

Probe Similarity and Recognition of Set Membership: A Parallel-Processing Serial- Feature-Matching Model

L. ROWELL HUESMANN

University of Illinois at Chicago Circle

AND

FREDRIC D. WOOCHEER

Stanford University

A model for memory scanning is proposed in which the encoded representation of a probe is compared in parallel with encoded representations of each item in the positive set. The within item matches are serial feature by feature comparisons that terminate when either a positive or negative criterion is reached. This model is shown to predict the results of a probe similarity experiment. The serial location of a similarity within an item affects negative reaction times, but the number of items in the positive set to which the probe is similar has no main effect. The model is also shown to yield predictions consonant with existing data on the relation between reaction times and set size and speed-accuracy trade offs.

In a memory scanning recognition task a subject must decide as rapidly as possible whether or not a presented test item was included in a previously presented set of items. Typically, a subject is asked to remember a list, s , of digits, letters or words presented visually or auditorily. He may or may not be allowed to rehearse the list. On a test trial he is presented with a test item, d , and asked to decide whether it was in the positive set, s . Usually, reaction time is the dependent variable of interest, and performance is designed to be nearly error free.

Sternberg's (1966, 1969) well known exhaustive scanning model for explaining recognition times in memory scanning was based on three major findings from his early experiments. First, recognition times increase linearly as a function of positive set size. Second, the slope of the positive and negative reaction times plotted against set size are about equal. Third, positive recognition times are independent of the serial position of the

The research reported in this paper was performed in part at Yale University where the senior author was on the faculty and the junior author was an undergraduate. It was supported in part by NSF Undergraduate Research Grant GY-9783. Requests for reprints should be sent to L. Rowell Huesmann, Department of Psychology, University of Illinois at Chicago Circle, Box 4348, Chicago, IL. 60680.

positive item in the to-be-remembered list. From these results Sternberg proposed an exhaustive serial-scanning model in which an encoded version of the test item, d , is compared with the memory representations of the items in s at a high rate (e.g., about 35 msec per item). Every item in the positive set, s , is generated and compared with the test item. Only after comparisons with all s items are completed does the subject make a decision and the appropriate response. Sternberg maintained that this exhaustive comparison process would not be inefficient if the generate and compare processes were rapid relative to the process which checks a register to see if a match occurred.

Recently several extensions of Sternberg's model and alternative theories have emerged for explaining recognition times in memory scanning. These extensions and theories have been proposed primarily to explain findings inconsistent with Sternberg's exhaustive scanning model. Most of the inconsistencies involved three relationships: (a) the relationship between recognition times and set size, (b) the relationship between reaction times for positive and negative responses, and (c) the relationship between positive reaction times and serial position of the positive item (Sternberg, 1975).

The proposed alternatives to Sternberg's theory mainly fall into four classes: extensions to the exhaustive scanning model to account for various phenomena, limited power parallel processing models, trace-strength models and self-terminating scanning models. The extensions to the exhaustive scanning model have included adding a process for retrieval of positive set items from long-term memory (Juola, Fischler, et al., 1971; Briggs & Blaha, 1969), adding a process for sampling components of the test item d at input (Swanson & Briggs, 1969), adding a self-terminating scan for categories within the positive set (Naus, Glucksberg, & Ornstein, 1971), and adding a priority-of-access hierarchy within the positive set (Burrows & Okada, 1971; Raeburn, 1972). The limited-power parallel comparison models assume that the items in the positive set are compared in parallel with the test item, but by a system whose comparison speed is limited by the number of items in the positive set. The more items, the slower the speed (Townsend, 1971; Murdock, 1971). Another type of model that has been proposed is the familiarity or trace-strength model which assumes subjects use a two-stage decision process. The first stage is a decision based on a trace or familiarity strength. The second stage is assumed to be a process requiring a linear increase in reaction time with set size (Atkinson & Juola, 1974; Baddeley and Ecob, 1970). Finally, Theios has argued that one cannot rule out a serial self-terminating scan from the available data (Theios et al., 1973).

One can see that a major element of almost every model has been either an exhaustive scanning system or a limited power parallel comparison system. Yet, so far, no one has offered a very convincing process model

specification for the efficiency and, hence, plausibility of an exhaustive scan or for the operation of a limited power parallel process. With regard to the exhaustive scan, Sternberg suggested that the procedure would be efficient if checking a signal bit to see if a match occurred takes longer than making the match. While engineers may have built similar systems (Steinberg, 1975), it is hard to imagine an automaton for which checking to see if a bit is zero or one would require more time than comparing two multiple feature items to see if they match. Recently Anderson (1973) proposed the more plausible explanation that a subject sums the result of all the comparisons to reach a decision. However, why the summing could not proceed concurrently with the comparisons was not clear. Similarly, it is no more sufficient to assert that a parallel process simply takes longer when more items are being checked. Any parallel processing model which assumes a varied distribution of comparison times for individual items could predict increases in response time with set size; so one must specify the underlying process to have a falsifiable model. As Gregg and Simon (1967) have argued, given equally accurate models, one should decide between them on the basis of just such considerations as a priori plausibility and falsifiability.

PROBE SIMILARITY AND SCANNING

One variable that has received relatively little attention but might distinguish between these alternative models is the probe's similarity to positive set items. The different models seem to predict quite different results for variations in the probe's similarity to the positive set items, particularly for negative probes.

In order to assess the effect of probe similarity, one must specify both the *between-item* scanning process and the *within-item* matching process used in the recognition task. Many models of item recognition have not included precise specifications of the within-item matching process, yet the within-item process can have a large influence on response time and how response time changes with positive set size.

If one considers a stimulus item to be defined by a vector of features, one can classify within-item matching procedures in three basic groups: serial-exhaustive feature matching, serial-self-terminating feature matching, and parallel feature matching. Of these, only the time for the serial self-terminating feature-matching process should be significantly influenced by a negative probe's similarity to one item in the positive set. Having a greater or lesser number of matching features should not affect matching time under the other two models, though false positive error rates would certainly be influenced. Furthermore, only with the serial-self-terminating within-item matching process should the comparison time and false positive error rate depend upon the location of the similarity within the set of features defining the negative probe.

The concept of a serial self-terminating feature matching process is not new. Egeth (1966) concluded on the basis of a series of experiments that this is a plausible model for same-different judgments with visual stimuli. Two studies manipulating probe similarity in item recognition have produced results suggestive of a serial within-item feature matching process. Lively and Sanford (1972) found that negative RT slopes were significantly less for dissimilar negative probes than for similar probes and concluded that a serial-within-item feature-matching process might be operating. A similar conclusion was reached by Duman, Gross, and Checkosky (1972) on the basis of recognition experiments with geometric patterns. Shavell and Atkinson (1974) analyzed the various combinations of within-item and between-item scanning models. They showed that while many such models predict nonlinear reaction time functions, the deviations from linearity would be extremely difficult to detect in the usual experimental paradigms.

The observed response time to a negative probe that is similar to a positive set item would depend, of course, on the between-item scanning process as well as the within-item matching process. Assuming a serial self-terminating within-item feature-matching procedure, how would negative reaction times be affected under the various between-item scanning models? For a serial-exhaustive scanning model, the effect of a similarity between the negative probe and one item would be to increase negative response time only a small amount. The increase would have to be less than the scanning rate, since only the time for checking the probe against the similar item could be affected. The effect with a serial self-terminating between-item scanning process would depend upon how the between-item and within-item processes were integrated. If each item were checked in turn, as is usually assumed, only a small increase in response time should be expected. If, on the other hand, the first feature of each item were checked, then the second, and so on, until all items mismatch, then a large effect would be found since the similar item would determine the time for the last mismatch. Likewise, if all items are processed in parallel, the similarity of the probe to any one item should have a large effect, since the time required for the last item to mismatch with the probe determines response time. These latter two models, in fact, are probably indistinguishable on the basis of reaction time data and will both be encompassed under the term parallel between-item scanning.

The degree to which the items in the positive set are similar to *each other* could also help discriminate between the various models on the basis of negative reaction times. Suppose all the items are very similar (homogenous) on one feature, be it syntactic or semantic. One compares the reaction time for a negative probe that has this feature with one that does not. Assuming a serial self-terminating within-item matching process, a serial exhaustive between-item scanning model would predict a large

increase in response latency since the matching time for each item is increased by the similarity. The parallel processing models also predict an increase, but one no larger than for the case of a single similar item in the positive set (assuming a fixed feature matching time).

Previous studies in which the positive set had a common feature have mostly shown that negative probes not having the common feature are rejected more rapidly than probes having the feature. (Ellis & Chase, 1971; Foss & Dowell, 1971; Dumas et al., 1972). Furthermore, two studies (Ellis & Chase, 1971; Foss & Dowell, 1971) indicate that negative reaction time plots become negatively accelerated when probes do not have the similar feature. However, these studies have not directly assessed how changes in the number of positive set items to which the probe is similar affect reaction time.

More exact predictions about the effect of probe similarity on reaction times and error rates can be made for specific experimental situations. In the next section we present the results of such an experiment on probe similarity. These results suggest a type of parallel-processing serial-feature-matching model which we develop in detail.

AN EXPERIMENT ON PROBE SIMILARITY

The experiment was a simple memory scanning task with short lists of words. The subjects were presented lists of familiar one-syllable, consonant-vowel-consonant words from a tape recorder. A fast presentation rate was used in an attempt to limit semantic associations. After each list, subjects were presented a series of test items and after each of these one-syllable CVC words a subject had to decide whether it was a member of the positive-set and make an appropriate response. His response and reaction time were recorded.

Features of stimuli. We will define a feature functionally as a part of a stimulus item that is processed as an entity and that requires ALPHA time units to generate and compare with another feature. This is a functional definition of feature and may differ considerably from structural definitions of feature used by others. It is derived from a view of information processing that presumes humans organize the same information in different ways in different situations. Thus, the model could be applied to the same stimuli under different conceptions of exactly what constitutes one feature and what value α has. In this experiment we will consider each CVC item to be defined by three features encoded respectively from (1) the beginning consonant, (2) the vowel, and (3) the end consonant. These three features used in processing are encoded from the phonetic characteristics of the three parts of the stimulus. The encodings are assumed to be performed sequentially as the word is processed and encoded. Of course, many other features of a stimulus may

be extracted and encoded, e.g., semantic features, but these should not have been important in the experiment.

Positive sets. Two factors were varied in constructing the positive sets: (a) size of positive set: three, four, or six words; and (b) list type: either homogeneous (*H*), in which the phonetic vowel sound was the same for every positive-set item, or nonhomogeneous (*N*), in which all vowel sounds were different. These two variables were combined orthogonally so that there were a total of six list conditions—*H3, H4, H6, N3, N4, and N6*. Each subject received three lists of each type, for a total of 18 lists.

Negative probes. The negative probes had two independent similarity factors: (a) consonant similarity—the probe item differed from *its target item* in the positive set on both beginning and end consonants (*BD, ,ED*), on only the beginning consonant (*BD, ,ES*) or on only the end consonant (*BS, ,ED*); and (b) vowel similarity—the test item either had the same phonetic vowel (*, VS,*) or a different phonetic vowel sound (*, VD,*) from its *target s*-set word. These were also combined orthogonally, yielding six different conditions for the test items. For example, if “pat” were the target item of the positive set, the six negatives could be “rack” (*BD, VS, ED*), “rock” (*BD, VD, ED*), “rat” (*BD, VS, ES*), “root” (*BD, VD, ES*), “pack” (*BS, VS, ED*), and “pick” (*BS, VD, ED*).

