

## List length and the time course of recognition in immediate memory

ADAM V. REED

*University of Oregon, Eugene, Oregon 97403*

The response signal method of Reed (1973) was used to study the time-course of list membership recognition after 2 sec of uncontrolled rehearsal, with lists of one, two, and four consonants. Fourteen specific hypotheses about the time course of this process were derived from various theories (Anderson, 1973; Anderson & Bower, 1973; Baddeley & Ecob, 1973; Corballis, Kirby, & Miller, 1972; Kirsner, 1972; Murdock, 1971; Sternberg, 1966, 1969; Theios, Smith, Haviland, Traupmann, & Moy, 1973; etc.) and additional assumptions about the effect of the response signal. When members of the to-be-learned lists are drawn from a small population of highly confusable items, as in the current experiment, list membership recognition appears to follow the model of Theios et al. (1973). Latency functions of signal lag appear to be particularly useful in differentiating among hypotheses which predict similar speed-accuracy tradeoff functions. The lag by positive-negative interaction for latencies of correct responses is highly significant for lists of one consonant, a result predicted by the hypothesis derived from the model of Theios et al. and incompatible with hypotheses derived from exhaustive search and single-threshold strength models.

Consider an experiment in which subjects are asked to decide, on each trial, whether or not a probe item belongs to a list which has just been presented. It is usually found in such experiments that mean reaction times, plotted separately for correct affirmative and negative responses, are linear functions of the number of items in the list. These functions usually have equal slopes, although their intercepts may or may not be equal. This effect was first reported by Sternberg (1966) for digits, and has been replicated since with random plane figures (Briggs & Blaha, 1969), letters (Yio & Santa, 1970), pictures (Klatsky, Juola, & Atkinson, 1971), etc. Several conditions must be met for reliable observation of the Sternberg effect. The items in the original list must be presented slowly enough to insure perfect acquisition (Sternberg presented his lists at a rate of 1.2 sec per item), an interval of several seconds must elapse between the presentation of the list and the probe, and subjects must be allowed to rehearse the list during this interval without being distracted by any other task. The Sternberg effect is observed whenever these

conditions are met (Bracey, 1969; Briggs & Blaha, 1969; Burrows & Murdock, 1969; Hoving, Morin, & Konick, 1970; Wingfeld & Bolt, 1970; Wingfeld & Branca, 1970; etc.).

Sternberg (1966, 1969) explains his findings by postulating an internal scan of the stored list. In order to account for the equal slopes of positive and negative reaction time (RT) functions of the number of items, Sternberg's model postulates that this scan is exhaustive; that is, even if the probe is encountered, the other items in the list will still be scanned before a decision is made. This model also accounts for the absence of serial position effects on RT in the Sternberg paradigm. The scan rate is assumed to equal the per item slope of the RT function, with digits or consonants, between 30 and 40 msec per item. The model is simple, although its exhaustive feature seems somewhat counterintuitive.

While Sternberg's experimental findings have stood the test of subsequent replications remarkably well, his explanation of these findings has not received universal acceptance, and several alternative mechanisms have been proposed for the Sternberg effect. These alternative proposals include differential encoding readiness (Kirsner, 1972), several strength mechanisms (J. A. Anderson, 1973; Baddeley & Ecob, 1973; Corballis, Kirby, & Miller, 1972), slowest-of-several parallel decision processes (Murdock, 1971), exhaustive search of a multidimensional semantic network (J. R. Anderson & Bower, 1973), self-terminating scan of a stochastically organized stack containing negative as well as positive items (Theios, Smith, Haviland, Traupmann, & Moy, 1973), and several additional models which have been discussed informally among workers in this field, but

This article is based on a dissertation submitted to the Graduate School of the University of Oregon, Eugene, in partial satisfaction of the requirements for the PhD degree. The research was supported by a National Science Foundation graduate fellowship to the author and by NIE Grant 3-0097 to W. A. Wickelgren. The manuscript was prepared for publication while the author was at Rockefeller University, with support from NIH Training Grant GM-01789. The author wishes to express his gratitude to Professor Wickelgren for advice and supervision, and to members of his dissertation committee, and other colleagues at Oregon and elsewhere, for suggestions which contributed to the present form of the article. Requests for reprints should be sent to Adam V. Reed, Mathematical Psychology Laboratory, The Rockefeller University, New York, New York 10021.

were never published. All of the above models were designed to account for the same set of RT findings. Thus, it is hardly surprising that the various models predict similar changes in RTs under many manipulations, and that differential predictions about RTs have proven difficult to derive. Even the most promising manipulation, that of stimulus probability (e.g., Miller & Pachella, 1973), appears to have only a limited utility in discriminating among the various mechanisms proposed to explain the Sternberg effect.

An alternative approach to the problem of differentiating experimentally among the various explanations of the Sternberg effect is to study the entire time course of the recognition process, making processing time an independent variable controlled by the experimenter. The accuracy of the subject's responses, as measured by percent correct, information transmitted, or signal-to-noise ratio, etc., can then be plotted for the several list lengths as a function of processing time. Since the different theories of the Sternberg effect propose different mechanisms, their predictions about the time course of the recognition process are usually different, and so are the accuracy functions derived from each. This approach to the problem was first attempted experimentally by Pachella (Note 1), who studied the accuracy of recognition under different deadlines. Unfortunately, the deadline method leaves much to be desired as a method of studying the time course of psychological processes. Its most serious shortcoming is that the subjects must be informed of the deadline in advance, and thus are enabled to adjust the strategy of the retrieval process to the deadline. It is therefore possible that accuracy curves obtained with the deadline method represent not increasing accuracy in the course of a specific process (or mix of processes), but rather terminal accuracies of several different processes or mixtures. In particular, one cannot discount the possibility that fast guess strategies (Ollman, 1966; Yellott, 1967, 1971) are used with some frequency when subjects are faced with the shortest deadlines. The presence of fast random guesses has been confirmed empirically in some visual discrimination experiments (Swenson, 1972), and may well be elicited with short deadlines in recognition memory experiments as well.

An alternative method of studying the time course of psychological processes was developed by Reed (1973). In the response-signal method of Reed (1973), there is a lag between the onset of the probe and a signal instructing the subject to respond. The subjects are asked to begin thinking (about whether the probe was more likely to have been or not to have been a member of the latest list) as soon as the probe appears; and to make a decision and respond as fast as they can after the response signal. Since the subject need not be told in advance what the response-signal lag is going to be on any trial, there can be no

systematic change in initial strategy as a function of the (as yet unknown) lag. There is no advantage to be gained by using a fast-guess strategy.

An additional advantage of the response signal method of Reed (1973) is the availability of data on latency of response (i.e., time since the response signal) as a function of response signal lag. If one assumes that this latency is accelerated when the subject is most ready to respond, the latency curve may be used as a measure, independent of response accuracy, of the subject's relative readiness to respond after a particular lag. A continuing process, with ever-increasing readiness to respond, would result in a monotonic decrease in latency with increasing signal lag. A local minimum in the lag latency curve, on the other hand, might result if a discrete process has just arrived at an answer. The latency curves are most useful in studying differential timing of processes underlying affirmative and negative responses. Because of bias effects, the accuracy of positive and negative responses often cannot be determined separately; and accuracy measures such as transmitted information or  $d'$  are a function of both positive and negative answers. But separate latency curves can be obtained for positive and negative answers, even if bias effects favoring one or the other are present.

Inherent in the response-signal method is the possibility that psychological processes taking place in the response-signal situation differ from those taking place in the absence of response signals, either because response signals interfere in some way with the process under investigation or because the subject adopts a different strategy under which processing is more easily interrupted or less susceptible to response-signal interference. Because of this possibility, it is advisable to perform, in addition to the response-signal experiment, a conventional RT control experiment using stimulus conditions similar to those used with response signals. Such a conventional RT control experiment will yield both reaction-time and accuracy data for each stimulus condition, and these data may be plotted, together with data from the response-signal experiment, in the time-accuracy plane. If the informational processes taking place in the response-signal and conventional-RT situations are sufficiently similar to yield the same speed-accuracy curves, then the time-accuracy points from the conventional-RT experiment will lie on the empirical speed-accuracy curves obtained with the response-signal method. If, on the other hand, the internal processes taking place in the two situations are different, then the conventional-RT time-accuracy points need have no particular relationship to the response-signal speed-accuracy curves. It is, of course, possible that the conventional-RT points might fall on these curves accidentally, but such accidents are relatively unlikely, since the immediate

vicinity of each speed-accuracy curve comprises only a small part of the relevant region of the time-accuracy plane. And, of course, simultaneous accidents causing all of the conventional-RT points to fall on the corresponding speed-accuracy curves would be even more unlikely. When time-accuracy points resulting from conventional-RT experiments fall on speed-accuracy curves obtained in response-signal experiments, the most likely explanation is that the same (or at least highly similar) informational processes are taking place in both situations.

