Psychological Review

VOLUME 85 NUMBER 2 MARCH 1978

A Theory of Memory Retrieval

Roger Ratcliff
University of Toronto, Ontario, Canada

A theory of memory retrieval is developed and is shown to apply over a range of experimental paradigms. Access to memory traces is viewed in terms of a resonance metaphor. The probe item evokes the search set on the basis of probe–memory item relatedness, just as a ringing tuning fork evokes sympathetic vibrations in other tuning forks. Evidence is accumulated in parallel from each probe–memory item comparison, and each comparison is modeled by a continuous random walk process. In item recognition, the decision process is self-terminating on matching comparisons and exhaustive on nonmatching comparisons. The mathematical model produces predictions about accuracy, mean reaction time, error latency, and reaction time distributions that are in good accord with experimental data. The theory is applied to four item recognition paradigms (Sternberg, prememorized list, study–test, and continuous) and to speed–accuracy paradigms; results are found to provide a basis for comparison of these paradigms. It is noted that neural network models can be interfaced to the retrieval theory with little difficulty and that semantic memory models may benefit from such a retrieval scheme.

At the present time, one of the major deficiencies in cognitive psychology is the lack of explicit theories that encompass more than a single experimental paradigm. The lack of such theories and some of the unfortunate consequences have been discussed recently by Allport (1975) and Newell (1973). Two important points are made by Newell: First, research in cognitive psychology is motivated and guided by phenomena (e.g., those of the Sternberg paradigm and the Brown–Peterson paradigm). Second, attempts to theorize result in the construction of binary oppositions (e.g., serial vs. parallel, storage vs. retrieval, and semantic vs. episodic). Thus, an appreciable portion of research in cognitive psychology can be described as testing binary oppositions within the framework of a particular experimental paradigm. This style of research does not emphasize or encourage theory construction; but without theory, it is almost impossible to relate experimental paradigms or to substantiate claims that the same processes underlie different experimental paradigms. Furthermore, the concern with binary oppositions tends to obscure the more interesting aspects of data, such as the form of functional relations.

The theory presented in the present article is concerned with providing an account of processes underlying retrieval from memory. Perhaps the most intensively investigated class

This research was supported by Research Grants APA 146 from the National Research Council of Canada and OMHF 164 from the Ontario Mental Health Foundation to Bennet B. Murdock, Jr. I would like to thank Ben Murdock for his support, help, and criticism and the use of data from his published and unpublished experiments. I would also like to thank the following: Ron Okada and David Burrows for use of their data; Peter Liepa for programming Experiment 2; David Andrews for help with the mathematical development; Gail McKoon, Bill Hockley, and Bob Lockhart for useful discussion; and Howard Kaplan for general programming assistance.

Requests for reprints should be sent to Roger Ratcliff, who is now at the Department of Psychology, Dartmouth College, Hanover, New Hampshire 03755.

of memory retrieval paradigms has been the item recognition paradigms, for example, the Sternberg paradigm (Sternberg, 1969b), the study–test paradigm (Murdock & Anderson, 1975), the continuous recognition memory paradigm (Okada, 1971), and the prememorized list paradigm (Burrows & Okada, 1975). In this area of research, several models have been developed to deal with results from each paradigm, but each model is inconsistent with some data (Ratcliff & Murdock, 1976; Sternberg, 1975). Also, the models are paradigm specific and do not generalize across paradigms. The theory presented here has been designed to deal with all aspects of the data in each paradigm (accuracy, response latency, and latency distributions) and has been designed to apply over a range of paradigms. In the first section, a qualitative account of the theory is presented; in the second section, the mathematical model is introduced and briefly described. The article is written so that an understanding of the mathematics is not necessary for comprehension of the theory. A complete development of the mathematical model is presented in the Appendix. In the third section, the theory is applied to five experimental paradigms, in which both response latency and accuracy are measured, and the same set of processing assumptions is used to account for performance in all of the paradigms. The fourth section relates the theory to neural network models, semantic memory models, and propositional models.

A Description of the Theory

In this section, I present an overview of what I term the *retrieval theory*. I have not presented any mathematical development in this section in order to give a qualitative account of the theory, which can be understood independently of the mathematical formulation.