For the nonhomogeneous positive sets, each negative probe was *completely dissimilar* (*BD, VD, ED*) to every other item in the positive set, except its target item, because the items themselves were completely dissimilar. However, when the positive set had a homogeneous vowel sound, any negative probe with the relation (*, VS,*) to one item would, of course, have the same relation to all other items in the set.

There were 12 test items for each list, one for each of the six negative conditions and six positives. When necessary (for *S* = 3 and 4), a positive item was repeated in order to retain a positive to negative test frequency ratio of one, but only the first occurrence was of experimental interest. Thus, each subject received a total of 216 test words for the 18 lists.

METHOD

Subjects. The subject population was randomly divided into two groups (*A* and *B*). Each group received a different set of word lists in order to eliminate and check on any effects which might be due to the particular selection of words used. Within each group, serial position of words within lists was counterbalanced, and order of presentation of lists and serial position of test items were randomized.

Lists and materials. Before the experiment, the experimenter prepared a general word pool of approximately 1,000 one-syllable words, each of the form consonant–vowel–consonant. The words were obtained from a rhyming dictionary (Wood, 1936) and were then cross-checked with Thorndike and Lorge (1944) to ensure that they were common enough to be familiar to every subject. These words were organized according to the phonetic vowel sound. The experimenter then prepared two sets of word lists (*A* and *B*): Each set consisted of 18 lists—three lists for each of the experimental conditions; i.e., three for *H3, H4, H6, N3,*

N4, and N6. In addition, 12 test items, also from the general word pool, were assigned to each list—six negatives corresponding to each of the experimental conditions and six positives. For set size three, each positive set item was repeated once; for set size four, two of the items were repeated; and for set size six, each positive-set word was tested once. No word from the word pool was used twice as either a positive-set or test item. The experimenter then made four tape recordings for presentation of lists. The first tape consisted of the first set of lists (A), each list being followed by its appropriate test items. The order of the lists was randomized, with the restriction that one list of each of the experimental conditions appeared in each third of the whole set. The order of the test items was also randomized, this with the constraint that each of the six negative conditions appeared equally frequently in each of the 12 serial positions. The second tape was identical to the first, except that the serial position of the words within the positive-set lists was randomly changed. Similarly, two tapes were made of the second set of lists (B). In addition, all four tapes began with the same two practice lists, one of condition N4, the other H3, and each of these lists contained and was followed by 16 (instead of the usual 12) test items.

Procedure. Subjects were 15 male and five female undergraduate students at Yale University, all enrolled in an introductory psychology course. All subjects were right-handed. The 20 subjects were divided into four groups; five subjects were presented each of the different tapes.

The experimenter first explained to the subject the general paradigm of a memory scanning task, and then gave the particular instructions for this experiment. The subject was then seated in front of three telegraph keys, and was separated from the experimenter and the equipment by a large screen to avoid distraction. The keys were arranged in a straight line, about 6 cm apart from each other. The subject was instructed to keep his right forefinger on the center key at all times and to indicate a positive or negative response by depressing one of the adjacent keys with the same finger as quickly as possible; the subject was cautioned not to sacrifice accuracy for speed and to report any difficulty in hearing the words immediately to the experimenter. Subjects were also told that the timer could not be activated unless the center key was depressed to ensure that they would not lift their finger before presentation of the test items. In reality, the wires from the center key were not attached to anything. In addition, the experimenter changed which keys indicated the positive and negative responses with each new subject.

The experimenter then presented the first practice list from the tape recorder through a set of headphones which the subject was wearing. The rate of presentation was about 90 words/min. The subject was then asked to repeat the list. If correct, the experiment continued; if not, the experimenter verbally repeated the list until the subject could repeat it perfectly himself. The equipment was arranged so that the onset of the test item triggered a voice key which began the timer. (Every attempt was made in recording and adjusting the volume to be sure that the voice key was triggered by the beginning of the first consonant sound.) When the subject then pressed one of the response keys, the electric timer automatically stopped, and a light notified the experimenter which response the subject had made. The experimenter then recorded the time, in milliseconds, and the response, telling the subject if he had made a mistake. The next 15 test items were then presented; before each one, however, the subject had to repeat the list correctly. About 4 sec elapsed between the time the subject repeated the list and the onset of the next test item. In this manner, the second practice list and the 18 experimental lists were presented to the subject, with a break of about 90 sec between the eighth and ninth experimental lists. The sequence of events was: (a) presentation of list; (b) repetition of list by the subject until correct; (c) presentation of test item; (d) response by the subject. Steps (b)–(d) were repeated 12 times for each experimental list. There were a total of 18 lists. The total experimental session lasted approximately 1 hr.

RESULTS AND DISCUSSION

Each subject received three lists in each experimental condition. The reaction times for those test words which the subject reported he could not

hear clearly were discounted, as were the times on those words on which the subjects made errors. The dependent measure in each condition, then, was the median of a subject's remaining times (three or less).

The differences between the mean reaction times for the different word sets used were found to be insignificant, except for the case of positive times with set size three where the means were slightly different ($p < .05$). As a result, the data from the two sets are combined in all analyses.

Set size. The mean reaction times for positive and negative responses are plotted as a function of set size in Fig. 1. Reaction times increased significantly with set size [$F(2,36) = 14.94, p < .01$] and negative reaction times were significantly slower than positives [$F(1,18) = 69.36, p < .01$] though we shall see that some types of negatives were responded to more rapidly than positives. The homogeneity of the stimulus list did not have a significant main effect on reaction time though it interacted significantly with set size [$F(2,36) = 3.63, p < .01$].

The relation between positive nonhomogeneous times and set size is linear with a slope of 29.7 msec/item. This condition is the one most similar to previous memory scanning experiments and the slope is comparable to that found in such studies. The positive homogeneous line, on the other hand, appears negatively accelerated with an overall slope of 13.4 msec/item while the homogeneous and nonhomogeneous negative RT lines have slopes of 15.9 and 4.7 msec/item, respectively. These three conditions differ from previous studies in list homogeneity and probe similarity.

Serial position curves. As Fig. 2 shows, bowed serial position curves were found for response times to positive probes. The serial position effects are significant for set sizes four and six [$F(3,27) = 3.66, p < .05$; $F(5,45) = 3.09, p < .05$], but insignificant for set size three [$F(2,36) = 1.45$]. The curve for set size six, by itself, might indicate that some subjects are using a self-terminating between-item scan, but the other curves do not

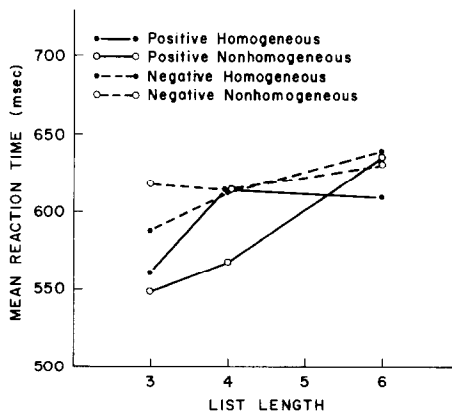


FIG. 1. The relation between positive set size (list length), set homogeneity, and reaction times in the probe similarity experiment.

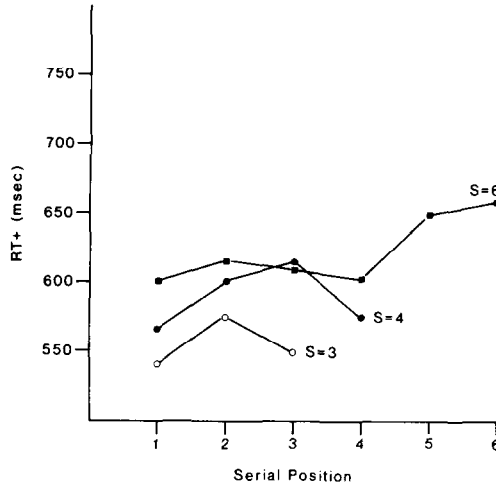


FIG. 2. The effect of serial position on positive reaction times as observed in the experiment.

support this interpretation. Bowed serial position curves like these have been found in many recent studies (e.g., Burrows & Okada, 1971).

Similarity Effects and Negative Reaction Times

Of greatest interest in this experiment are the effects of similarity on negative reaction times. The mean negative reaction time for each type of similarity is plotted in Fig. 3. One can see that there is an increase in

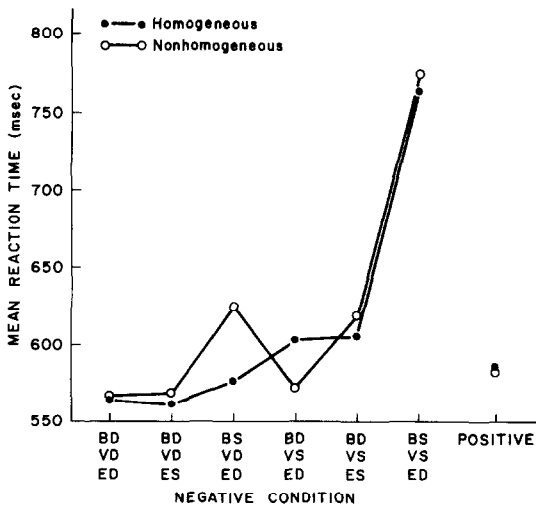


FIG. 3. The mean reaction times to negative probes as a function of the similarity between the probe and the most similar item in the positive set.

reaction time as the probe's differences with the target item occur farther toward the end of the negative probe. Furthermore, as Table 1 more clearly shows, the effects revealed in Fig. 3 are very significant. Consonant similarity [$F(2,26) = 111.7, p < .01$] and vowel similarity [$F(1,18) = 124.3, p < .01$] both produce significant main effects on reaction time and interact significantly [$F(2,36) = 70.7, p < .01$]. The more the similarity and the closer the similarity is to the beginning of the probe (i.e., the farther the differences are toward the end), the slower is the negative reaction time.

While overall mean reaction times to negative probes are no different for homogeneous and nonhomogeneous lists, homogeneity interacts significantly with both vowel similarity [$F(1,18) = 4.76, p < .05$] and consonant similarity [$F(2,36) = 6.87, p < .01$]. For example, the mean VD reaction time was 90.5 msec quicker than the mean VS time on homogeneous lists, but only 69.3 msec quicker than VS on nonhomogeneous lists.

These results suggest that the within-item feature matching process is a serial self-terminating process and that the between-item scanning procedure is a parallel comparison process. The significant consonant similarity effect was primarily due to the fact that response times for the (BD, ,ED) and (BD, ,ES) probes were much faster than for the (BS, ,ED) probes. In both the (BD, ,ED) and (BD, ,ES) conditions, the difference between the negative item and target positive-set item appears at

TABLE 1
MEAN NEGATIVE REACTION TIMES AS A FUNCTION OF HOMOGENEITY
AND PROBE SIMILARITY

Vowel similarity	Consonant similarity			Mean
	(BD, ,ED)	(BD, ,ES)	(BS, ,ED)	
Homogeneous lists				
VS	603.8	606.8	763.9	658.2
VD	564.0	562.1	577.1	567.7
Mean	583.9	584.5	670.5	612.7
Nonhomogeneous lists				
VS	573.0	620.2	775.5	656.2
VD	566.9	568.6	625.3	586.9
Mean	570.0	594.4	700.4	621.6
All lists				
VS	588.4	613.5	769.7	657.2
VD	565.4	565.3	601.2	577.3
Mean	577.0	589.5	685.5	617.3

the beginning of the item. Thus a serial self-terminating feature-testing process would produce mismatches early and reject the probe rapidly. For the (*BS*, *,ED*) condition, however, the difference between the probe and target item is in the end consonant; the subject must then process at least to the vowel, and perhaps all the way to the last feature, before he gets enough mismatches to reject the probe.