In the current experiment, the response signal method was used to differentiate among 14 specific hypotheses about the time course of recognition in immediate memory, which were derived from the various models of the Sternberg effect. The several models and hypotheses, and their respective predictions, are discussed below.

### MODELS AND HYPOTHESES

In this section, the several theories of the Sternberg effect will be used to derive the 14 specific hypotheses tested in the current experiment. The correspondence between theoretical models and the 14 hypotheses is not one-to-one: some theories lead to more than one hypothesis, depending on additional assumptions about the effect of the response signal on the process proposed in the model; in other cases, several models make predictions which are not distinguishable experimentally in the context of the present study, converging on the same set of hypotheses. For this reason, the section will be alternately concerned with theoretical models, or types of models, and with experimental hypotheses, or types of hypotheses, derived from these models. The several theoretical models of the list-recognition process (and the specific, numbered hypotheses about the expected outcome of the present experiment derived from these models) may be classified into three rough groups: exhaustive serial process models (Group 1: Hypotheses 1-3); direct access and parallel process models (Group 2: Hypotheses 4-11); and self-terminating serial process models (Group 3: Hypotheses 12-14).

The predictions of most of the 14 hypotheses are formulated in terms of the speed-accuracy tradeoff curves. The most striking differences among the hypotheses are in the predicted changes, with list length, in three parameters of the speed-accuracy curve: its time intercept, its rate of growth, and its asymptote. The time intercept, corresponding to the moment when responses can first be made with above-chance accuracy, is the point on the abscissa from which the speed-accuracy curve begins to rise. The growth rate is the inverse of the "time constant," i.e., the time necessary to reach a predetermined

fraction of final accuracy. The accuracy asymptote is the level of accuracy toward which the speed-accuracy curve tends with very long lags between the onset of the probe and the signal to respond.

In the discussion of particular hypotheses below, parameters such as scan rates are treated as though they were deterministic rather than random. It is likely that in reality these parameters are random variables, and the hypotheses below were originally derived in random-variable terms. However, the extent of the empirically measured variability in these parameters (see results) was found to be too small to affect the differential predictions of the several hypotheses, or to change the shape of the predicted time-accuracy plots in any way other than rounding out some of the sharp corners. Random-variable descriptions often obscured the differences among the various hypotheses and therefore the random-variable aspects were for the most part eliminated from the hypothesis descriptions below. Readers who prefer to think in terms of random variables are welcome to generate random-variable counterparts of these descriptions, keeping in mind the relatively small effects of actual variability (see Table 1).

#### Group 1: Exhaustive Serial Process Models

##### Exhaustive Scan Model

The first model of the Sternberg effect was Sternberg's own (1966, 1969) exhaustive scan model. According to this model, the subject stored the to-be-learned items in a set of linked internal locations. During retrieval, the probe would be compared in succession to all items stored in locations corresponding to the most recent list. This process was assumed to be exhaustive, i.e., to proceed even after a match had been obtained, until the probe had been compared with all the items in the list. Only then would a decision be made—affirmative if a match had been found, negative otherwise—and the corresponding response executed.

If the decision process which follows scanning were assumed to be completely deterministic, scanning theories would predict that subjects in scanning experiments should never execute an incorrect response, as long as the original list has been correctly acquired and preserved. This is clearly not the case, and it seems more reasonable to assume that the decision is made on the basis of a familiarity variable, the distribution of which changes when a match is found. Under this assumption (and additional assumptions about the effect of the signal to respond), one can make predictions about  $d'$ , the distance between the means of the two familiarity distributions, as a function of the lag between presentation of the probe and the signal to respond.

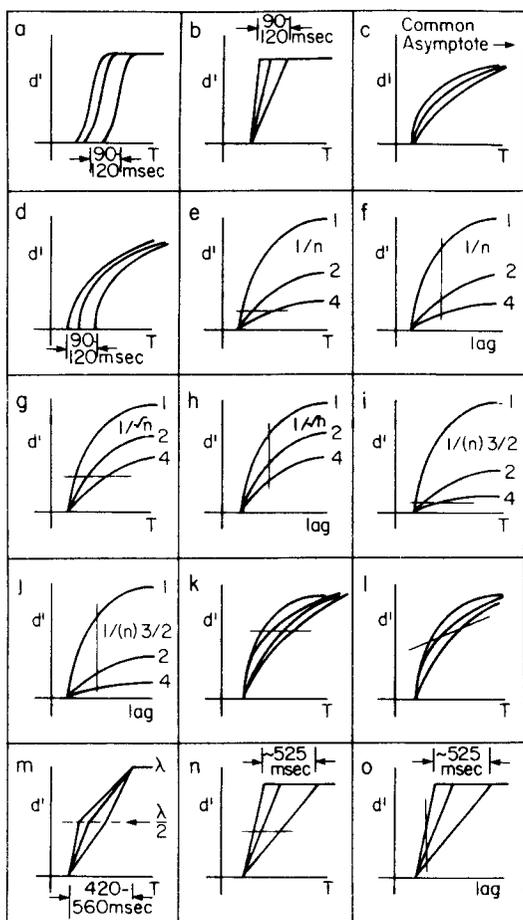


Figure 1. A schematic representation of the speed-accuracy functions predicted by the 14 hypotheses. Panels a through j, speed-accuracy curves predicted by Hypotheses 1 through 10. Panel k, the speed-accuracy curves of the individual parallel processes postulated by the model of Hypothesis 11. Panel l, the overall speed-accuracy curves predicted by Hypothesis 11. Panels m through o, speed-accuracy curves predicted by Hypotheses 12 through 14. Thin lines represent the expected locations of the operating points obtained in a conventional RT experiment.

The three different assumptions one may make about the effect of the response signal under this model lead to Hypotheses 1, 2, and 3 below.

**Exhaustive Search Models**

The exhaustive search model of Anderson and Bower (1973) will be mentioned at this point because the hypotheses about the behavior of the subject (as a function of lag or processing time) which can be derived from this model do not differ measurably from those which may be developed from the Sternberg model (above). Anderson and Bower postulate a deterministic search through an associative network representing list membership relationships in memory, instead of the linear scan proposed by Sternberg. However, in order to predict identical slopes of RT with set size for positive and negative probes, Anderson and Bower are compelled

to assume, like Sternberg, that the search is exhaustive rather than self-terminating (p. 376). Their time-to-search function is a polynomial which approximates Sternberg's straight line very closely over the interval in question (p. 375). While Hypotheses 1 and 2 below—uninterruptible search or scan—are equally compatible with the Sternberg and the Anderson and Bower models, Hypothesis 3—continuing scan—does not appear to be compatible with the Anderson and Bower model, since a search is not likely to restart after its object has been located.

**Hypothesis 1: Uninterruptible Exhaustive Search or Scan**

The simplest assumption about the effect of a signal to respond on a subject engaged in an exhaustive search or scan is that no information will be available before the search or scan is completed, and full information thereafter. Thus, except for random fluctuations in decision speed, the time-accuracy curve will be essentially a step function. The time at which this step takes place should increase by one Sternbergian scan time for each additional item in the list. According to Sternberg, the scan time per item should be equal to the slope of the RT vs. number-of-items function, i.e., between 30 and 40 msec per item. Thus the horizontal distance between the one-item and four-item time-accuracy functions should be between 90 and 120 msec, the curves being otherwise identical (Figure 1a). Although the speed-accuracy functions predicted by this hypothesis need not have true time intercepts, the predicted horizontal displacement is most similar to an intercept increment. No rate or asymptote changes with list length are predicted.

**Hypothesis 2: Interruptible Fast Scan or Search**

An alternative assumption is that partial information will be available to the decision process when an early response signal forces the decision to occur before the scan or search has terminated. The familiarity variable will be distributed as in case of a match if a match has been obtained already, but not otherwise. Thus, the rise of  $d'$  will be linear, from the time it begins to rise above chance until it reaches its terminal value  $n$  scan times later. If Sternbergian scan times (or, equivalently, Anderson and Bower search times) are assumed, then the time-accuracy curve for one-item lists ought to reach its terminal value 30-40 msec after it leaves the abscissa, and the time-accuracy curve for four-item lists should join it 90-120 msec after that (Figure 1b). This hypothesis predicts a slower growth rate with increasing list length, with the asymptote and intercept remaining constant.