Let us consider a typical item recognition task in which a group of items in memory has been designated the memory search set and a single probe item is presented for testing. According to the theory, the probe is encoded and then compared with each item in the search set simultaneously (i.e., in parallel). Each individual comparison is assumed to be accomplished by a random walk (more specifically, a diffusion) process. A decision is made when any one of the parallel comparisons terminates with a match (self-terminating on positives) or when all of the comparisons terminate with a nonmatch (exhaustive on negatives). When a decision has been made, a response is initiated and performed. Figure 1 illustrates the overall retrieval scheme.

Search or Evoked Set

In some models developed to account for item recognition data, it is assumed that the only memory items accessed in the comparison process are the items in the experimenter-designated search set, that is, the positive set (Schneider & Shiffrin, 1977; Sternberg, 1966). There is a problem with such models because it can be shown that items outside this positive set may be accessed. For example, Atkinson, Herrmann, and Wescourt (1974) report two experiments in which the recency of negative probes was manipulated. The first used a Sternberg (1966) procedure. Subjects were presented with a memory set of either 2, 3, 4, or 5 words, which was followed by a probe word. It was found that latency of response to a negative probe word was greater and its accuracy lower the more recently it had occurred. The second experiment used a 25-word study list and a 100-word test list with no repeated negatives. After 50 test words, subjects were required to read written instructions; 10 words in the instructions served as negative items in the remainder of the test list. Latencies for the negative items that had been in the instruction set were significantly longer than latencies for other negative items. Therefore, to explain the processes involved in item recognition, a model must include some mechanism that allows items from outside the experimenter-designated search set to be accessed in the comparison process.

I will adopt the notion of search set suggested by a resonance metaphor (Neisser, 1967, p. 65). Suppose each item in memory corresponds to a tuning fork and the informational basis on which comparisons are made corresponds to frequency. Then, the comparison process can be viewed as the probe tuning fork ringing and evoking sympathetic vibrations in memory tuning forks of similar frequencies. Whether or not the probe evokes sympathetic vibrations in a memory-set item depends only

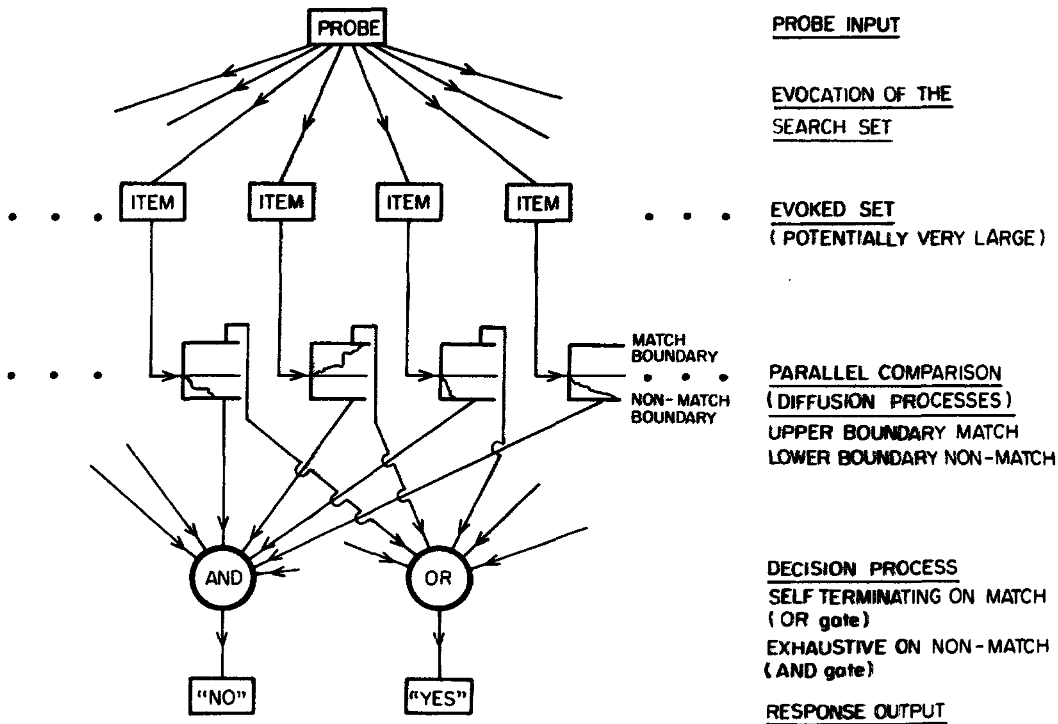

Figure 1. An overview of the item recognition model.