Assuming, on the basis of the above analysis, that a serial-self-terminating within-item feature matching process is being used, the same data suggest that the between-item scanning process is a parallel comparison. If a serial between-item process were being used, the increase in matching time for a (*BS*, *,ED*) probe over a (*BD*, *,ED*) probe should be small relative to the overall scanning rate per item since only one item is affected. But the observed difference was very large, on the order of 100 msec, while the estimated scanning rate under an exhaustive model would be 5 to 30 msec/item, and twice that under a self-terminating model (See Fig. 1).

The observed interaction of homogeneity and similarity also suggests the serial feature-matching, parallel between-item scanning process. The significant interaction between homogeneity and vowel similarity reflects the fact that negative reaction times to *VS* probes were not different for homogeneous and nonhomogeneous lists, while negative reaction times to *VD* probes were different (See Table 1). In the *VS* condition, even though the probe's vowel matched the vowel of every item in the homogeneous positive set and only one item in the nonhomogeneous set, there was no difference in response time. This is consistent with parallel processing of the positive set but not with serial scanning. In the *VD* condition, the mean reaction time was about 20 msec faster for homogeneous than nonhomogeneous lists. With a serial within-item feature matching model, this result would be expected if a substantial number of subjects detected that all the positive items had a common feature (the vowel) and thus processed that feature sooner in the homogeneous than nonhomogeneous conditions.

So far, we have seen that the results of the experiment seem to provide stronger support for a serial-self-terminating within-item feature matching procedure coupled with a between-item parallel comparison process than for any other model. But how well do the data really fit such a model? It is impossible to predict exact response times without assigning values to parameters and elaborating the model. However, by making only one additional assumption, we can predict the ordinal relations that should exist between the mean times in Table 1. The assumption is that a subject always says "no" in this experiment if the probe has mismatched every positive item on at least two (of the three) features, and he may say "no" even if only one feature has been mismatched on some of the items. In other words, we are assuming that every subject's negative criterion (*NC*) is either two or one mismatches.

The tentative model, then, is that a subject compares a probe in parallel against each item in the positive set. The within-item comparison consists of a serial-self-terminating feature-matching process. Again, we must emphasize that "feature" is defined functionally as a unit of information that the subject processes as an entity. Like the concept of "chunk" in memory models, a feature may consist of many smaller elements grouped together for processing. In our experiment the stimuli were defined in terms of three features which of course were actually "chunks" of smaller elements, the manipulations were based on the definitions, and the analysis has proceeded under this definition. Yet the specific definition is not an essential part of the process model now being suggested. In fact, for the ordinal predictions below, it is not even essential that the same quantity of time (α) is used in processing each feature. In the case of a homogeneous list, the vowel feature is processed first; otherwise the features are processed in input order. The subject says "no" when the probe has been found to mismatch with all items. A probe is said to mismatch with one item when the number of feature mismatches reaches a number called Negative Criterion (NC). Negative Criterion in this experiment may vary from subject to subject and list to list, but is always assumed to be one or two. As a result, there is some possibility of a subject not reaching NC with an item after all features have been tested. In such a case a subject may make a false-positive response (in which case his time was not included in the data for Table 1) or may guess correctly or recheck correctly. In any of the cases, some additional time is consumed to reach a decision.

Ordinal predictions. On the basis of this tentative model, the equations shown in Table 2 were derived. In each cell there are two equations for the time a subject would need to respond "no." The two equations specify the time a subject would need to reach $NC = 1$ and $NC = 2$ in terms of times to process each part of the item.

A few examples will help to show how Table 2 was constructed. Let t_B be the time needed to process the beginning consonant, t_V be the time for the vowel, t_E be the time for the end consonant, and t_R be the time for the "rechecking" or "guessing" process. Consider condition $N(BD, VD, ED)$. First, the beginning consonant is tested. Since it differs from every item in the positive set including the target item on its beginning consonant, the subject would respond "no" if NC were one. If so, the time would be less than (or equal to) the time component for the beginning consonant ($\leq t_B$). Hence, the first equation. If $NC = 2$, the vowel's elements are tested, and since the vowel also differs, NC will be reached before all the vowels' elements are tested. The total time necessary would then be $t_B + \leq t_V$. Hence, the second equation. These are the only two possibilities. As a second example, consider condition $H(BD, VS, ES)$. The elements of the vowel are tested first because the list is homogeneous. Since the vowel is the same as the target item, a match occurs. Thus, the subject must

TABLE 2

RT EQUATIONS FOR ALTERNATIVE WAYS TO REACH *NC* UNDER THE PROPOSED MODEL^a

Vowel similarity	Consonant similarity		
	(BD, , ED)	(BD, , ES)	(BS, , ED)
Nonhomogeneous lists			
VS	$t_1 = \leq t_B$	$t_1 = \leq t_B$	$t_1 = t_B + t_V + \leq t_E$
	$t_2 = t_B + t_V + \leq t_E$	$t_2 = t_B + t_V + t_E + t_R$	$t_2 = t_B + t_V + t_E + t_R$
VD	$t_1 = \leq t_B$	$t_1 = \leq t_B$	$t_1 = t_B + \leq t_V$
	$t_2 = t_B + \leq t_V$	$t_2 = t_B + \leq t_V$	$t_2 = t_B + t_V + \leq t_E$
Homogeneous lists			
VS	$t_1 = t_V + \leq t_B$	$t_1 = t_V + \leq t_B$	$t_1 = t_V + t_B + \leq t_E$
	$t_2 = t_V + t_B + \leq t_E$	$t_2 = t_V + t_B + t_E + t_R$	$t_2 = t_V + t_B + t_E + t_R$
VD	$t_1 = \leq t_V$	$t_1 = \leq t_V$	$t_1 = \leq t_V$
	$t_2 = t_V + \leq t_B$	$t_2 = t_V + \leq t_B$	$t_2 = t_V + t_B + \leq t_E$

^a Let t_B = time necessary to process beginning consonant's features, t_V = time necessary to process vowel's features, t_E = time necessary to process end consonant's features, t_R = time necessary for rechecking process.

process the beginning consonant, which, since different, would provide the required number of mismatches if $NC = 1$. The time would then be $t_V + \leq t_B$, shown in the first equation. If NC were 2, however, not only must the end consonant's elements be tested, but since it matches, the rechecking operation must occur. The total time would then be $t_V + t_B + t_E + t_R$, as shown in the second equation.

Based on the times in Table 2, we can make predictions of the reaction times of most conditions relative to the other conditions within the same list type. In doing so, we assume that the probability of rejecting a probe after one mismatching feature is found (i.e., $NC = 1$) is independent of whether or not, other, as yet undiscovered feature mismatches exist. Hence, for example, the probability of reaching NC from only the beginning consonant's discrepancy is assumed to be the same in any nonhomogeneous condition with BD . The observed reaction-time in each cell should be the weighted sum of the RT for each equation multiplied by its probability of occurrence. We make no assumptions concerning the relative magnitudes of t_B , t_V , t_E , and t_R , not any concerning the specific probabilities of any of the equations within each condition (except that they are all greater than zero and less than one).

The various predictions resulting from these equations are contained in Table 3. Note that the only unspecified relation is between $N(BS, VD, ED)$ and $N(BD, VS, ES)$. When these predictions are compared with the results

TABLE 3

ORDINAL PREDICTIONS FOR OBSERVED NEGATIVE RTs UNDER THE PROPOSED MODEL

Homogeneous lists		
H(BD,VD,ED) = H(BD,VD,ES) < H(BS,VD,ED) < H(BD,VS,ED) < H(BD,VS,ES) < H(BS,VS,ED)		
Nonhomogeneous lists		
N(BD,VD,ED) = N(BD,VD,ES) < N(BD,VS,ED) < N(BS,VD,ED) < N(BS,VS,ED) < N(BD,VS,ES)		
Interlist		
H(BD,VD,ED) < N(BS,VD,ED) < N(BS,VS,ED)	H(BD,VD,ES) < N(BS,VD,ED) < N(BS,VS,ED)	H(BS,VD,ED) < N(BS,VD,ED) < N(BS,VS,ED)
H(BD,VS,ED) > N(BD,VD,ED) > N(BD,VD,ES) > N(BD,VS,ED) < N(BS,VS,ED)	H(BD,VS,ES) > N(BD,VD,ED) > N(BD,VD,ES) > N(BD,VS,ED) > N(BD,VS,ES) ^a < N(BS,VS,ED)	H(BS,VS,ED) > N(BD,VD,ED) > N(BD,VD,ES) > N(BS,VD,ED) > N(BD,VS,ED) > N(BD,VS,ES) = N(BS,VS,ED)

^a This was the only prediction not confirmed by the data.

of the experiment (Table 1), the ordinal fit is perfect, even to the extent that predicted equal times are within a few milliseconds of each other. In other words, all 29 of the intralist predictions made in Table 3 were verified by the data.

Predictions can also be made on the relationships between most conditions for the different list types, although they are harder to make without any assumptions about the relative time components for each of the word parts or the probabilities of the different alternatives. Table 3 contains 21 interlist predictions. When these are once again compared with the experimental results, all but one of the predictions are verified.

Why was the mean reaction time for the (BS, VS, ED) negatives so much slower than for the (BD, VS, ES) negatives? The model certainly predicts that negatives of the type (BS, VS, ED) should be the slower, but it is not clear why they should be so much slower.

We offer two possible reasons for the large difference. First, the (BS, VS, ED) condition had a false positive error rate of 18% compared to the overall false positive rate of 1.8%. Eighty percent of all false positives occurred in this condition. The mean reaction time for these false positives (522.4 msec) was almost 250 msec faster than the mean time for the (BS, VS, ED) condition and 60 msec faster than the mean time for "yes" responses. Therefore, it seems possible that the faster subjects were making these false positive errors and the exclusion of their times raised the observed mean. Another possibility is that the three features structurally defined for the experiment (beginning consonant, vowel, and end consonant) are not really functionally equally important in recognition. This seems even more plausible. In particular, if the end consonant were

less important than the beginning consonant, a mismatch on the end consonant would be less likely to surpass *NC*. Hence, rechecking would be more likely to occur in the (*BS*, *VS*, *ED*) condition.

We have not yet explicitly specified how subjects would make positive recognition decisions under the model being proposed. Yet in the above discussion we have tacitly assumed that the procedure must be one of saying "yes" when enough matches with an item's features are made to satisfy some positive criterion (*PC*). We will not elaborate this proposal until later, but we will proceed under the assumption that a "yes" response is given whenever sufficient matches to reach *PC* are achieved before *NC* has been reached for every item.

Similarity Effects and Set Size

To this point we have based our conclusions on the effects of similarity over all set sizes; yet there were differences in the similarity effects under the three set sizes.

