (140 msec) agrees with the corresponding estimate in the long-term study (137 msec). The α values are also ordered as expected, with a much slower search rate for the ST set than for the LT set. Note that the estimate of α_S (35.0 msec) is close to the α value estimated for the short-term study (33.9 msec), and that α_L (9.8 msec) is virtually identical to the α value estimated for the

[17] The curvilinear component in the data of the left panel of Figure 16 [excluding $t_S(P)$ for $m = 1$] was unexpected, because a study by Juola and Atkinson (1971), using a similar procedure but employing only S-type trials, yielded quite straight lines. (For a comparison of the two procedures, see Wescourt & Atkinson, 1973.) The model presented in this chapter can be generalized to yield curvilinear predictions. One possibility is that the subject adjusts his decision criteria as a function of the ST-set size: when the large memory set is presented, he anticipates a slow response and attempts to compensate by adjusting the criteria to generate more fast responses based on familiarity alone. Another possibility is that, under certain experimental conditions, the familiarity of the target items depends on their serial position in the study list (Burrows & Okada, 1971). This assumption would lead to serial position effects and could also account for the curvilinear effects noted here. For a discussion of these possibilities see Atkinson, Herrmann, and Wescourt (1974).

long-term study (9.9 msec). Differences in response keys and stimulus displays make it doubtful that ($\ell + \rho + r_1$) or r should agree across the three studies reported in this chapter. The parameters that one might hope to be constant over experiments do indeed seem to be, providing some support for the model beyond the goodness-of-fit demonstration.

SUMMARY AND CONCLUSIONS

In this chapter we have considered a model for recognition memory. The model assumes that, when a test stimulus is presented, the subject accesses the lexical store and retrieves a familiarity value for the stimulus. Response decisions based only on this familiarity value can be made very quickly, but result in a relatively high error rate. If the familiarity value does not provide the subject with sufficient information to respond with confidence, a second search of a more extended type is executed. This latter search guarantees that the subject will arrive at a correct decision, but with a consequent increase in response latency. By adjusting the criteria for emitting responses based on familiarity versus those based on an extended memory search, 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 search and decision stages proposed in the present chapter are indicative of possible mechanisms involved in recognition. We do not, however, believe that they provide a complete description of the processes involved; the comparisons of data with theoretical predictions are reported mainly to demonstrate that many features of our results can be described adequately by the model.

There are several additional observations, however, that suggest that the memory and decision components of the model correspond to processing stages of the subject. Introspective reports indicate that subjects might indeed output a rapid response based on tentative, but quickly retrieved, information about the test stimulus. Subjects report that they are sometimes able to respond almost immediately after the word is presented without '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 responding. The fact that subjects are always aware of their errors supports the general outline of the model; even if the initial familiarity of an item produces a decision to respond immediately, the search of the appropriate memory store continues and, when completed, permits the subject to confirm whether or not his response was correct.

Additional support for the model comes from its generality to a variety of experimental paradigms (for examples, see Atkinson & Juola, 1973). As

reported here, the model can be used to predict response times in recognition tasks with target sets stored in LTS or STS, or in both. It can also handle results from other classes of recognition experiments, such as those employing the Shepard-Teghtsoonian paradigm (e.g., Hintzman, 1969; Okada, 1971). The differences in results from these various types of tasks can be explained in terms of the extended memory search stage; the likelihood that the subject delays his response and makes an extended search of memory is determined by the criteria he adopts to minimize errors while still insuring fast responses. Once the extended search is initiated, its exact nature depends on how the target set is stored in memory (Smith, 1968). If the target set is a well-ordered and thoroughly memorized list of words, the extended search will involve systematic comparisons between the test stimulus and the target items. On the other hand, the target set may be represented in memory as a list of critical attributes (Meyer, 1970). In this case, the extended search would involve checking features of the test stimulus against the attribute list (Neisser, 1967). The dependency of latency on target-set size then would be determined by the relationship between the number of attributes needed to unambiguously specify a target set and the set's size. Finally, target items may be weakly represented in memory (e.g., because they received only a single study presentation); then the extended search might be aimed at retrieving contextual information, with search time relatively independent of target-set size (Atkinson, Herrmann, & Wescourt, 1974; Atkinson & Wescourt, 1974).