**Hypothesis 3: Continuing Fast Scan**

Another alternative is to assume that if the subject

is waiting for an external response signal, the scan will not stop on reaching the end of the list, but instead will continue, increasing the signal-to-noise ratio on successive scans. Letting  $\lambda$  be the linear amplitude of the retrieved trace,  $\sigma^2$  the associated storage noise variance,  $s^2$  the variance of retrieval noise, the power signal-to-noise ratio,  $(s/n)_0$ , after  $m$  scans should equal:

$$(s/n)_0 = \frac{m^2 \lambda^2}{m^2 \sigma^2 + m s^2} = \frac{\lambda^2}{\sigma^2 + \frac{s^2}{m}}$$

But since:

$$m = \frac{(T - \eta)}{n t_s},$$

where  $t_s$  is a single scan time and  $n$  is the number of items being scanned, and  $\eta$  is the time intercept of the accuracy curve,

$$(s/n)_0 = \frac{\lambda^2}{\sigma^2 + \frac{s^2 n t_s}{(T - \eta)}}$$

we may, without loss of generality, choose units so that  $s^2 = 1/t_s$ ; and  $d'$ , the linear signal-to-noise ratio, which is simply the square root of  $(s/n)_0$ , equals

$$d' = \frac{\lambda}{[\sigma^2 + n/(T - \eta)]^{1/2}}$$

The curves generated by this equation for  $n = 1, 2,$  and  $4$ , with typical values of  $\lambda$  and  $\sigma$ , are shown in Figure 1c. With increasing list length, Hypothesis 3, like Hypothesis 2, predicts a difference in the rate parameter only.

### Group 2:

#### Direct-Access and Parallel Process Models

##### Differential Encoding Readiness Model

Kirsner (1972) suggested that all or part of the increase in RT with list size in the Sternberg paradigm may be due to differences in readiness to encode particular probes, and consequent differences in the speed of encoding, rather than be due entirely to differences in the speed of retrieval processes. Kirsner (1972) cites his observation that the latency of probe naming increases with list length, for both words and letters, in support of this possibility.

##### Hypothesis 4: Encoding Time Differential

When accuracy of response is plotted as a function of available time in an experiment using the

response-signal method, encoding time is a part of the minimum amount of time which must elapse before the accuracy can rise above chance. This minimum is the time intercept of the time-accuracy curve. According to the Kirsner hypothesis, it is this intercept which should increase when one increases the size of the learned list. This prediction is shown schematically in Figure 1d. It should be noted that this hypothesis says nothing about the shape of the time-accuracy curve past the time intercept, so that Kirsner's mechanism is not incompatible with any of the other mechanisms discussed in this section. In general, several of the suggested mechanisms may always be operating at once, each accounting for part of the growth of RT with length of the to-be-learned list.

##### Strength Models

In the context of the present study, the term "strength model" is used to refer to those models of the recognition process which postulate that the only change in the accuracy curve, as a result of increasing the number of items in the to-be-remembered list, is a decrease in the terminal accuracy level, or "strength." Thus, other things being equal, the time-accuracy curve obtained with any number of  $n$  of items in the to-be-remembered list ought to be some multiple of the time-accuracy curve obtained with a single to-be-remembered item:

$$d'(n, T) = f(n) d'(1, T)$$

The first study to introduce changes in terminal strength as an explanation of the Sternberg effect was that of Corballis, Kirby, and Miller (1972). Corballis et al. invoked strength effects to explain the finding of primacy and recency effects in their study. It may be argued that Corballis et al. violated one of the preconditions of the Sternberg effect, since they presented their lists at a rate of 300 msec per item (vs. 1,200 msec in Sternberg's experiments) and allowed only 600 msec more (vs. 2,000) before presenting the probe. Thus, it is possible that their subjects simply did not have time to organize a list for scanning. However, their explanation could be applied to the results of experiments in which Sternberg's conditions were met as well as to those where they were not. Corballis et al. did not provide details about the relative terminal strengths and the mechanisms responsible for transforming differences in strength into differences in RT, beyond suggesting that RT decreased with increasing strength, while strength decreased with increases in the number of items in the to-be-remembered list.

One possible mechanism for the inverse relationship between list length and strength was suggested by Baddeley and Ecob (1973). (Baddeley and Ecob actually postulated two models, the second

of which combined attentional capacity limitations with a response mechanism based on the Hick-Hyman law. Unfortunately, their second model does not provide any experimental predictions which would permit one to differentiate it empirically from strength models. As the authors noted, their second model needs further elaboration before it can be adequately evaluated.) In their Model 1, Baddeley and Ecob suggest that a linear relationship between RT and set size could be accounted for by postulating that the total amount of strength available to the system during storage is limited, and must be divided among the items in the positive set. Baddeley and Ecob also postulate a specific quantitative relationship between strength and RT, but no specific mechanism. Thus, the key prediction from their Model 1 is that terminal strength on retrieval should be inversely proportional to the size of the positive set.

Another possible focus of an inverse relationship between strength and the size of the positive set is during the retrieval process. Reed (Note 2) postulated that recognition decisions were based on a variable resulting from the integration of a stored signal over time. If one assumes a simple exponential leaky integrator (Model 100 of Reed, Note 2), with a fixed integrating capacity shared among the several list items, one may derive the expression

$$(d')^2 \sim \frac{1}{n} \cdot \frac{1 - e^{-\beta(T-\eta)}}{1 + e^{-\beta(T-\eta)}}$$

from which it follows that final strength, as measured by  $d'$ , would be inversely proportional to the square root of positive set size.

An inverse square root relationship between strength and set size also results from the discrete retrieval mechanism postulated by J. A. Anderson (1973). The latter model predicts a slightly different  $d'$  vs. time curve: instead of a continuous function tending smoothly to an asymptote, Anderson's model predicts that  $d'$  should be a staircase function, approximately proportional to square root of  $(T - \eta)/n$ , until the summing mechanism reaches the end of a finite set of memory cells, and constant thereafter. In principle, one could differentiate between the models of J. A. Anderson (1973) and Reed (Note 2) by means of time-accuracy curves obtained with the response-signal method. However, the author was not aware of the Anderson model when designing the present experiment, and the design does not have sufficient power to differentiate between the two models. For this reason, separate hypotheses are not derived from the two.

A third set of hypotheses about the relationship between strength and positive set size results from combining the storage model of Baddeley and Ecob (1973) with the retrieval model of Reed (Note 2). The

two effects are multiplicative, and the combined model predicts terminal  $d'$  levels inversely proportional to the 3/2 power of set size.

J. A. Anderson (1973) also spells out one of the possible mechanisms by which different terminal strengths might cause RT differences. This mechanism postulates that a response is executed when enough time has elapsed to permit responses to be made with a preset minimum accuracy. Time-accuracy curves building up to a higher terminal strength will reach this preset threshold earlier than curves building up to a lower one. The resulting relationships are clearly monotonic, and with appropriate assumptions about the shape of the curves, such as those of J. A. Anderson (1973), a linear relationship between RT and positive set size may be predicted. Even in those cases where the relationship predicted from a particular assumption about the shape of the time-accuracy curves is not linear, it usually tends to approximate a linear relationship over small ranges of  $n$  as a result of analytic continuity. These thresholds are shown as horizontal lines in Figure 1 panels corresponding to hypotheses which use the threshold mechanism.

An alternative mechanism which might mediate relationships between strength and RT is the confidence mechanism of Norman and Wickelgren (1969). Norman and Wickelgren (1969) demonstrated that more confident responses were faster than less confident ones, and that confidence and response latency were well enough correlated for operating characteristics to be plotted using latency measures instead of confidence ratings. Thus, if the retrieval itself were stopped after a fixed amount of time regardless of the size of the positive set, reaction times could still be slower with larger set sizes because of a decrease in strength and a corresponding decrease in confidence. This implies that in the response signal situation, the latency of the response following the decision might be a function of its accuracy and the resulting confidence level. For this reason, the speed-accuracy curves for the confidence-mediation hypotheses are plotted not as a function of total response time (signal lag plus response latency), but as a function of lag alone. Consider, moreover, the set of points at which these speed-accuracy functions are intersected by horizontal lines representing the accuracies obtained in the spontaneous RT section of the current experiment. If spontaneous RT accuracies are the result of retrieval processes which stop after a fixed amount of time regardless of the size of the positive set, as proposed by the confidence mediation theory, then the values of the abscissa at these intersection points should all be the same, i.e., they should lie on the same vertical line. These vertical lines are shown in the panels of Figure 1 corresponding to hypotheses which use a confidence-latency mechanism. The confidence mediation

hypotheses also predict that when an analysis of variance is performed on reaction times as a function of both rated confidence and positive set size, the effect of confidence should exceed the effect of set size.