As Table 4 reveals, the ordering of negative RTs was almost identical for all three set sizes—the only exception being that the mean (*BD*, *VD*, *ED*) time was not the fastest for set size three as it should have been. Nevertheless, set size interacted significantly with both vowel similarity [$F(2,36) = 8.9, p < .01$] and, more importantly, consonant similarity

TABLE 4
MEAN NEGATIVE REACTION TIMES AS A FUNCTION OF POSITIVE SET SIZE
AND PROBE SIMILARITY

Vowel similarity	Consonant similarity			Mean
	(BD, , ED)	(BD, , ES)	(BS, , ED)	
Set size 3				
VS	549.5	567.9	779.2	632.2
VD	561.3	548.5	616.9	575.6
Mean	555.4	558.2	698.1	603.9
Set size 4				
VS	604.3	629.1	757.0	663.4
VD	549.0	553.8	586.8	563.2
Mean	576.6	591.5	671.9	613.3
Set size 6				
VS	611.5	643.6	773.0	676.0
VD	586.1	593.7	599.8	593.2
Mean	598.8	618.6	686.4	634.6

[$F(4,72) = 9.5, p < .01$]. While the reaction time to *BD* probes increased with set size, the reaction time to (*BS*, *ED*) probes did not. Apparently, when the mismatching features between the probe and the positive set occur near the end of the item, the effect of set size is reduced. Why should this be the case under the hypothesized model?

Under the hypothesized model, the increase in overall mean negative reaction time with set size [$F(2,36) = 10.32, p < .01$] results from an increase in the number of features that must be checked before negative criterion has been reached with each item. There is a nonzero probability that any feature of a negative probe may match the corresponding feature of a positive set item either because of similarities, encoding errors, or processing errors. As a result, the expected number of features that need to be processed before *NC* is reached with all items in the positive set will increase with set size as a negatively accelerated exponential. The asymptote of this function will be determined by the number of features in an item. Therefore, if all the features in an item already need to be processed before a decision can be made, increasing set size should have little effect on the negative reaction times under the proposed model. The data in Table 4 conform fairly well to this interpretation.

Similarity Effects and Error Rates

The error data of Table 5 provide additional support for the proposed model in two ways. Perhaps most significant is the fact that the greatest number of false positive errors occurred in the (*BS*, *VS*, *ED*) condition where the opportunities for incorrectly reaching a positive criterion (*PC*)

TABLE 5
OBSERVED ERRORS IN EACH EXPERIMENTAL CONDITION^a

Set size	Set homogeneity	Negative probes						Positive probes						Total	Rate
		Probe similarity						Serial position of target							
		(BD, .ED)		(BD, .ES)		(BS, .ED)									
		VS	VD	VS	VD	VS	VD	1	2	3	4	5	6		
3	H	0	1	1	0	3	1	0	0	0	—	—	—	6	.011
	N	0	0	0	0	7	0	0	0	0	—	—	—	7	.013
4	H	0	0	1	0	1	1	1	3	1	0	—	—	8	.013
	N	1	0	1	0	12	0	1	3	0	0	—	—	18	.030
6	H	0	0	0	0	5	0	1	4	0	2	2	1	15	.021
	N	0	0	1	0	4	0	1	3	1	0	3	7	20	.028
Total		1	1	4	0	32	2	4	13	2	2	5	8	74	.020
Rate		.003	.003	.011	.000	.089	.006	.011	.036	.006	.008	.042	.067		

^a Based on 20 subjects who each received three different versions of each positive set type yielding a total of 60 responses per cell.

before a negative criterion (*NC*) were at a maximum if a subject were using a within-item serial feature scan. In particular, note that more errors occurred in this condition than in the (*BD*, *VS*, *ES*) condition which was equally similar but provided less opportunity for the subject to reach *PC* before *NC*. The mean reaction times for the false positives support this interpretation that false positives are usually caused by too rapid a response. While the mean RT for correct responses in the (*BS*, *VS*, *ED*) condition was 769.7 msec, the mean RT for false positives in this condition was 522.4 msec. The second interesting finding is that there are significantly fewer errors in the homogeneous (30) than nonhomogeneous (46) condition ($z = 1.84, p < .05$). This is true for both false positives and misses. Such a finding lends additional credence to the proposal that subjects extract common positive set features and process them first. The repeated, redundant presentation of the common vowel would reduce the chances of encoding or processing errors and, therefore, the chances of false positives or misses.

Similarity Effects and Positive Reaction Times

The effect of homogeneity (inralist similarity) on positive reaction times also provides evidence against serial-exhaustive scanning. The relationship between set size, homogeneity, and mean positive reaction time is shown in Fig. 1. The negatively accelerated curve for the homogeneous condition is similar to results reported elsewhere (Foss & Dowell, 1971). With a serial exhaustive scan (coupled with a serial self-terminating within-item feature matching process), one would expect to find that the mean positive RT would be significantly greater for homogeneous lists. If each item in the positive set has the same vowel the within-item serial scan will take longer on every item than it would in the nonhomogeneous case, so the slope would be greater. These predictions are not borne out. Positive RTs for homogeneous lists were not significantly slower than for nonhomogeneous lists (594.9 versus 584.7 msec), and the positive RT slope for the homogeneous lists was actually lower than for nonhomogeneous (*H* slope = 13.4 msec/item, *NH* slope = 29.7 msec/item). We cannot make predictions about positive reaction times from the parallel-scanning, serial feature-matching model until we formalize the model in the next section.

The findings of this experiment on similarity effects have provided substantial support for an item recognition process consisting of within-item serial feature matching and the parallel processing of all items. The significant differences in negative reaction times produced by consonant and vowel similarity could best be explained by such a process. The effect of list length and list homogeneity on positive and negative reaction times also suggested this type of process. Finally, the observed effect of similarity on error rates was consistent with such a model. We will

now specify such a model more formally and show that it can predict the other important characteristics of item recognition data.

FORMALIZING THE MODEL

A diagrammatic flow-chart explaining the basic recognition procedure being proposed is presented in Fig. 5. The procedure assumes that the positive set has been stored in nodes on a page of episodic memory and can be accessed as shown in Fig. 4.

When a positive-set is presented to the subject, he processes each item and stores its features at a node in memory. For auditory presentation of words, the features would be chunks of information encoded in sequence from the auditory input. Since extraction and storage of the features occurs rapidly, some features will inevitably not be stored and some incorrect features may be stored. Rehearsal would improve the encoding of the features. Those items whose serial position is in the middle of the positive-set would be encoded less completely with more errors than those items at the ends.

When the recognition part of the experiment begins, and a test item is presented, it is also encoded in terms of its features in the recognition system's input buffer. In this comparison stage, the test item is compared in parallel with all the items in the positive-set. Normally, each feature of the test item is accessed in sequence by the cell-selector and compared in parallel with the corresponding features of the items in the positive set. However, the cell-selector can be directed to access features in alternative orders if that is more efficient. If the subject has noted that the positive set has a common feature, he may use a recognition program in which the

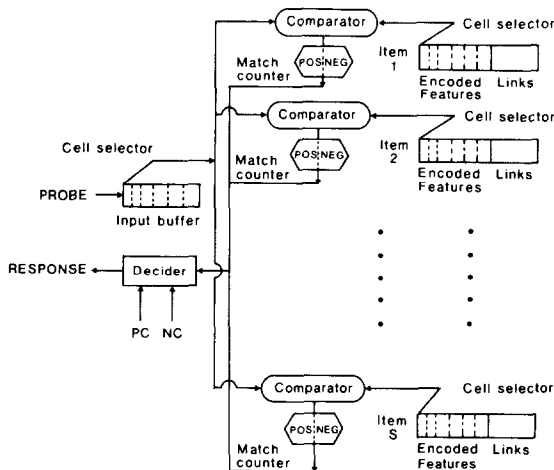


FIG. 4. A diagram of the hypothesized retrieval system for recognizing an item's presence in a page of episodic memory.

HUESMANN AND WOOCHER

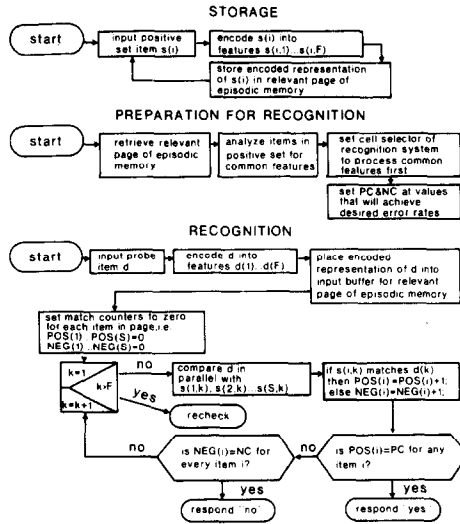


FIG. 5. A flow chart of the parallel-processing serial feature-matching model for recognition of set membership. $s(i,k)$ represents the k th feature of the i th item in the positive set. S is the positive set size and F is the number of features into which an item is encoded.

cell-selector processes that feature first (e.g., a common vowel, a semantic feature). The first feature selected in the test item is compared in parallel with the corresponding feature in all of the nodes of the page for the positive set. For every match, a positive counter associated with the matching node is incremented. For every mismatch, a negative counter is incremented. Then another feature is selected and compared.

Since there always may be incomplete encodings, encoding errors, or other processing errors, two encodings of the same item may not match perfectly. Similarly, encodings of different items may match on a number of features because of inherent similarities and errors. Thus, in order to make a decision as to whether the test item is a member of the positive-set, the subject sets criteria. If enough features in the test item match a positive-set item, the positive criterion (PC) is reached, and the subject terminates the comparison process and makes a positive response. If enough features mismatch, the negative criterion (NC) is reached and comparison for that item stops. Only after *all* the items have reached NC , though, can the subject respond negatively. In order to determine whether an item has reached criterion, therefore, the subject compares the positive and negative counters in parallel to the respective criterion after every feature comparison stage, and then continues with the next feature comparison stage if necessary.

This model, though derived independently, could be viewed as a generalization of one of the models suggested by Shavell and Atkinson (1974). They outlined a set of models based on the conceptions that stimuli

are encoded as vectors of features, that recognition requires an error-prone within-item feature matching process, and that reaction times depend upon the number of features that must be checked to make a decision. However, they assumed that a subject's negative criterion is always one feature mismatch and his positive criterion is always F feature matches (where F is the total number of features). We propose that NC and PC are varied to meet the requirements of the specific situation.

Criteria and Error Rates

How would a subject select appropriate positive and negative criteria for a recognition problem? We propose that a subject sets the criteria to obtain the desired speed-accuracy trade off. A subject should use the lowest possible criteria he can while maintaining acceptable error rates. The values selected would depend upon the perceived base rates of positive and negative items and the perceived cost of the two different kinds of errors, as well as the perceived probabilities of incorrect feature matches and mismatches. A subject's criteria may change during the course of an experiment as his perception of these factors changes.

In developing the relations between error rates and criteria we will concentrate on the case where positives and negatives are perceived to be equally likely, and false positives and misses are perceived to be equally costly. Generalization to other situations should pose no special problem.

We must now introduce several parameters: let F be the number of features in an item; S be the size of the positive set; PC be the positive criterion, $0 < PC \leq F$; NC be the negative criterion, $0 < NC \leq F$; $1 - q$ be probability that feature i of a positive probe *does not match* the corresponding feature of its target item in the positive set; and $1 - p$ be probability that feature i of a negative probe *matches* the corresponding feature of item j in the positive set.