These speculations about recognition memory and the nature of the specific task lead to certain testable hypotheses. If the subject adjusts his criteria to balance errors against response speed, different instructions could be used to alter the criteria. For example, if the target set is a well-memorized list of words, and the subject is instructed to make every effort to avoid errors, the appropriate strategy would be to always conduct the extended search before responding. Because the time necessary to complete this search depends on target-set size, both overall latency and list-length effects should increase. Alternatively, if response speed is emphasized in the instructions, the subject should respond primarily on the basis of familiarity. In this case, responses would be emitted without an extended search, and overall latency would decrease and there should be few, if any, list-length effects.

For the theory described in this chapter, the encoding process that permits access to the appropriate node in the lexical store is assumed to occur without error and at a rate independent of the size and make-up of the target set. For highly familiar and minimally confusable words, this assumption appears to be reasonable and is supported by our data. However, for many types of stimuli, increases in target-set size will lead to greater confusability and consequently to slower, as well as less accurate, responses (Juola et al., 1971). When this is the case, the explanation of the set-size effect given here will not be sufficient, for we have assumed that it is due entirely to the extended mem-

ory search. Analyses of set-size effects in the framework of this theory would be inappropriate if the experiment were not designed to minimize confusions among stimuli. The theory can be extended to encompass confusion effects by reformulating the encoding scheme and perhaps the extended search process. However, the result would be a cumbersome model with so many interacting processes that it would be of doubtful value as an analytic tool. Trying to account for stimulus confusability in a theory of recognition memory is too ambitious a project, given our current state of knowledge. Greater progress can be made by employing experimental paradigms specifically designed to study recognition memory and other paradigms specifically designed to study confusions among stimuli.

ACKNOWLEDGMENTS

This research was supported by grants from the National Institute of Mental Health (MH 21747) and the National Science Foundation (NSFGJ 443X3). The authors are indebted to J. C. Falmagne and D. J. Herrmann for comments and criticisms on an early draft of this paper.

REFERENCES

Anderson, J. R., & Bower, G. H. Recognition and retrieval processes in free recall. *Psychological Review*, 1972, **79**, 97-123.

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 Assoc., 1974.

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

Atkinson, R. C., & Juola, J. F. Factors influencing speed and accuracy of word recognition. In S. Kornblum (Ed.), *Fourth international symposium on attention and performance*. New York: Academic Press, 1973.

Atkinson, R. C., & Shiffrin, R. M. Human memory: A proposed system and its control processes. In K. W. Spence and J. T. Spence (Eds.), *The psychology of learning and motivation*. Vol. II. New York: Academic Press, 1968.

Atkinson, R. C., & Shiffrin, R. M. The control of short-term memory. *Scientific American*, 1971, **225**, 82-90.

Atkinson, R. C., & Wescourt, K. T. Some remarks on a theory of memory. In P. Rabbitt and S. Dornic (Eds.), *Fifth international symposium on attention and performance*. London: Academic Press, 1974.

Baddeley, A. D., & Ecob, J. R. *Reaction time and short-term memory: A trace strength alternative to the high-speed exhaustive scanning hypothesis*. Technical Report No. 13. San Diego: Center for Human Information Processing, University of California, 1970.

Banks, W. P. Signal detection theory and human memory. *Psychological Bulletin*, 1970, **74**, 81–99.

Burrows, D., & Okada, R. Serial position effects in high-speed memory search. *Perception & Psychophysics*, 1971, **10**, 305–308.

Corballis, M. C., Kirby, J., & Miller, A. Access to elements of a memorized list. *Journal of Experimental Psychology*, 1972, **94**, 185–190.

Doll, T. J. Motivation, reaction time, and the contents of active verbal memory. *Journal of Experimental Psychology*, 1971, **87**, 29–36.

Egeth, H., Marcus, N., & Bevan, W. Target-set and response-set interactions: Implications for models of human information processing. *Science*, 1972, **176**, 1447–1448.

Fischler, I., & Juola, J. F. Effects of repeated tests on recognition time for information in long-term memory. *Journal of Experimental Psychology*, 1971, **91**, 54–58.

Forrin, B., & Morin, R. E. Recognition time for items in short- and long-term memory. *Acta Psychologica*, 1969, **30**, 126–141.

Herrmann, D. J. The effects of organization in long-term memory on recognition latency. Unpublished doctoral dissertation, University of Delaware, 1972.