### **Hypotheses 5 Through 10: Strength Effect Mechanism**

The three models of the variation of terminal strength with the size of the positive set may be combined with the two models of strength-RT mediation to form six hypotheses about the results of varying list length in a study using the response-signal method and a control spontaneous RT condition. These are  $1/n$  terminal strength with threshold mediation (Hypothesis 5),  $1/n$  terminal strength with confidence mediation (Hypothesis 6),  $1/\sqrt{n}$  terminal strength with threshold mediation (Hypothesis 7),  $1/\sqrt{n}$  terminal strength with confidence mediation (Hypothesis 8),  $1/(n)^{3/2}$  terminal strength with threshold mediation (Hypothesis 9), and  $1/(n)^{3/2}$  terminal strength with confidence mediation (Hypothesis 10). Figures 1e through 1j show schematically the accuracy curves and spontaneous accuracy predicted by the six strength hypotheses.

### **Parallel Process Model**

Murdock (1971) proposed an explanation of the growth of RT with the size of the positive set in terms of a set of parallel comparison processes, proceeding simultaneously but at different speeds. Under this model, a spontaneous response would not take place until the slowest of these processes reached a preset level of accuracy. When the to-be-learned list contains one item, only the fastest of these processes is brought into use; with two items, the fastest and the second fastest, etc.

### **Hypothesis 11: Slowest of $n$ Parallel Processes**

By the time the slowest of the several comparison processes reaches a fixed preset accuracy threshold, the other processes will have reached an even higher level of accuracy. Thus, if the individual comparison processes followed the time-accuracy curves proposed by Murdock (1971) (Figure 1k), the average time-accuracy curves would follow a course shown schematically in Figure 1l. Note that this model, unlike other models of the Sternberg effect, predicts that the accuracy of spontaneous responses will rise monotonically (or, at *minimum*, remain constant) with increasing  $n$ .

### **Group 3:**

#### **Self-Terminating Serial Process Models**

#### **Self-Terminating Scan of Ordered Whole-Population Stack**

Theios, Smith, Haviland, Traupmann, and Moy

(1973) demonstrated that parallel RT slopes with positive set size could be obtained with a self-terminating rather than exhaustive scan, if the scanned list contained all the items in the given population, both positive and negative, and were reordered stochastically whenever the positive set was changed.

According to Theios et al. (1973), the subject maintains throughout the experiment a stack containing entries for *all* the items in the population from which the members of the positive set may be drawn. When presented with a new positive set, the subject pulls its members out of the old stack, and replaces them at the head of the stack after attaching to them markers which identify these items as positive. Negative markers are then attached to the remaining items in the stack. Theios et al. also make other assumptions which have to do with the order of items in the stack when the different items have different probabilities of being used in the probe position. These latter assumptions are not relevant to the present experiment. When the subject is presented with a probe, the Theios model postulates that he scans this stack, comparing the probe to successive items. When the item matching the probe is found, the positive or negative marker attached to this item is read off, and the corresponding response executed. One may easily show, using elementary algebra, that the slope of RT with the size of the positive set under a Theios scan should equal one-half the scan time per item for both positive and negative responses. Thus, according to the Theios et al. (1973) model, Sternberg effect slopes of 30 to 40 msec per item correspond to scan times of 60 to 80 msec per item. If the items are drawn from a population of seven different items, the time necessary to scan the entire stack may be as high as 560 msec, i.e., considerably more than the 160 msec maximum necessary to scan a positive list of four items under the assumptions of Sternberg (1966, 1969). For this reason, the model of Theios et al. (1973) may be referred to as a slow-scan model.

### **Hypothesis 12: Theios et al. [1973] Scan**

In order to use the model of Theios et al. to derive a specific prediction about the behavior of  $d'$  as a function of available time, it is necessary to postulate a specific effect of marker readout, displacing the distribution by a fixed amount in the positive direction, while the readout of a negative marker results in an equal negative displacement. In an experiment in which the probe is positive half of the time, then, the  $d'$  will reach one-half its terminal value in the time necessary to scan the positive section of the stack, and in the time necessary to scan the rest of the stack it will go the whole way. The time-accuracy curves predicted under these assumptions when the positive sets of one, two, and four items are drawn from a total population of seven, are shown

schematically in Figure 1m. While it is not necessary to assume that the familiarity distribution displacement resulting from marker readout will be independent of positive set size, this is clearly the simplest assumption in this case.

### Rehearsal Buffer Scan

W. A. Wickelgren has suggested to the author (personal communication) yet another model of the effect of list length on processing time and accuracy. According to this model, a cyclical rehearsal buffer containing the positive list is read and compared to the probe as it passes, at the usual rehearsal rate, past a stationary readout. Under assumptions about the effect of comparison on the familiarity distribution similar to the assumptions made in the case of the fast scanning models discussed earlier, this model, like that of Hypothesis 2, predicts a linear rise from zero  $d'$  to its terminal value in the time it takes to scan the positive set. The difference between this model and fast-scan models is in the time scale: the rate of rehearsal is about 175 msec per item (Landauer, 1962), i.e., about five times as long as the time postulated by Sternberg. This model is consistent with the need to permit free rehearsal for the Sternberg effect to be observed; and it has the additional advantage of not requiring any mechanisms beyond the rehearsal buffer, which is already a common feature of many theories of immediate memory.

### Hypotheses 13 and 14

The rehearsal readout model discussed above does not incorporate a specific spontaneous RT mechanism, and an additional mechanism must be postulated to account for the Sternberg effect using the time-accuracy relationship postulated under that model. Either of the two mechanisms discussed in connection with strength theories may be applied for this purpose. The rehearsal readout model combined with a threshold RT mechanism gives rise to Hypothesis 13, shown schematically in Figure 1n; combined with confidence mediation of RT, to Hypothesis 14, shown in Figure 1o. Note that Hypothesis 14, like other confidence-mediation hypotheses, requires the plotting of accuracy as a function of signal lag rather than total time.

## METHOD

### Trial Procedure

Each trial began with the presentation of the to-be-remembered items. The list consonants were displayed on electroluminescent panels, simultaneously, for a total duration of 1.2 sec per item, and the subjects were required to read them out loud during the presentation. This was followed by a period of 2 sec during which the activity of the subjects was not controlled. The probe then appeared, and was followed, after a variable lag, by a 1,000-Hz tone of 100 msec duration, which served as the signal to respond. The probe remained on display for 1 sec after the onset of this signal. The subjects were asked to, "start trying to remember, as soon as a

consonant comes on, whether or not it was one of the consonants which you read on the display at the start of the trial. If, at the moment you hear the beep, you think it more likely that it was, press the 'yes' button; if you think it more likely that it was not, press the 'no' button. Try to press the appropriate button as fast as you can after hearing the beep. Then think back to how confident you felt about your answer as you were giving it. If you felt completely positive, press Button 9 on your card punch; if you felt you were guessing, press '2', if you felt you had pressed the wrong yes-no button, press '1.' In between, use the other numbers: the closer you felt to being positive, the larger the number you should use. Try to use all numbers between 2 and 9 equally often." The console buttons were labeled "←YES- NO→," so that all subjects pressed the "YES" button with their left hands, and the "NO" with their right. The yes-no answers and latencies were measured and recorded by an off-line digital system built for this purpose. The subjects recorded their confidence ratings using manual computer card punches.

### Sessions

Each subject participated in two 210-trial sessions per day, 4 days a week, for a total of 58 sessions per subject. The sessions lasted approximately 70 min each, and were separated by a 20-min break. A University end-of-term break of 3 weeks occurred between Sessions 34 and 35.