The rechecking process. In order to calculate the expected error rates for a particular PC and NC , we need to make an assumption about what happens when neither PC nor NC is reached. This could happen whenever $PC + NC > F$. Is a guess made? Is the matching process repeated? Is a new process attempted? We suggest that experimental conditions affect what the subject does in this situation. However, in order to develop more detailed predictions from the model, we must make some assumptions about the rechecking process. Unfortunately, since different rechecking procedures may be used in different situations, the effect of any assumption about rechecking is to reduce the generality of the model. The assumption we prefer is that rechecking is a slow but very accurate process during which all F features in the probe are recompared at the original rate with the items in the positive set that did not reach criterion. After all F features have been reprocessed, the correct decision is made. In other words, processing time is doubled and Prob (error/recheck) is zero.

The advantages of these assumptions are that (a) no new free parameters are introduced into the model, and (b) the predicted error rates are not affected by what happens during rechecking. To minimize the importance of these assumptions, though, we will restrict our model to experimental situations in which only a small number of responses require rechecking. In fact, this is not much of a restriction since in most memory scanning experiments the positive set is learned well enough so that $1 - p$ and $1 - q$ are quite small; so the probability of rechecking does not become significant until S is very large.

Expected error rates. For the model described above,

$$\text{Prob (False pos/neg probe)} \quad (1.1)$$

$$= \text{Prob (reach } PC \text{ with at least one of } S \text{ items in positive set)}$$

$$= 1 - \text{Prob (reach } PC \text{ with none of } S \text{ items in positive set)}$$

$$= 1 - (\text{Prob (not reach } PC \text{ with item } i))^{S-1}$$

$$= 1 - (1 - \text{Prob (reach } PC \text{ with item } i))^{S-1}$$

$$= 1 - \left(1 - \sum_{j=PC}^{\min(F, PC+NC-1)} \binom{j-1}{PC-1} (1-p)^{PC} p^{j-PC} \right)^{S-1}.$$

$$\text{Prob (Miss/pos probe)} \quad (1.2)$$

$$= \text{Prob (reach } NC \text{ with target item)} \times \text{Prob (not reach } PC \text{ with any other item in positive set)}$$

$$= \left[\sum_{j=NC}^{\min(F, PC+NC-1)} \text{Prob (reach } NC \text{ with target item on } j \text{th feature test)} \right]$$

$$\times [1 - \text{Prob (False pos/neg probe for set size } S - 1)]$$

$$= \left[\sum_{j=NC}^{\min(F, PC+NC-1)} \binom{j-1}{NC-1} (1-q)^{NC} q^{j-NC} \right]$$

$$\times \left[\left(1 - \sum_{j=PC}^{\min(F, PC+NC-1)} \binom{j-1}{PC-1} (1-p)^{PC} p^{j-PC} \right)^{S-1} \right].$$

We can see from these equations that the PC is most important in determining the rate of false positives and that NC is most important in determining the rate of misses. In fact, if $PC + NC > F$, then NC will not affect the rate of false positives. (Assuming again that error rates are negligible during rechecking.)

Constructing criteria. We proposed above that people choose NC and PC to produce the fastest responding possible with the desired error rates. However, a subject cannot know the values for the noise parameters $1 - p$ and $1 - q$ ahead of time, and certainly would not compute the desired

criteria analytically even if he could. We must assume that subjects initially construct criteria on the basis of prior experience, and change criteria as the experiment proceeds to obtain the desired error rates. We propose the subject tries to set his positive criterion (PC) to a value that will give the desired rate of false positives, and then he tries to set his negative criterion (NC) to a value that will give the desired rate of misses. If the sum of the resulting PC and NC is greater than F , the desired rates can be approximated in this way. If $PC + NC$ turn out to be no greater than F , the error rates will be *less* (or better) than the desired rates; however, we would expect that on the average $PC + NC$ would be greater than F , so that reaching one criterion implies that the other criterion can no longer be achieved.

As Equation (1.1) shows, the rate of false positives increases with set size if criteria remain constant. In order to maintain an approximately constant rate of false positives, a subject would have to increase his PC as set size increases. The rate of misses, however, is only affected by set size through the second component of Equation (1.2) which represents the probability of a false positive for $S - 1$ items. Since from Equations (1.1) and (1.2),

$$\text{Prob}(FP|S - 1) = 1 - [1 - \text{Prob}(FP|S)]^{(S-1)/S}$$

for small set sizes and large $\text{Prob}(FP|S)$,

$$\text{Prob}(FP|S - 1) < \text{Prob}(FP|S)$$

and the observed miss rate should decrease slightly with set size as $\text{Prob}(FP|S - 1)$ increases to $\text{Prob}(FP|S)$. However, for large S (i.e., $S > 10$) or small $\text{Prob}(FP|S)$ (i.e., $\text{Prob}(FP|S) < .2$),

$$\text{Prob}(FP|S - 1) \approx \text{Prob}(FP|S),$$

and the rate of misses would not change significantly with set size, since $\text{Prob}(FP|S)$ is being kept constant through changes in PC . As a result a subject could control his error rates reasonably well by altering his positive criterion directly with set size and leaving his negative criterion unchanged.

Reaction Times for Positive and Negative Probes

We can now derive expressions for expected reaction times under the parallel-processing serial feature-matching model.

Let

m be the minimum ($F, PC + NC - 1$).

α be the time needed to generate and compare feature i of the probe in parallel with the corresponding feature of all the items in the positive set, $\alpha > 0$.

γ be the time needed for the processes unrelated to the memory scanning process, e.g., input and output time, $\gamma \geq 150$.

Then

$E(\text{RT yes for pos probe})$

$$= \alpha E(\text{features compared to reach } PC / \text{"yes"} \text{ resp to pos probe}) \\ + 2\alpha F P(\text{recheck} / \text{"yes"} \text{ resp to pos probe}) + \gamma. \quad (2.1)$$

Since reaching PC or rechecking always produce a yes response to a positive probe,

$$E(\text{RT yes for pos probe}) = \alpha \frac{\left\{ \sum_{j=PC}^m j \cdot P(\text{reach } PC \text{ on trial } j) \right\}}{P(\text{"yes"} \text{ resp to pos probe})} \\ + 2\alpha F \frac{P(\text{recheck})}{P(\text{"yes"} \text{ resp to pos probe})} + \gamma.$$

For the typical experimental situation the probability of making a correct response due to reaching PC with the wrong positive set item is small enough to have a negligible effect on expected reaction times,¹ so

$E(\text{RT yes for pos probe})$

$$= \alpha \frac{\left\{ \sum_{j=PC}^m j \cdot \binom{j-1}{PC-1} q^{PC}(1-q)^{j-PC} \right\} + 2\alpha F P(\text{recheck})}{P(\text{"yes"} \text{ resp to pos probe})} + \gamma \\ = \alpha \frac{\left\{ \sum_{j=PC}^m j \cdot \binom{j-1}{PC-1} q^{PC}(1-q)^{j-PC} \right\}}{(1 - P(\text{reach } NC \text{ with target}))} \\ + 2\alpha F \frac{(1 - P(\text{reach } NC \text{ with target}) - P(\text{reach } PC \text{ with target}))}{(1 - P(\text{reach } NC \text{ with target}))} + \gamma.$$

Since

$$P(\text{reach } PC \text{ with target}) = \sum_{j=PC}^m \binom{j-1}{PC-1} q^{PC}(1-q)^{j-PC},$$

¹ With a set-size S the probability of having a positive probe reach PC with the wrong item would be the same as the probability of a false positive response to a negative probe for set-size $S-1$. This probability has already been assumed to be low. However, its effect is negligibly small because reaching PC with a nontarget item only changes reaction times when PC is not reached with the target item or is reached with the target item after it is reached with the nontarget item. The joint probability of such events is negligibly small for typical parameter values until set size becomes very large (so that PC approaches F).

and

$$P(\text{reach } NC \text{ with target}) = \sum_{j=NC}^m \binom{j-1}{NC-1} (1-q)^{NC} q^{j-NC},$$

$E(\text{RT yes for pos probe})$

$$= \alpha \frac{\left\{ \sum_{j=PC}^m j \cdot \binom{j-1}{PC-1} q^{PC} (1-q)^{j-PC} \right\}}{\left\{ 1 - \sum_{j=NC}^m \binom{j-1}{NC-1} (1-q)^{NC} q^{j-NC} \right\}}$$

$$+ \left[2\alpha F \left\{ 1 - \sum_{j=NC}^m \binom{j-1}{NC-1} (1-q)^{NC} q^{j-NC} - \sum_{j=PC}^m \binom{j-1}{PC-1} q^{PC} (1-q)^{j-PC} \right\} \right] /$$

$$+ \left[\left\{ 1 - \sum_{j=NC}^m \binom{i-1}{NC-1} (1-q)^{NC} q^{j-NC} \right\} \right] + \gamma.$$

For negative probes,

$$E(\text{RT no for neg probe}) = \alpha E(\text{features compared to reach } NC \text{ with all items/“no” resp to neg probe}) + 2\alpha F P(\text{recheck/“no” resp to neg probe}) + \gamma \quad (2.2)$$

Since reaching NC with all items or rechecking always produce a no response to a negative probe,

$E(\text{RT no for neg probe})$

$$= \alpha \frac{\left\{ \sum_{i=NC}^m j \cdot P(\text{last item reaches } NC \text{ on feature } j) \right\}}{P(\text{“no” resp to neg probe})}$$

$$+ 2\alpha F \frac{P(\text{recheck})}{P(\text{“no” resp to neg probe})} + \gamma$$

$$= \alpha \frac{\left\{ \sum_{j=NC}^m j \cdot P(\text{last item reaches } NC \text{ on feature } j) \right\}}{\{1 - P(\text{reach } PC \text{ with at least one item})\}}$$

$$+ 2\alpha F \frac{\{1 - P(\text{reach } PC \text{ with at least one item}) - P(\text{reach } NC \text{ with all items})\}}{\{1 - P(\text{reach } PC \text{ with at least one item})\}} + \gamma.$$

Since

$P(\text{reach } NC \text{ with all items})$

$$= \left\{ \sum_{j=NC}^m \binom{j-1}{NC-1} p^{NC}(1-p)^{j-NC} \right\}^S,$$

and

$P(\text{reach } PC \text{ with at least one item})$

$$= \{1 - P(\text{reach } PC \text{ with no item})\} \\ = \left\{ 1 - \left(1 - \sum_{j=PC}^m \binom{j-1}{PC-1} (1-p)^{PC} p^{j-PC} \right) \right\}^S.$$

and

$P(\text{last item reaches } NC \text{ on feature } j)$

$$= P(\text{all items have reached } NC \text{ by feature } j) \\ - P(\text{all items have reached } NC \text{ by feature } j-1) \\ = \{P(\text{item } i \text{ has reached } NC \text{ by feature } j)\}^S \\ - \{P(\text{item } i \text{ has reached } NC \text{ by feature } j-1)\}^S \\ = \left\{ \sum_{k=NC}^j \binom{k-1}{NC-1} p^{NC}(1-p)^{k-NC} \right\}^S \\ - \left\{ \sum_{k=NC}^{j-1} \binom{k-1}{NC-1} p^{NC}(1-p)^{k-NC} \right\}^S,$$

thus

$E(\text{RT no for neg probe})$

$$= \alpha \left[\left\{ \sum_{j=NC}^m j \right\} \left\{ \sum_{k=NC}^j \binom{k-1}{NC-1} p^{NC}(1-p)^{k-NC} \right\}^S \right. \\ \left. - \left\{ \sum_{k=NC}^{j-1} \binom{k-1}{NC-1} p^{NC}(1-p)^{k-NC} \right\}^S \right] / \\ + \left[\left\{ 1 - \left\{ 1 - \left(1 - \sum_{j=PC}^m \binom{j-1}{PC-1} (1-p)^{PC} p^{j-PC} \right) \right\}^S \right\} \right] \\ + 2\alpha F \left\{ 1 - \left\{ 1 - \left(1 - \sum_{j=PC}^m \binom{j-1}{PC-1} (1-p)^{PC} p^{j-PC} \right) \right\}^S \right\} \\ - \left\{ \sum_{j=NC}^m \binom{j-1}{NC-1} p^{NC}(1-p)^{j-NC} \right\}^S \right] / \\ + \left[\left\{ 1 - \left\{ 1 - \left(1 - \sum_{j=PC}^m \binom{j-1}{PC-1} (1-p)^{PC} p^{j-PC} \right) \right\}^S \right\} \right] + \gamma.$$

Bound and free parameters. None of the parameters in the above equations are completely free. The set size, S , is determined by the experimental situation; NC and PC are calculated from the observed error rates; F , the number of features, may be measurable in some situations and not in others. That leaves p , q , α and γ as free parameters whose values can be varied to obtain the best fit to observed data. Since α and γ represent times, we can place some gross constraints on their values. Gamma cannot be less than minimal physical reaction time; so $\gamma \geq 150$ msec. Alpha must be positive; so $\alpha > 0$. Since $1 - p$ and $1 - q$ represent noise levels (due to encoding errors or similarities), they should both be much smaller than .5 in a typical experiment. In summary, depending upon whether or not F can be measured, the model has four or five parameters each of which is constrained in some way.