Hintzman, D. Recognition time: Effects of recency, frequency, and the spacing of repetitions. *Journal of Experimental Psychology*, 1969, **79**, 192–194.

Juola, J. F. Repetition and laterality effects on recognition memory for words and pictures. *Memory & Cognition*, 1973, **1**, 183–192.

Juola, J. F., & Atkinson, R. C. Memory scanning for words vs. categories. *Journal of Verbal Learning and Verbal Behavior*, 1971, **10**, 522–527.

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. Memory and decision aspects of recognition learning. *Psychological Review*, 1967, **74**, 496–504.

Kintsch, W. *Learning, memory, and conceptual processes*. New York: John Wiley, 1970. (a)

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

Klatzky, R. L., Juola, J. F., & Atkinson, R. C. Test stimulus representation and experimental context effects in memory scanning. *Journal of Experimental Psychology*, 1971, **87**, 281–288.

Mandler, G., Pearlstone, Z., & Koopmans, H. S. Effects of organization and semantic similarity on recall and recognition. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 410–423.

McCormack, P. D. Recognition memory: How complex a retrieval system. *Canadian Journal of Psychology*, 1972, **26**, 19–41.

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

Miller, G. A. The organization of lexical memory: Are word associations sufficient? In G. A. Talland and N. C. Waugh (Eds.), *The pathology of memory*. New York: Academic Press, 1969.

Mohs, R. C., Wescourt, K. T., & Atkinson, R. C. Effects of short-term memory

contents on short- and long-term memory searches. *Memory & Cognition*, 1973, **1**, 443–448.

Morton, J. The interaction of information in word recognition. *Psychological Review*, 1969, **76**, 165–178.

Morton, J. A functional model for memory. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970.

Murdock, B. B., Jr. A parallel-processing model for scanning. *Perception & Psychophysics*, 1971, **10**, 289–291.

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

Nickerson, R. S. *Binary-classification reaction time: A review of some studies of human information-processing capabilities*. Technical Report No. 2004. Cambridge, Mass.: Bolt, Beranek, and Newman, 1970.

Norman, D. A. (Ed.) *Models of human memory*. New York: Academic Press, 1970.

Okada, R. Decision latencies in short-term recognition memory. *Journal of Experimental Psychology*, 1971, **90**, 27–32.

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

Rothstein, L. D., & Morin, R. E. The effects of size of the stimulus ensemble and the method of set size manipulation on recognition reaction time. Paper presented at the meetings of the Midwestern Psychological Association, Cleveland, Ohio, May 1972.

Rubenstein, H., Garfield, L., & Millikan, J. A. Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 487–492.

Schank, R. Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology*, 1972, **3**, 552–631.

Schvaneveldt, R. W., & Meyer, D. E. Retrieval and comparison processes in semantic memory. In S. Kornblum (Ed.), *Fourth international symposium on attention and performance*. New York: Academic Press, 1973.

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

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

Shevell, S., & Atkinson, R. C. A theoretical comparison of list scanning models. *Journal of Mathematical Psychology*, 1974, in press.

Smith, E. E. Choice reaction time: An analysis of the major theoretical positions. *Psychological Bulletin*, 1968, **69**, 72–110.

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

Sternberg, S. The discovery of processing stages: Extensions of Dander's method. In W. G. Koster (Ed.), *Attention and performance II. Acta Psychologica*, 1969, **36**, 276–315. (a)

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

Suppes, P. Stimulus-sampling theory for a continuum of responses. In K. J. Arrow, S. Karlin, and P. Suppes (Eds.), *Mathematical methods in the social sciences*. Stanford, Calif.: Stanford University Press, 1960.

Thomas, E. A. C. Sufficient conditions for monotone hazard rate: An application to latency-probability curves. *Journal of Mathematical Psychology*, 1971, **8**, 303-332.

Townsend, J. T. A note on the identifiability of parallel and serial processes. *Perception & Psychophysics*, 1971, **10**, 161-163.

Tulving, E. Episodic and semantic memory. In E. Tulving and W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972.

Wescourt, K., & Atkinson, R. C. Memory scanning for information in short- and long-term memory. *Journal of Experimental Psychology*, 1973, **98**, 95-101.

Wilde, D. J. *Optimum seeking methods*. Englewood Cliffs, N.J.: Prentice-Hall, 1964.