### Design

In the main experiment there were 42 factorially combined conditions: two types of probes (positive and negative); three list lengths (one, two, and four items); and seven response signal lags (7, 51, 103, 204, 411, 1,030, and 4,131 msec). In addition, the same subjects participated in a conventional RT control experiment performed under moderate speed emphasis (the subjects were asked to respond "reasonably fast"). In the conventional RT control experiment, there were six conditions (three list lengths; positive and negative probes). The design of each section of the experiment consisted of multiple Position by Session Latin squares, with 210 trials per session and per condition. Two complete sequences were generated, with the second used in checking out the equipment and in training sessions. The second half of each design used the same sequence of lists as the first, a new list being presented on each trial. Lists followed by positive probes in the first half of the design were followed by negative probes in the second, and vice versa. The sessions were presented in the following order: 5 main-experiment training sessions, followed by 21 sessions of the first half of the main experiment, followed by 3 conventional RT control experiment training sessions, followed by the 6 sessions of the conventional RT control experiment, followed by a perceptual control session in which the subjects wrote down stimuli presented on the display panels, followed by a main-experiment retraining session, followed by the 21 sessions of the second half of the main experiment.

### Materials

In order to bring  $d'$  within measurable range, the items used in the experimental lists were selected from the seven highly confusable English consonants, B, C, D, P, V, T, and Z. Each consonant was presented equally often in each list position and as a probe.

### Subjects

The four subjects were University of Oregon undergraduates, female, right-handed, nonsmokers, native speakers of English. In addition to hourly pay, they received a bonus for completing all sessions of the experiment.

### Data Processing

The response and confidence ratings obtained from each subject for each list length and signal lag were used to compute  $d_T$ , an estimate of  $d'$  at the moment of the yes-no response. In the procedure for computing  $d_T$ , the cumulative confidence rating data are used to determine the parameters of a family of ROC curves,

Table 1  
Data Summary

List Length	Lag (Msec)	Geometric Mean $d_T$	Mean Latency (Including Errors)		Number of Responses Counted*	Mean Latency of	
			Latency	SD		Correct Positives	Correct Negatives
1	7	.36	228.69	56.82	1625	222.54	278.49
	51	1.03	215.28	53.65	1609	201.14	255.49
	103	1.54	199.38	49.89	1624	184.16	218.96
	204	2.28	175.23	42.32	1526	172.99	173.64
	511	2.59	188.64	57.68	1430	203.71	172.60
	1033	2.99	191.34	46.82	1585	193.02	187.46
	4104	2.98	184.84	46.09	1610	186.06	181.60
	Conven. RT	1.58	280.21	53.49	1640	260.01	317.20
2	7	.38	236.77	59.09	1610	226.73	288.65
	51	.47†	225.93	55.26	1621	214.62	265.21
	103	.99	210.88	51.18	1628	195.02	229.61
	204	1.59	184.09	42.95	1586	178.53	186.53
	511	2.61	180.25	51.74	1477	187.33	172.99
	1033	2.74	189.34	48.59	1576	187.28	190.47
	4104	2.82	186.78	47.64	1587	186.63	183.11
	Conven. RT	.98	304.73	58.17	1651	288.66	333.32
4	7	.31†	243.16	62.10	1612	238.72	298.75
	51	.51	230.63	59.39	1625	220.59	281.60
	103	.66	224.97	58.59	1631	212.79	250.68
	204	1.02	201.68	48.58	1591	190.19	211.54
	511	2.11	184.67	45.73	1557	182.31	182.53
	1033	2.53	188.86	44.64	1613	184.40	191.29
	4104	2.61	188.30	47.51	1621	187.70	186.28
	Conven. RT	.77	325.32	69.75	1594	308.80	373.54

\*100 msec <  $T_2$  < 500 msec out of 1680.

having a unique member curve through each point in the ROC plane. The  $d_T$  is the  $d'$  of the ROC curve belonging to this family and passing through the yes-no point. A detailed derivation and description of the procedure for computing  $d_T$  is given in Reed (1973) and will not be repeated here. The  $d_T$ s were averaged geometrically over subjects. Geometric, rather than arithmetic, means were used in order to satisfy the constraint that the square of the means be the mean of the squares of  $d_T$ . This constraint was imposed because some of the models used to derive the hypotheses tested in this experiment had been formulated in terms of the power signal-to-noise ratio  $(s/n)_0$ , which corresponds to the square of the linear signal-to-noise ratio,  $d'$ , measured by  $d_T$ . Because logarithmic averaging tends to exaggerate decremental errors for low values of  $d_T$ ,  $d_T$  values less than .2 were replaced by the value of .2 before averaging. Averages which include such values are marked in the tables by a dagger. In order to eliminate trials on which anticipation responses occurred, data from responses with latencies less than 100 msec (measured from the onset of the signal to respond) were discarded. In order to eliminate trials on which a subject missed the response signal and initiated her response on hearing the others respond, data from responses with latencies exceeding 500 msec were similarly discarded. Latencies were averaged linearly.

Theoretical curves were fitted to  $d_T$  data using an iterative hillclimbing algorithm which minimizes squared logarithmic error. Estimates of variance fraction accounted for,  $r^2$ , were adjusted for the number of free parameters, using the formula:

$$r^2 = 1 - \frac{\sum_{i=1}^h (\ln x_i - \ln \hat{x}_i)^2 / (h - k)}{\sum_{i=1}^h (\ln x_i - \ln \bar{X})^2 / (h - 1)}$$

where  $h$  is the number of empirical points  $x_i$ ,  $k$  is the number of free parameters in the theoretical function,  $\hat{x}_i$  are the theoretical values corresponding to  $x_i$ , and  $\bar{X}$  is the grand mean of  $x_i$ .

## RESULTS AND DISCUSSION

Before considering those aspects of the experimental results potentially relevant to discrimination among the several hypotheses outlined in the first half of this article, it is necessary to check the assumptions which might underlie the claim of their relevance. These assumptions are that (1) the use of a highly confusable stimulus set, while increasing the error rate sufficiently to insure that accuracy is readily measurable, does not eliminate the Sternberg phenomenon of linear growth of reaction time with list length; and (2) that there is reason to believe that the informational processes taking place in the response signal situation are identical, or at least highly similar, to those taking place, with identical stimuli, in a conventional RT situation. The data from the experiment are summarized in Table 1. The  $d'$  does not exceed 2.99 in any condition, corresponding in the symmetrical case to a minimum error rate of .067. This means that with 1,680 trials per condition, the relative error in the measurement of error rate is never greater than 9%, i.e., the use of highly confusable stimuli did increase the error rate sufficiently to insure adequate measurability. The mean reaction times of correct

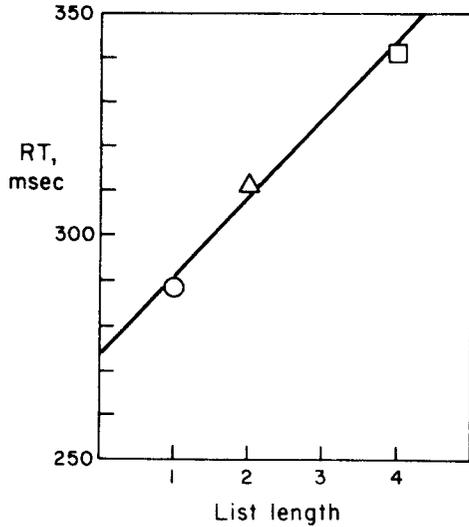


Figure 2. Average reaction time of correct responses in the conventional RT control experiment, as a function of list length. See Table 1 for a more detailed breakdown of data.

responses in the conventional RT control experiment are shown, as a function of list length, in Figure 2. The approximately linear growth of RT with list length is evident, the slight curvature having the same direction as in Sternberg (1966). The slope, 18 msec per item, is lower than in most experiments dealing with the Sternberg effect—possibly a consequence of speed emphasis, in place of the usual accuracy emphasis, in the instructions.

In order to check Assumption 2, the time-accuracy data from the conventional RT control experiment were plotted (as filled points) in the time-accuracy plane, together with the speed-accuracy curves obtained using the response signal method, as shown in Figure 4. It is clear that these control points lie quite close to the corresponding response-signal curves for all three list lengths. As was pointed out in the introduction, such a result is relatively unlikely if the informational processes taking place in the response-signal situation and in the conventional-RT situation are different. A crude test of the statistical significance of the finding that the conventional-RT control points lie closer to the corresponding time-accuracy curves than would be expected by chance if they actually lay on different time-accuracy functions may be made by comparing their deviation from these curves with the deviations of nearby points known to lie on different curves. The deviations of the response-signal points bracketing each conventional-RT point from time-accuracy functions (in Figure 4) other than those fitted to these points were used for this purpose. The F ratio is 43.2 (12,3);  $p < .006$ . In the absence of evidence to the contrary, the simplest explanation of the finding that the conventional-RT time-accuracy points lie on corresponding response-signal speed-accuracy curves is that the same

informational processes take place in both conditions.