Estimating Criteria from Data

One would like to be able to estimate the criteria used with each set size in an experiment from the observed error rates for each set size. Equations (1.1) and (1.2) specify the relation between criteria and expected error rates. However, the *observed* error rates will be an average of the rates produced by the different criteria different subjects are using at different times. Since the criteria for any one subject must be integer valued, one usually will not be able to find criteria that yield the exact average error rates. What one can compute, however, is the average PC and average NC used. To do this though, one needs to make an assumption about the distribution of criterion values used. The simplest distributional assumption to make is that everyone uses one of the two integer values bounding the mean criterion value. With this procedure we can duplicate any observed frequency of false positives, $ERRFP$, through the following procedure.

- (1) Compute $PC1$ and $PC2$ such that
 - (a) $PC1 + 1 = PC2$
 - (b) $PC1$ yields frequency F_{P1} , $PC2$ yields frequency F_{P2} , and $ERRF_{P1} < ERRFP < ERRF_{P2}$
- (2) Use $PC1$, $(ERRF_{P2} - ERRFP) \times 100\%$ of the time.

A similar procedure can be used to estimate the average NC being used in an experiment. The resulting distribution of criteria is the simplest distribution that would yield the desired rates but by no means the only such distribution that would.

Positive criterions versus positive set size. In Fig. 6 the predicted PC (positive criterion) using the above estimation procedure is plotted against S (positive set size) for the case of stimuli with five features. One can see that, for any given rate of false positives and given value of $1 - p$ (probability of an erroneous feature match), PC increases with set size in a

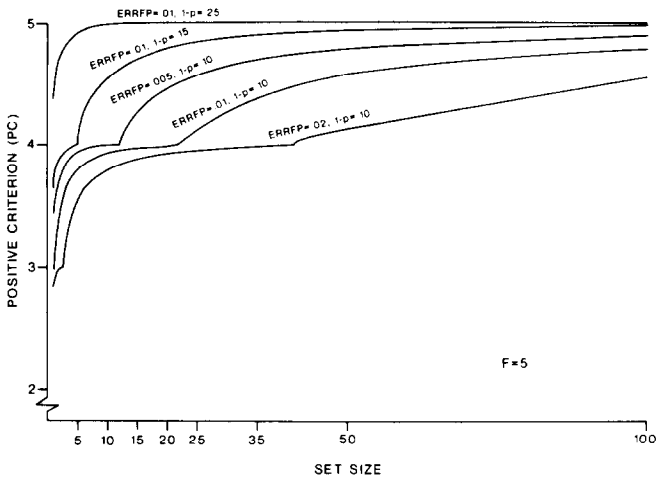


FIG. 6. The model's prediction of how the average positive criterion should increase as a function of positive set size. *ERRFP* is the average rate of false positives observed, and $1 - p$ is the probability that a feature of a negative probe erroneously matches the corresponding feature of a positive item. F , the number of features into which an item is encoded, is assumed to be five for this example.

negatively accelerated manner except for inflection points at each criterion boundary. As $1 - p$ (probability of an erroneous feature match) increases toward 0.50, the value of *PC* increases (except at its asymptote). Similarly, as the observed false positive error rate, *ERRFP*, decreases, the value of *PC* increases.

It is worthwhile to ask how a different assumption about the distribution of criteria would affect the estimate of the mean positive criterion used. For set sizes where the observed error rate can be produced by integer criteria, the estimate should always be the integer value. Thus, the inflection points in Fig. 6 should remain unchanged under any acceptable distributional assumption. As a result, the general form of the relation between set size and the mean positive criterion needed for an observed error rate will always be a negatively accelerated exponential function. What could be influenced by different assumptions is the degree of concavity in the curves between inflection points. It follows that specific predictions about reaction times over short ranges of set sizes are sensitive to the assumption about criterion distributions while the general prediction of a negatively accelerated exponential relation is independent of the assumption.

Predicted reaction times and set size. In Figs. 7 and 8 the relationship between predicted mean reaction times and set size is illustrated for both positive and negative probes. The situation simulated was again one in which the stimuli were coded into five features and the observed error rates were 0.01 (for both false positives and misses). Alpha, the time for a feature

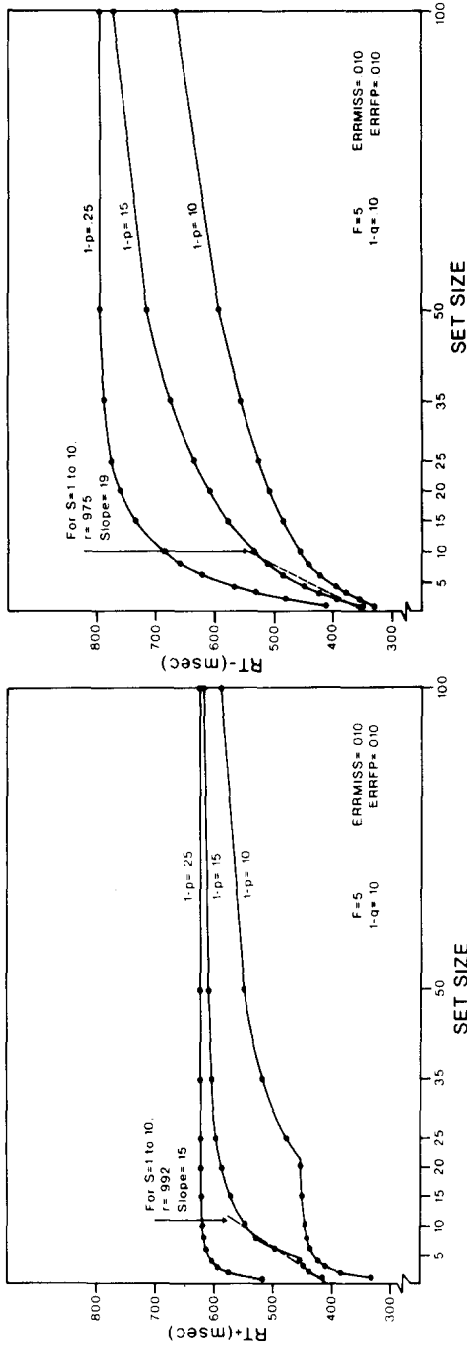


FIG. 7. Predicted reaction times to positive and negative probes as a function of set size for three values of $1 - p$, the probability that a feature of a negative probe erroneously matches the corresponding feature of a positive item.

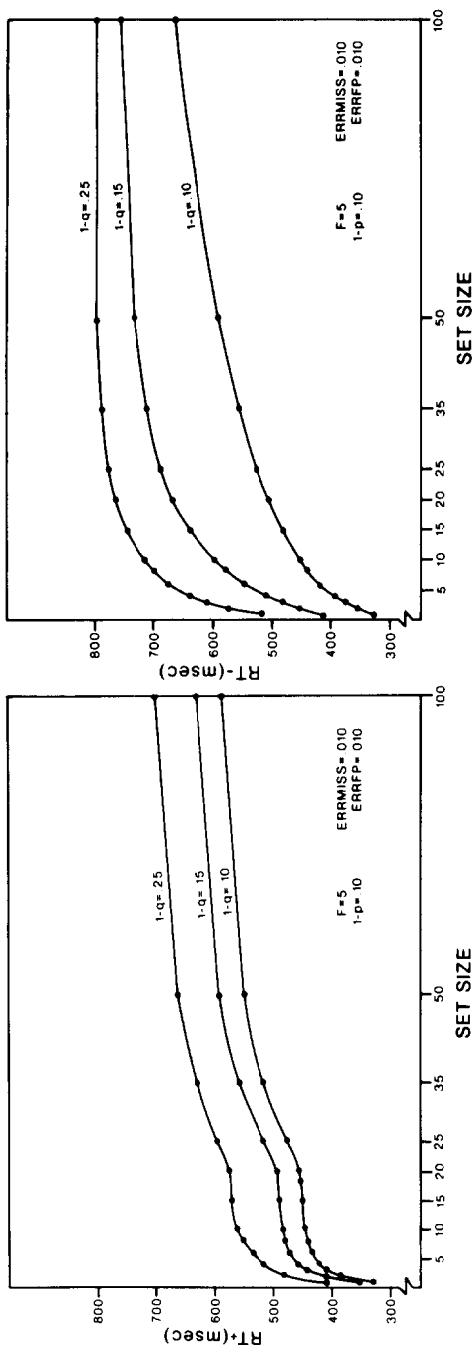


Fig. 8. Predicted reaction times to positive and negative probes as a function of set size for three values of $1-q$, the probability that a feature of a positive probe erroneously mismatches the corresponding feature of the target item in the positive set.

comparison, was arbitrarily set to be 100 msec, and γ , the fixed reaction time, was set at 0 msec.

Regardless of the values of the parameters $1 - p$ and $1 - q$, both positive and negative reaction times increased with set size in a generally negatively accelerated manner until an asymptote was reached. The shape of the positive curve for a particular value of $1 - q$ directly mirrors the way in which the positive criterion increased with set size for that value (See Fig. 6). The increase in expected negative reaction time with set size (while NC remains constant) is due to the increased amount of time needed for the last of the S independent parallel comparisons to reach NC (Equation 2.2). The RT is determined by the slowest, comparison, and with probability comparisons, as S increases, the chances for slow comparison increase.

Increasing the probability of an erroneous feature match (increasing $1 - p$) increased expected positive reaction time by increasing the positive criterion, and increased expected negative reaction time by increasing the expected number of feature comparisons needed to achieve NC (Fig. 7). Increasing the probability of an erroneous feature mismatch (increasing $1 - q$) increased expected positive reaction time by increasing the expected number of feature comparisons needed to achieve PC , and increased expected negative reaction time by increasing the negative criterion (Fig. 8).