**Temporal Variability**

The use of the deterministic form of the several theoretical hypotheses is justifiable only if (1) for each list length, the portion of the total variability due to components prior to the response signal is at least an order of magnitude smaller than the rise time of the corresponding speed-accuracy curve, and (2) the increase in variability with increasing list length is similarly small in comparison with the temporal differences between curves. A conservative estimate of the pre-response-signal portion of total variability for each list length may be obtained by subtracting the minimum post-response-signal latency variance from the total RT variance in the conventional-RT control condition. The square roots of these residual variances are 32.71, 39.23, and 53.59 msec for lists of one, two, and four items, as compared to speed-accuracy curve rise times of the order of 1/2 sec. The increase in variability is approximately 7 msec/item, compared to temporal separation of about 67 msec/item. Thus, the measurements of temporal variability appear to meet both preconditions for the use of deterministic approximations in discriminating among the 14 hypotheses described earlier.

**Time-Accuracy Curves**

The geometric mean values of  $d_T$  for the different list lengths are plotted against signal lag, in Figure 3, and against the total time between the onset of the probe and the execution of the response, in Figure 4. The  $d_T$  values for individual subjects are shown, as a

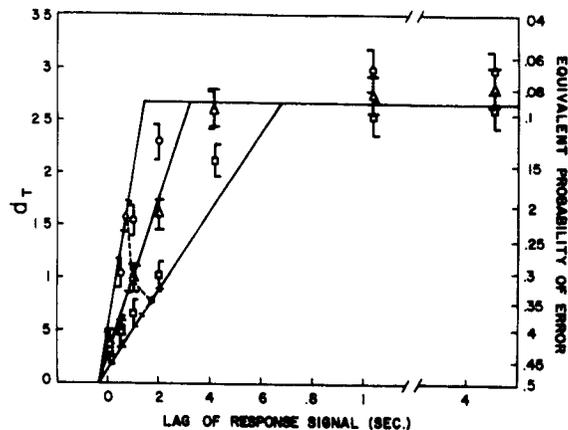


Figure 3. Response accuracy as a function of response signal lag. Circles, one-item list data; triangles, two-item list data; squares, four-item list data. Brackets represent two standard errors, calculated from equivalent error rates using the  $\sqrt{p(1-p)/n}$  standard error in error rate formula. Solid lines represent the best-fitting functions of the form predicted by the rehearsal scan with confidence RT mediation hypothesis (Hypothesis 14). Diamonds connected by dashed lines represent the intersections of fitted functions with accuracy levels of conventional RT responses.

function of total time, in Figure 5. These results permit us to discard all but 3 of our 14 initial hypotheses. Consider first the 6 hypotheses derived from strength models. These predict that terminal levels of  $d'$  for lists of length one and length four should be in the ratio of 4:1 (Hypotheses 5 and 6), 2:1 (Hypotheses 7 and 8), or 8:1 (Hypotheses 9 and 10). In other words, even the most conservative strength model predicts that the terminal value of  $d'$  for list length one should exceed the terminal value of  $d'$  for list length four by 100% of the latter. The results of the present experiment, however, show a measured difference of only 14%, i.e., terminal strength appears to vary little, if at all, as a function of list length. Hypothesis 4, based on postulated differences in encoding time (Kirsner, 1972), is contradicted by the fact that the extrapolated time intercepts of the three accuracy curves are very nearly the same, their small differences being opposite in direction to that predicted by the encoding time model. Kirsner's (1972) finding that naming latency increases with list size must be explained, therefore, by an effect on some component other than encoding time—possibly, e.g., a priming of the naming response.

The spontaneous response points are clearly not horizontal. Their accuracy decreases with increasing list length, approaching the vertical when plotted on lag-accuracy curves. Thus, the two remaining hypotheses which postulate a threshold RT mechanism (Hypotheses 11, based on the parallel model of Murdock [1971], and 13, in which a threshold RT mechanism was postulated in conjunction with rehearsal scan) are eliminated by the spontaneous RT control data. The three curves clearly

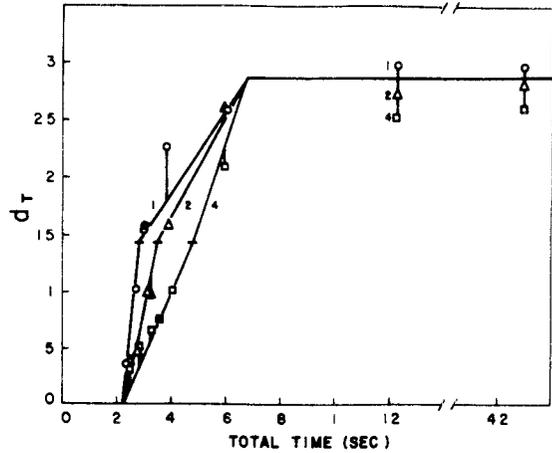


Figure 4. Response accuracies plotted against total time between the onset of the probe and the execution of the response. Circles, data for one-item lists; triangles, data for two-item lists; squares, data for four-item lists; open symbols, data from response-signals experiment; filled symbols, data from a conventional-RT control experiment. In the response signal experiment, the differences between positive and negative response latencies were not consistent in direction, and total time was computed by adding signal lag to mean response latency from onset of signal. In the conventional RT experiment, the mean latencies of negative responses consistently exceeded the mean latencies of positive responses. Because the measured accuracy is not reached until the negative responses have been executed, the total time used for conventional-RT experiment points is the mean latency of negative responses. The conventional-RT operating points appear to be located on the speed-accuracy functions generated in the response signal experiment, indicating that the processing of response signals did not interfere significantly with the main task. Solid lines represent the best fitting functions of the form predicted by Hypothesis 12, which was derived from the model of Theios, Smith, Haviland, Traupmann, and Moy (1973).

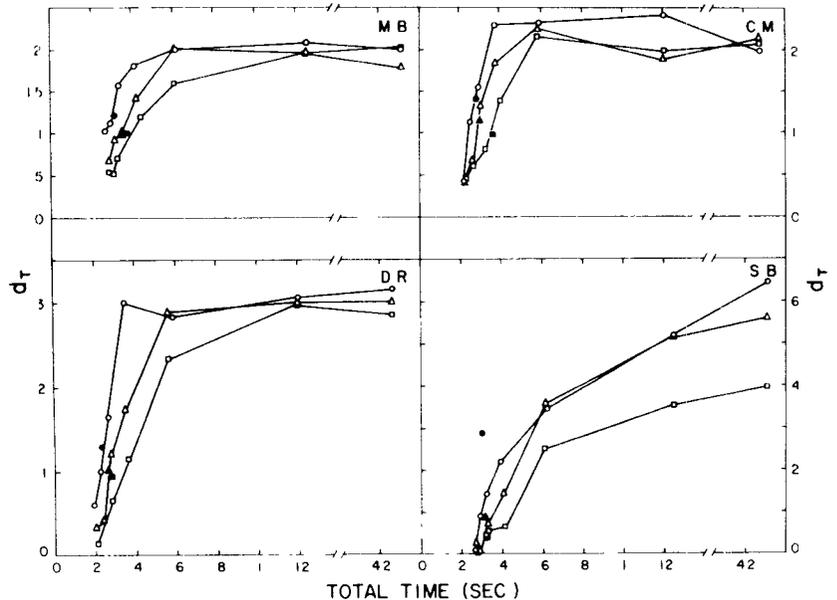


Figure 5. Response accuracies for the individual subjects, plotted in the same manner as the geometrically averaged accuracies in Figure 3.

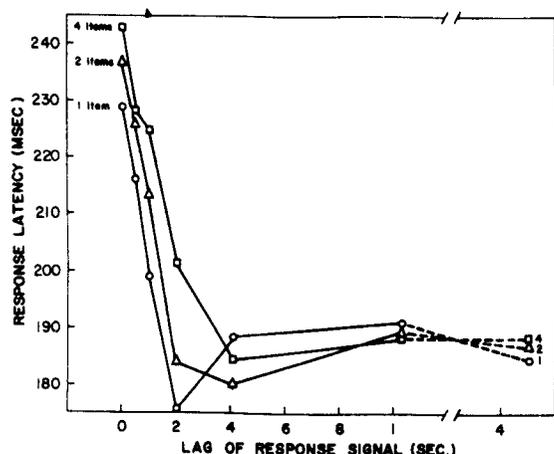


Figure 6. Mean response latencies from onset of the response signal, as a function of response signal lag after the onset of the probe. The data are plotted separately for the different list lengths.

have different shapes, rather than the same shape horizontally displaced, eliminating Hypothesis 1 (the obligatory exhaustive scan hypothesis). Finally, a curve-fitting test was performed to differentiate between Hypotheses 2 and 14, which postulate similar curves but with different time parameters: Sternbergian scan (Hypothesis 2), with scan times of 30-40 msec per item predicts a maximum difference of 120 msec between arrival of one-item and four-item curves at the terminal level, as against a scan time of 120 msec per item and above, for a minimum total difference of 360 msec, predicted by rehearsal scan (Hypothesis 14). When curves with initial slopes in the ratio of 4:2:1 are fitted to the data, scan times are 178.3 msec per item, for a total difference of 535 msec. Even without the slope ratio constraint, the difference is well above 300 msec. This finding is not consistent with the Sternbergian scan, but does accord with the predictions of the rehearsal scan hypothesis. As was noted earlier, the spontaneous response accuracy levels, when plotted on the lag-accuracy curves, are nearly vertical, i.e., consistent with the confidence mediation feature of Hypothesis 14.

In summary, it appears that all hypotheses except 3, 12, and 14 may be excluded as inconsistent with the accuracy data obtained in the present experiment. The theoretical curves of these three hypotheses were fitted to data, resulting in  $r^2$ s of .935 for Hypothesis 3 (continuing exhaustive scan), .945 for Hypothesis 12 (Theios et al. stack), and .875 for Hypothesis 14 (rehearsal scan with a confidence-effect RT mechanism). The scanning rate per item which best fits the accuracy data under Hypothesis 12 is about 67 msec, i.e., close to the middle of the expected 60- to 80-msec range. The fitted scanning rate for Hypothesis 14, 178 msec per item, is in good agreement with available data on rehearsal rates (e.g., Landauer, 1962, who reported a rate of 175 msec per item).

### Signal Lag-Response Latency Data

The mean response latencies for the three list lengths are shown, as a function of signal lag, in Figure 6. All three curves show local minima at lags of 204 or 411 msec. Such local minima are not readily explainable if a continuing process, such as that postulated under Hypothesis 3, is assumed. Under the self-terminating scan assumption of Hypotheses 12 and 14, on the other hand, these minima lend themselves fairly readily to an explanation, provided it is assumed that response latency is accelerated when the response signal coincides with a time of greater readiness to respond. Such a period of increased readiness to respond would follow the termination of a discrete recognition process, such as the self-terminating scans proposed under Hypothesis 12 (Theios et al. whole-population stack scan) and 14 (rehearsal scan). These two hypotheses are also capable of accounting for the relative depths of the minima observed in Figure 6. The explanation, in terms of both of these hypotheses, is shown schematically in Figure 7. Both self-terminating scan hypotheses predict that terminations, and therefore response accelerations, should be more concentrated with shorter lists. This concentration results in a deeper minimum of average response latency.

Before discussing the differential lag-latency plots for correct negative and affirmative responses, it is necessary to consider a second phenomenon likely to contribute to a decrease in response latency with increasing signal lag: the refractory period effect. This effect is generally monotonic, but the latency of responses made with the dominant hand may dip earlier than that of responses made with the nondominant one. In order to exclude laterality differences in refractory period as the main cause of the relative temporal positions of positive and negative response latency minima, the positive response was

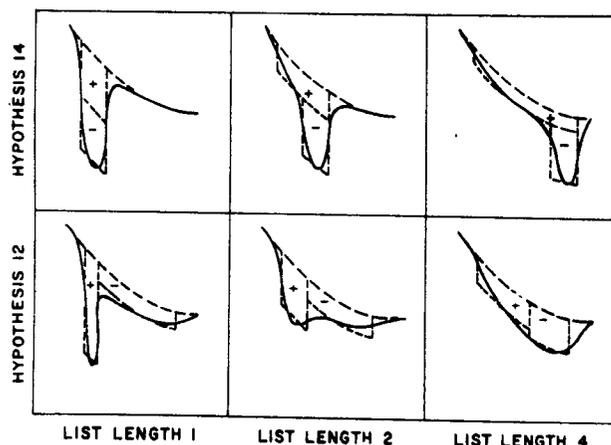


Figure 7. A schematic representation of the average latency predictions derived from Hypotheses 12 and 14. To simplify the derivation, it was assumed that, as a first order approximation, latency accelerations due to decision process termination are uniform and last for one scan cycle.

assigned to the left, i.e., nondominant, hand of the right-handed subjects. All theories which predict that maximum readiness to respond should occur at different times for positive and negative responses, also predict that the positive, i.e., left hand, responses should dip earlier. Laterality differences in refractory period lead to the opposite prediction. As Figure 9 shows, correct positive (left-hand) responses do indeed dip earlier, so that refractory-period effects are not likely to be the sole or main cause of the relative location of observed minima in the lag-latency curves. In Figures 7 and 8, refractory period effects are represented by the monotonically decreasing dashed baseline.

Although Hypotheses 12 and 14 make similar predictions about overall latency as a function of lag, they differ sharply in their predictions about the separately averaged latencies of correct positive and correct negative responses as a function of lag, compared separately for the various list lengths. The two sets of predictions are shown schematically in Figure 8.

According to Hypothesis 14, the scanned rehearsal buffer contains positive list items only. Thus, termination resulting in a negative response will always follow a complete scan, coinciding with positive terminations based on the last item in the buffer to be scanned. Therefore, with list length one, there should be no interaction between signal lag and response type. The interaction should increase with list length. Specifically, the width of the positive response latency dip and the separation between minima should increase, while the width of the negative response latency dip stays the same.

In contrast, Hypothesis 12, based on the model of Theios et al. (1973), assumes scanning of a stack containing both positive items (on top) and negative items. Termination always occurs on encountering the probe item during a scan of the stack. Thus, with list length one, there should be a considerable interaction

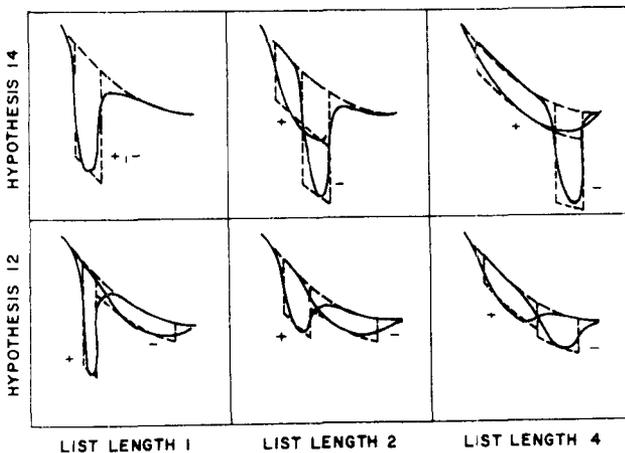


Figure 8. A schematic representation of differential latency curves for positive and negative responses, under the same assumptions as for Figure 6.

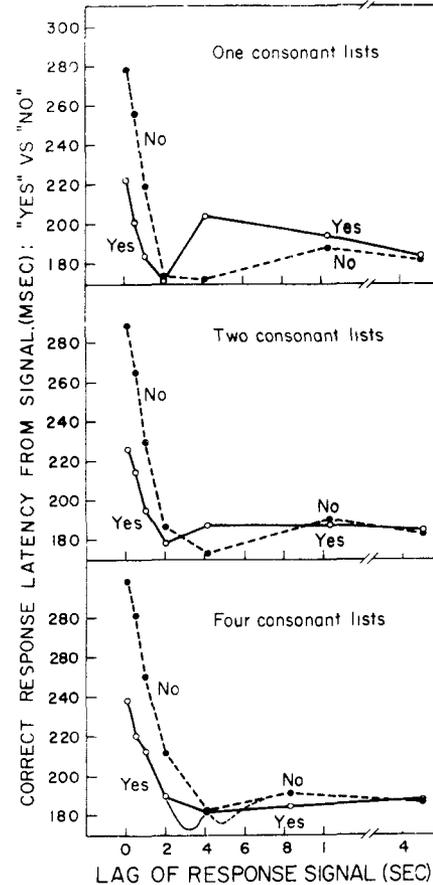


Figure 9. The empirical response latency curves, plotted separately for correct positive and negative responses for each list length. The lower curves in the bottom panel represent alternative extrapolations of the two lag-latency curves between empirical points.

with a positive dip considerably narrower than the negative one and occurring 3.5 item-scan times earlier. The separation between positive and negative response latency minima should stay constant with the list length, the width of the positive dip should increase, and that of the negative dip should decrease. The overall interaction between lag and response type would be expected to decrease as the widths of the two latency dips become more nearly equal. Thus, with respect to the F test on the signal lag-response type of interaction component of latency variance, the two hypotheses have directly opposite predictions. Hypothesis 14 predicts that the F ratio for the interaction will not be significant with list length one, and if the error term does not vary with list length, the F ratio will increase with longer list lengths. Hypothesis 12, on the other hand, predicts that the interaction F ratio will be greatest with list length one and will decrease with increases in the size of the positive set.