PREDICTING OBSERVED ITEM RECOGNITION PHENOMENA

Linear slopes. A generally linear increase in both positive and negative reaction times with increases in set size has been observed in most major studies with small ranges in set size, e.g., two to 12. In Fig. 7 we have plotted a least-squares line for one set of positive reaction times generated by our model over a similar range of set sizes. The fit is good even though the reaction-time function is in reality nonlinear. Furthermore, the slope of the least-squares line will decrease as larger and larger set sizes are included, a prediction that is born out by some data. While the tradition has been to fit a straight line to a small number of means and achieve a linear fit, a negatively accelerated exponential would fit most such data equally well and would fit the data from many studies with large ranges of set sizes better. With larger set sizes either negatively accelerated trends (Wingfield & Branca, 1970) or flatter slopes (on initial presentations) (Juola et al., 1971) have been found. Flatter or negatively accelerated trends have also been apparent in many studies with smaller set sizes, particularly on negative probes similar to items in the positive set (Dumas, 1972; Ellis & Chase, 1971; Foss & Dowell, 1971; Lively, 1972). From Fig. 7 one can see that the negative reaction times predicted by our model also increase linearly over the range of set sizes from one to 10. As one would expect from most empirical research, the slope of the negative curve is very similar to the slope of the positive curve.

The slopes and intercepts, of course, would vary depending upon parameter values and error rates. Which response is quicker in a particular experiment would depend upon the relative values of $1 - p$ and $1 - q$ and the relative rates of false positives and misses. In general, the greater the number of observed false positives relative to misses, and the smaller $1 - p$ (probability of an erroneous feature match) relative to $1 - q$ (probability of an incorrect feature mismatch), the faster will be the positive response times relative to the negative response times. However, regardless of these factors the model predicts that negative reaction times will eventually exceed positive reaction times as set size increases.² This prediction of the model has been supported by data showing that for large S positive, RT curves flatten out more than negative RT curves (Joula et al., 1971; Wingfield & Branca, 1970).

Speed-accuracy tradeoff. As Fig. 6 illustrates, PC (positive criterion) decreases as the false positive error rate increases. As a result, positive reaction times will decrease as the false positive rate increases. Similarly, NC (negative criterion) decreases as the miss rate increases; so negative reaction times will decrease as the miss rate increases. These predicted changes in reaction times as a function of error rates are shown in Fig. 9.

If one plots the best fitting lines between set size and our model's predicted times, one finds decreasing intercepts with increasing error rates. One also finds decreasing slopes for negative reaction times and for positive reaction times if slopes are computed over a large range of set sizes. For any small range, however, as the false positive rate decreases, the slope of the positive reaction time curve may remain unchanged or even increase depending upon the concavity of the local curve. Again, one should remember that the local concavity depends upon the distribution of criteria among subjects (or at least the assumed distribution) and is not determined by the model.

² When PC reaches F , PC can no longer be increased as S increases. As a result, the expected positive reaction time will asymptote while the rate of false positives rises with further increases in S . From (2.1) we can see that expected positive reaction time will asymptote at

$$E(\text{RT yes for pos probe}) = \alpha F \frac{q^F}{1 - \sum_{j=NC}^F \binom{j-1}{NC-1} (1-q)^{NC} q^{j-NC}} + 2\alpha F \frac{1 - q^F - \sum_{j=NC}^F \binom{j-1}{NC-1} (1-q)^{NC} q^{j-NC}}{1 - \sum_{j=NC}^F \binom{j-1}{NC-1} (1-q)^{NC} q^{j-NC}} + \gamma.$$

On the other hand, the expected negative reaction time will continue to increase with set size until the probability of reaching NC for all S items becomes zero. At this point, from Equation (2.2), $E(\text{RT no for neg probe}) = 2\alpha F + \gamma$.

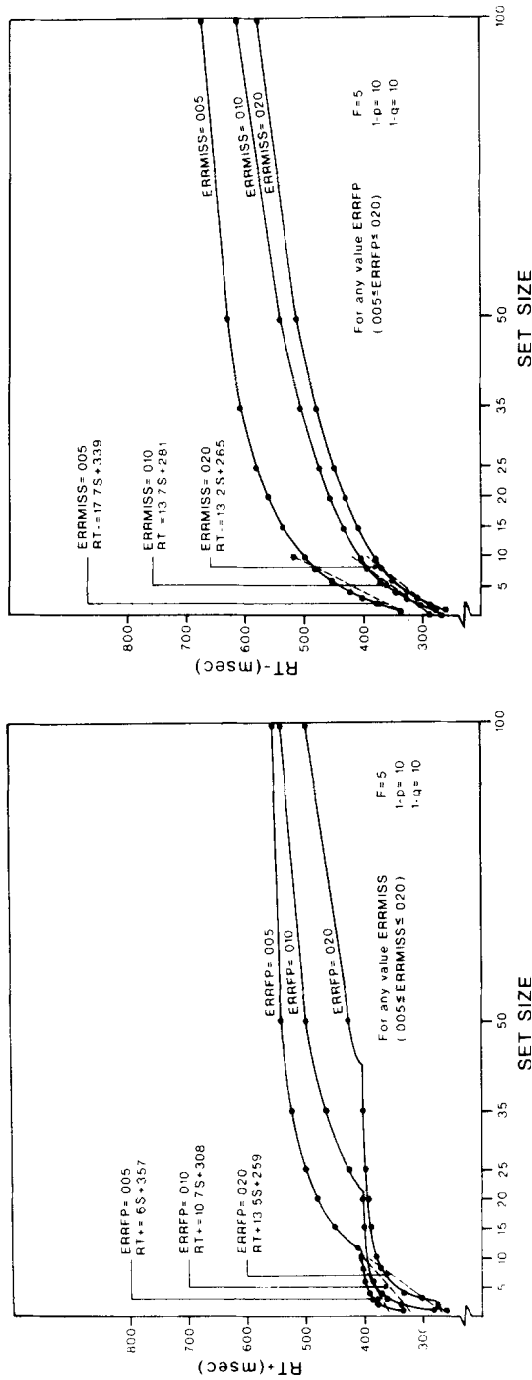


FIG. 9. How the predicted relation between reaction times and set size changes as a function of error rates. The predicted slopes are plotted over set sizes one to 10.

Most researchers have found slope and intercept changes consistent with these predictions. Intercepts always seem to decrease as error rates increase (Swanson & Briggs, 1969; Briggs & Swanson, 1970; Lyons & Briggs, 1971), and several workers have found that slopes decrease with increasing error rates (Weaver, 1972; Banks & Atkinson, 1973).

Under the model, a subject sets his criteria by an interactive process that eventually achieves the desired error rates. In setting his criteria, the subject assumes, we proposed, that he could perform errorlessly by taking enough time to recheck every item. He changes his positive criterion as set size changes to maintain the desired error rate. Therefore, if $PC + NC > F$, error rates should not change much with set size. Suppose, however, that the subject is instructed to guess instead of rechecking. Since the probability of a subject not reaching a criterion and guessing would increase with set size and the chances of guessing wrong would increase with set size or at least not decrease, one would expect to find error rates increasing with set size. Banks and Atkinson (1973) have found evidence of such an effect.

Finally, according to the model, one could also find error rates increasing as a function of set size when $PC + NC < F$. These increases would asymptote when PC increased enough with S so that $PC + NC > F$. But for large set sizes then $PC = F$, the rate of false positives, would again increase.

Serial position effects. As long as $1 - p$ and $1 - q$ do not vary with serial position, the model predicts that positive reaction times will not be affected by the serial position of the positive item in the to-be-remembered set. But, if the set is lengthy and not overlearned, one would expect to find more encoding errors in the stored representations of items in the middle of the to-be-remembered list. In other words, the value of $1 - q$ (probability of an erroneous feature mismatch) would be larger for serial positions in the middle of a quickly memorized list of moderate length. This results in a prediction of primacy and recency effects for positive reaction times on such lists. Such a prediction is consistent with available experimental evidence (Corcoran, 1971; Burrows & Okada, 1971).

Memory span and scanning time. Cavanagh (1972) has reported that the amount of space an item seems to occupy in immediate memory is linearly related to the scanning rates observed for the same kind of item (e.g., digits, colors, letters, words, shapes, nonsense syllables). If one assumes that the amount of space an item occupies is proportional to how many features must be stored to unambiguously define the item, Cavanagh's findings are consistent with our proposed model. As F increases, the overall slopes of both positive and negative reaction times against positive set size increase. For positive RTs, the slope depends upon the change in PC with set size. To maintain a constant false positive rate, PC is increased with S , lowering the chances of a false positive on any one item. The greater the number of

features, F , the greater will be the increase needed in PC to produce the desired lowering. For negative RTs, the slope depends upon the increase with set size of the expected number of features checked before NC is achieved for all items. For a fixed miss rate, the greater the number of features, F , the greater will be the increase with set size of the number of features checked.

Duplicate target effect. Sternberg (1975) and Baddeley and Ecob (1973) have found that the duplication of an item in the positive set decreases the positive reaction time for that target but does not affect other times. Our proposed model also predicts this. The chances of reaching PC at an early time are enhanced by having more than one replicate of the target being compared with the target simultaneously.

We have seen that the serial-feature-matching parallel-processing model can predict the basic findings of memory scanning recognition experiments. The elementary information processing operations a subject uses are unambiguous and plausible. The recognition algorithm the model suggests is efficient and produces the speed-accuracy trade off required by the situation. However, the distributional assumptions necessary to estimate criteria from error rates, the problems in measuring the number of features a subject processes, and the problems in measuring the probabilities of feature matching errors [$(1 - p)$ and $(1 - q)$] make it difficult to evaluate the model on its specific quantitative predictions and their goodness of fit. Rather, major support for the model must come from its ability to predict such observed general relationships as the set size-reaction time effects, speed-accuracy trade offs, serial position effects, and the probe similarity effects that stimulated the development of the model. Nevertheless, let us apply the model to the data of our experiment and one other to show that acceptable quantitative predictions can be made.

FITTING OBSERVED DATA

To fit the data from our experiment quantitatively with the model, we must estimate α (the average time needed for a feature comparison), γ (the average base reaction time), $1 - p$ (the probability of an erroneous feature match), and $1 - q$ (the probability of an erroneous feature mismatch). Of course, in fitting group data, one is fitting an average subject rather than any one individual. Using the average observed error rates of .018 for false positives and .022 for misses, the parameter values were chosen by an interactive least-squares procedure that could not be guaranteed to produce an optimal fit. The predicted values are compared with the observed values in Fig. 10. The discrepancies between predicted and observed are small percentages of the observed standard deviations indicating a reasonable fit. The derived values for $1 - p$ and $1 - q$ were .05 and .11, respectively. In other words, the apparent likelihood that a feature

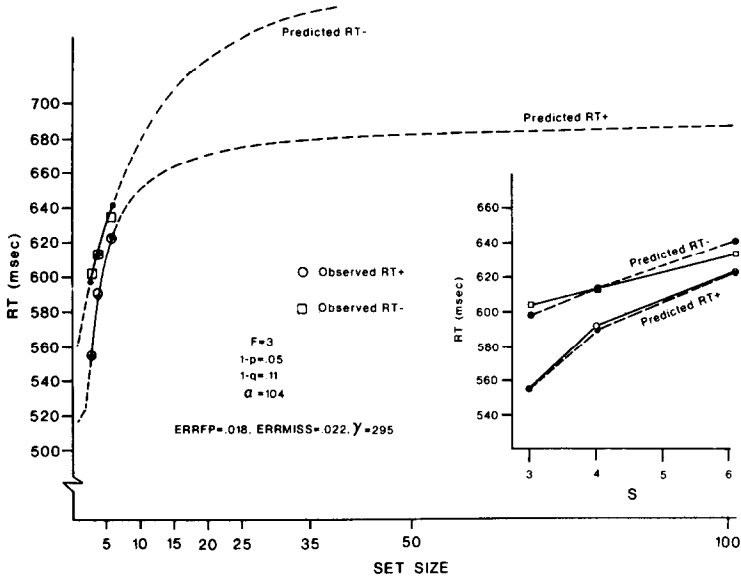


FIG. 10. A comparison of predicted and observed mean reaction times for the probe similarity experiment.

in a positive probe would mismatch its corresponding feature in the positive set was .11, while the apparent likelihood of an erroneous feature match with a negative probe was .05. Actually, for the set of similar negative probes used in the experiment, the percentage of matching features was 15%; so the estimate of $1 - p$ was surprisingly low. One reasonable possibility is that our definition of identical features was too liberal. In other words, two apparently identical vowel or consonant sounds may actually be pronounced noticeably differently because of the differing contexts.