The separate comparisons of the correct positive and correct negative curves are shown in Figure 9. The minima for list length one are indeed separated,

Table 2  
F Test on Latency Curves for Correct "Yes" and "No" Responses

Variable	List Length 1		List Length 2		List Length 4		df
	F	p	F	p	F	p	
Yes-No	1.390	n.s.	3.337	n.s.	6.852	< .05	1/6
Signal Lag	2.683	n.s.	4.594	< .05	6.704	< .025	6/6
Interaction	135.866	~.0	104.329	~.0	58.945	~.0	6/8000+

as predicted by the Theios et al. (1973) model; and all the other qualitative predictions of that model are borne out. The signal lag and response type F ratios are shown in Table 2. Because of unequal cell frequencies, Cochran's (1951) F' procedure was used to compute the interaction F ratio. In accord with the Theios et al. (1973) model, and in contradiction to the rehearsal scan hypothesis, the F ratio for the lag-response type of interaction decreases with list length. (There is no noticeable change in the error term) Moreover, the interaction F ratio for list length one is 135.87, with 6 over more than 8000 degrees of freedom. The probability of this F ratio occurring under any hypothesis predicting no interaction for this list length, such as Hypothesis 14, is literally too small to compute with equipment available to the author.

It should be noted that this result is of a very different kind than the positive-negative response differences measured with spontaneous RT methods. Positive and negative responses are necessarily different, so that any overall RT differences could always be ascribed to response execution time rather than to cognitive timing differences. In the present experiment, however, any differences in response execution time would be reflected in response type main effect F ratios rather than response type/signal lag interaction F ratios. The response-type main effect F ratios are, however, very small compared to the interaction F ratios: not significant for list lengths one and two, and barely significant at the .05 level for list length four. Thus, the present result makes untenable any theory which predicts that, under conditions obtaining for lists of length one in the present experiment, positive and negative recognition processes should terminate at the same time. Among theories incompatible with this result are single-threshold strength theories, exhaustive scanning theories, and theories postulating a self-terminating scan of positive items only. Of theories considered a priori, only that of Theios et al. (1973) appears to be consistent with the above result.

**Further Discussion**

In spite of what appears to be fairly conclusive evidence for the Theios et al. (1973) model under the conditions of the present experiment, caution is indicated in applying the results of the present study when discussing other situations in which the Sternberg effect has been found. It is possible that two

aspects of the population of items used in the present experiment—its small size and high internal confusability—would tend to encourage the formation of a Theios et al. stack, and that other mechanisms might be responsible for the Sternberg effect when larger item populations or less confusable items are used. The author hopes that other experimenters will explore this question, possibly employing the response-signal method used in the present study.

While a general model of recognition probably ought to incorporate a Theios et al. (1973) stack for use in some situations, it must obviously incorporate much more: it would take over a minute to scan through a 1,000-word vocabulary (at the rate found for the Theios et al. (1973) scan in this experiment), and we can understand spoken English considerably faster than that. One likely candidate for a general model of recognition is that posited by Juola and Atkinson (Atkinson & Juola, 1973, in press; Juola, Fishler, Wood, & Atkinson, 1971). In that model, a search process is resorted to only when a sufficiently reliable decision cannot be reached on the basis of simple strength. When highly confusable materials are used, the subject may have to resort to a search on all trials. The results of the present study indicate that at least in some situations, the search resorted to follows the model of Theios et al. (1973).

**REFERENCE NOTES**

1. Pachella, R. G. *Memory scanning under speed stress*. Paper presented at the Midwestern Psychological Association Meeting, 1972.
2. Reed, A. V. *Signal-integration models of speed-accuracy trade-off in recognition memory*. Paper presented at the Mathematical Psychology Meeting, 1973.

**REFERENCES**

ANDERSON, J. A. A theory for the recognition of items from short memorized lists. *Psychological Review*, 1973, **80**, 417-438.  
 ANDERSON, J. R., & BOWER, G. H. *Human associative memory*. Washington, D.C: Winston, 1973.  
 ATKINSON, R. C., & JUOLA, J. R. Factors influencing speed and accuracy of word recognition. In S. Kornblum (Ed.), *Attention and performance IV*. New York: Academic Press, 1973.  
 ATKINSON, R. C., & JUOLA, J. R. Search and decision processes in recognition memory. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology*. San Francisco: Freeman, in press.  
 BADDELEY, A. D., & ECOB, J. R. Reaction time and short-term memory: Implications of repetition effects for the high-speed exhaustive scan hypothesis. *Quarterly Journal of Experimental Psychology*, 1973, **25**, 229-240.

- BRACEY, G. W. Two operations in character recognition: A partial replication. *Perception & Psychophysics*, 1969, **6**, 357-360.
- BRIGGS, G. E., & BLAHA, J. Memory retrieval and central comparison times in information processing. *Journal of Experimental Psychology*, 1969, **79**, 395-401.
- BURROWS, D., & MURDOCK, B. B., JR. Effects of extended practice on high-speed scanning. *Journal of Experimental Psychology*, 1969, **82**, 231-237.
- COCHRAN, W. G. Testing a linear relation among variances. *Biometrics*, 1951, **7**, 17-32.
- CORBALLIS, M. C., BIRBY, J., & MILLER, A. Access to elements of a memorized list. *Journal of Experimental Psychology*, 1972, **94**, 185-190.
- HOVING, K. L., MORIN, R. E., & KONICK, D. S. Recognition reaction time and size of the memory set: A developmental study. *Psychonomic Science*, 1970, **21**, 247-248.
- 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.
- KIRSNER, K. Naming latency facilitation: An analysis of the encoding component in recognition reaction time. *Journal of Experimental Psychology*, 1972, **95**, 171-176.
- KLATSKY, 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.
- LANDAUER, T. K. Rate of implicit speech. *Perceptual & Motor Skills*, 1962, **15**, 646.
- MILLER, J. O., & PACHELLA, R. G. Locus of the stimulus probability effect. *Journal of Experimental Psychology*, 1973, **101**, 227-231.
- MURDOCK, B. B., JR. A parallel-processing model for scanning. *Perception & Psychophysics*, 1971, **10**, 289-291.
- NORMAN, D. A., & WICKELGREN, W. A. Strength theory of decision rules and latency in retrieval from short-term memory. *Journal of Mathematical Psychology*, 1969, **6**, 192-208.
- OLLMAN, R. T. Fast guesses in choice reaction time. *Psychonomic Science*, 1966, **6**, 155-156.
- REED, A. V. Speed-accuracy trade-off in recognition memory. *Science*, 1973, **181**, 574-576.
- STERNBERG, S. High-speed scanning in human memory. *Science*, 1966, **153**, 652-654.
- STERNBERG, S. Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist*, 1969, **4**, 421-457.
- SWENSSON, R. G. The elusive tradeoff: Speed vs. accuracy in visual discrimination tasks. *Perception & Psychophysics*, 1972, **12**, 16-32.
- THEIOS, J., SMITH, P. G., HAVILAND, S. E., TRAUPMANN, J., & MOY, M. C. Memory scanning as a serial self-terminating process. *Journal of Experimental Psychology*, 1973, **97**, 323-336.
- WINGFIELD, A., & BOLT, R. A. Memory search for multiple targets. *Journal of Experimental Psychology*, 1970, **85**, 45-50.
- WINGFIELD, A., & BRANCA, A. A. Strategy in high-speed memory search. *Journal of Experimental Psychology*, 1970, **83**, 63-67.
- YELLOTT, J. I., JR. Correction for guessing in choice reaction time. *Psychonomic Science*, 1967, **8**, 321-322.
- YELLOTT, J. I., JR. Correction for fast guessing and the speed-accuracy tradeoff in choice reaction time. *Journal of Mathematical Psychology*, 1971, **8**, 159-199.
- YIO, J. H., & SANTA, J. L. Reaction time in short-term recognition with digits and letters. *Psychonomic Science*, 1970, **20**, 121-122.

(Received for publication January 27, 1975;  
revision received March 26, 1975.)