Many of the ways in which the predictions of the model differ from those of exhaustive scanning models become most obvious when set sizes are very large. The parallel-processing serial-feature-matching model, for example, predicts that the increase in positive RT with set size is negatively accelerated and reaches an asymptote for large set sizes while negative RT curves, though also negatively accelerated, can continue to increase indefinitely with set size. It seems appropriate, therefore, to try and fit the model to data from an experiment with large set sizes.

To predict reaction times we must have measures of error rates and an estimate of the number of features in the stimuli. An experiment for large set sizes where such information is available is Atkinson and Juola's (1974) experiment with set sizes of 16, 24, and 32. In this experiment the stimuli were one syllable nouns; so again we will represent the stimuli with three features: beginning consonant, vowel, and end consonant. The stimuli

were presented visually; so the assumption must be that each stimulus was encoded into the features which then could be processed serially. The initial error rates were about 3% false positives and 18% misses.

Using this information we could predict reaction times in one of two ways. We could set the values of the remaining parameters ($1 - p$, $1 - q$, α , and γ) to maximize the goodness of fit, or, since the stimuli were one syllable words as in the probe similarity experiment, we could assume that the parameter values derived from our experiment would be adequate approximations. This should certainly be true of α , the rate of feature scanning if we adopt the same functional definition of a feature. However, $1 - p$ and $1 - q$ are more suspect. The probability of encoding errors and chance matches might be quite different in a visual experiment where similarity is not being manipulated. Nevertheless, let us first try to fit Atkinson and Juola's data *with all the parameters fixed to the values derived for the probe similarity experiment except for γ , the base reaction time, which would be different.*

The parallel-processing serial-feature-matching model's predictions were compared with Atkinson and Juola's data for two situations, the initial presentation of probes and repeated presentations of probes. The parameters F , p , q , and α were not altered to obtain the best fit; rather the values derived in fitting our probe similarity data were used. The base reaction time, γ , was increased to 520 msec for the initial presentation condition and 450 msec for repeated presentations. The criteria used were varied to obtain the observed error rates in each condition. In Fig. 11 the

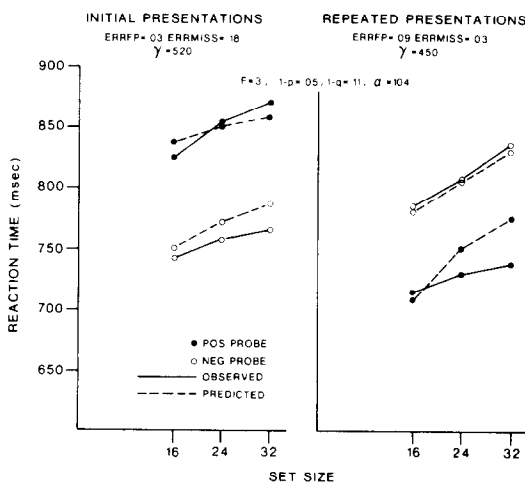


FIG. 11. A comparison of the reaction times observed in Atkinson and Juola's (1974) experiment with the predictions of the parallel-processing serial feature-matching model. The values used for the model's parameters were the same as those used to predict the results of the similarity experiment (Fig. 10) with the exception of the base reaction time ($GAMMA$) which was recomputed for the new experimental situation.

predicted and observed data are shown. What is interesting to note is how the relative positions of the RT lines in the four conditions were predicted accurately by our model. In other words, the changes in the pattern of positive and negative RTs in their experiment can be explained by the model solely in terms of changes in error rates and the corresponding speed-accuracy trade off. The increase in false positive error rate produced a decrease in the model's computed *PC* (Equation 1.1), and the decrease in misses produced an increase in the computed *NC* (Equation 1.2). In fact, a subject's $1 - p$ and $1 - q$ probably also would change for repetition trials, and better fits could be obtained with altered values for these parameters.

SUMMARY

In this paper we have developed a model of memory scanning based on a conception of memory as pages of cells associated in list structures and accessed by a special recognition processor. We have proposed that individual items are recognized by serial feature-matching with stored representations. An item is recognized when enough features match so that the positive criterion is reached; an item is rejected when enough features mismatch so that the negative criterion is reached; a rechecking process is undertaken if neither criterion is reached.

According to this model, set membership is recognized by comparing a probe in this way with every item in the set at the same time. Thus, the model is called a parallel-processing serial-feature-matching model. It was proposed that desired error rates would determine the positive and negative criteria. In order to maintain a constant rate of false positives, the positive criterion would have to increase with set size. The negative criterion would not have to change. It was shown that this model yielded predictions consonant with observed empirical data about the relation between positive and negative reaction times and set size, about serial position curves, and about speed-accuracy trade offs. Most importantly, the model predicted observed probe similarity effects.

The similarity of a negative probe to items in the positive set was discovered to affect recognition times in several ways predicted by the model. The serial location of the similarity within an item affected negative recognition times, but the number of items to which the probe was similar had no main effect. It was argued that these and other observed results provided strong support for the proposed model.

REFERENCES

- Anderson, J. A. A theory for the recognition of items from short memorized lists. *Psychological Review*, 1973, **80**, 417-438.
- 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*. New York: Wiley, 1974, 101–146.
- Atkinson, R. C., Holmgren, J. E., & Juola, J. F. Processing time as influenced by the number of elements in a visual display. *Perception and Psychophysics*, 1969, **6**, 321–327.
- Atkinson, R. C., & Juola, J. F. 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, 1974, Vol. 1.
- Baddeley, A. D., & Ecob, J. R. Reaction time and short term memory: A trace strength alternative. Technical Report. No. 12, 1970, Center for Human Information Processing, University of California, San Diego.
- 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.
- Banks, W., & Atkinson, R. *Accuracy and speed strategies in scanning active memory*. Paper presented at Psychonomic Society Meetings, November 1973, St. Louis.
- Briggs, G. E., & Blaha, J. Memory retrieval and central comparison times in information processing. *Journal of Experimental Psychology*, 1969, **79**, 395–402.
- Briggs, G. E., & Swanson, J. M. Encoding, decoding, and central functions in human information processing. *Journal of Experimental Psychology*, 1970, **86**, 296–308.
- Burrows, D. B., & Okada, R. Serial position effects in high-speed memory search. *Perception and Psychophysics*, 1971, **10**, 305–308.
- Cavanagh, J. P. Relation between the immediate memory span and the memory search rate. *Psychological Review*, 1972, **79**, 525–530.
- Chase, W. G., & Calfee, R. C. Modality and similarity effects in short-term recognition memory. *Journal of Experimental Psychology*, 1969, **81**, 510–514.
- Checkosky, S. Speeded classification of multidimensional stimuli. *Journal of Experimental Psychology*, 1971, **87**, 313–318.
- Corcoran, D. W. J. *Pattern Recognition*. Baltimore: Penguin, 1970.
- Crowder, R. G., & Morton, J. Precategorical acoustic storage (PAS), *Perception and Psychophysics*, 1969, **5**, 365–373.
- Dumas, J., Gross, E., & Checkosky, S. Effects of attribute probability in a memory search task. *Journal of Experimental Psychology*, 1972, **93**, 327–332.
- Ellis, S. H., & Chase, W. G. Parallel processing in item recognition. *Perception and Psychophysics*, 1971, **10**, 379–384.
- Foss, D. J., & Dowell, B. E. High-speed memory retrieval with auditorily presented stimuli. *Perception and Psychophysics*, 1971, **9**, 465–468.
- Gregg, L. W., & Simon, H. A. Process models and stochastic theories of simple concept formation. *Journal of Mathematical Psychology*, 1967, **4**, 246–277.
- Juola, J. F., & Atkinson, R. C. Memory scanning for words versus 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 and Psychophysics*, 1971, **10**, 8–14.
- Kennedy, R. A., & Hamilton, D. Time to locate probe items in short lists of digits. *American Journal of Psychology*, 1969, **82**, 272–275.
- Lively, B. L., & Sanford, B. J. The use of category information in a memory search task. *Journal of Experimental Psychology*, 1972, **93**, 379–385.
- Lyons, J. J., & Briggs, G. E. Speed-accuracy trade-off with different types of stimuli. *Journal of Experimental Psychology*, 1971, **91**, 115–119.
- Murdock, B. B., Jr. A parallel-processing model for scanning. *Perception and Psychophysics*, 1971, **10**, 289–291.
- Naus, M. J., Glucksberg, S., & Ornstein, P. A. Taxonomic word categories and memory search. *Cognitive Psychology*, 1972, **3**, 643–54.

- Raeburn, V. *Priorities in item recognition*. Unpublished dissertation. Yale University, 1972.
- Shavell, S. K., & Atkinson, R. C. A theoretical comparison of list scanning models. *Journal of Mathematical Psychology*, 1974, **11**, 79–106.
- Sternberg, S. High speed scanning in human memory. *Science*, 1966, **153**, 652–654.
- Sternberg, S. Two operations in character recognition. *Perception and Psychophysics*, 1967, **2**, 45–53.
- Sternberg, S. Memory Scanning: Mental processes revealed by reaction time experiments. *American Scientist*, 1969, **57**, 421–457.
- Sternberg, S. Memory scanning: new findings and current controversies. *Quarterly Journal of Experimental Psychology*, 1975, **27**, 1–32.
- Swanson, J. M., & Briggs, G. E. Information processing as a function of speed versus accuracy. *Journal of Experimental Psychology*, 1969, **81**, 223–229.
- Theios, J., Smith, P. G., Haviland, S. E., Traupmann, J., & Moy, M. C. Memory scanning is a serial, self-terminating process. *Journal of Experimental Psychology*, 1973, **97**, 323–336.
- Thorndike, E. L., & Lorge, I. *The Teachers word book of 30,000 words*. New York: Teachers College, 1944.
- Townsend, J. T. A note on the indentifiability of parallel and serial processes. *Perception and psychophysics*, 1971, **10**, 161–163.
- Weaver, F. S. *Retrieval processes in short-term memory*. Unpublished dissertation. University of Delaware, 1972.
- Wingfield, A., & Branca, A. A. Strategy in high speed memory search. *Journal of Experimental Psychology*, 1970, **83**, 63–67.
- Wood, C. *The complete rhyming dictionary and poet's craft book*. Garden City, NY: Garden City Books, 1936.

(Accepted October 23, 1975)