on the similarity of frequencies and is independent of whether or not the memory item is in the experimenter-designated memory set. The amplitude of resonance drives the random walk comparison process: The greater the amplitude (better match), the more bias there is toward a "yes" response; the smaller the amplitude (poorer match), the more bias toward a "no" response (see Figure 2).

So far, I have not indicated how many memory items can be contacted by the probe. It turns out that predicted performance will be little affected even if the number of items contacted is large, perhaps comparable to the number of items in memory. The reason is that "no" responses are based on exhaustive processing of items in the search set, so the decision time for a "no" response is a function of the group of slowest individual nonmatch comparisons. Thus, many fast finishing processes may enter the decision process without affecting performance.

I will designate items giving rise to this slower group of latency- and accuracy-determining processes as the *search set*. In later applications, the search set will usually be approximated by the memory set. A demonstration that the evoked set (defined as all

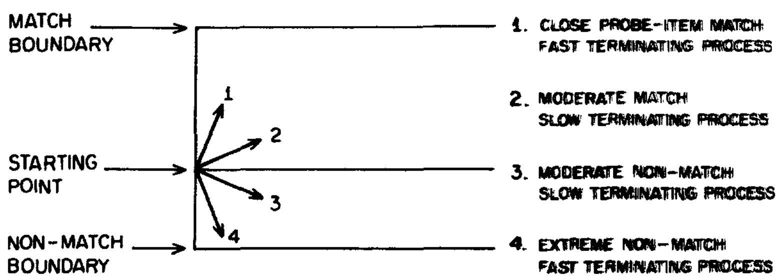

Figure 2. The relation between relatedness and amount of match in the diffusion random walk process. (Relatedness varies from high [Process 1] to low [Process 4].)

items contacted by the probe) can theoretically be very large is presented later in this section. There, it is shown that many probe–item comparisons with very small resonant amplitudes can be added to the few performance-determining processes (probe–nontarget item comparisons with larger resonances) without changing latency or accuracy.

It is worth noting at this point that the retrieval theory outlined in Figure 1 in conjunction with the resonance metaphor can be considered an implementation (for item recognition) of content-addressable memory models or broadcast models (Bobrow, 1975). The notion of resonance also suggests similarities to the pandemonium model (discussed in Neisser, 1967) for feature encoding.

Information Entering the Comparison Process

Here I examine in more detail factors that influence the "amplitude of resonance." Seward (1928, cited in Woodworth, 1938) performed an experiment in which subjects were presented with 30 patterns ("fancy papers"). Following 10 minutes of unrelated activity, a yes–no recognition test was given in which 10 test items were identical to the study items, 10 rather similar, 10 slightly similar, and 10 very different. It was found that as similarity decreased, the proportion of "yes" responses decreased and latency of "yes" responses increased. Similarly, as similarity decreased, the proportion of "no" responses increased and latency of "no" responses decreased.

More recently, Juola, Fischler, Wood, and Atkinson (1971) performed an experiment with words as stimulus and probe items. Stimulus words were memorized to a criterion of perfect recall, and test words were presented sequentially at a fixed rate. There were three types of negative items: homophones and synonyms of target words and neutral words. It was found that synonyms were 60 msec slower than neutral words, and homophones were 120 msec slower; although when homophones were broken down into two types, visually similar and visually dissimilar to target items, differences were 200 msec and 40 msec, respectively. These results show that semantic, visual, and phonemic similarities between probe and memory items affect recognition performance.

Schulman (1970) studied a complementary task, that is, one requiring the subject to make the judgment as to whether a probe word was identical to, a homonym of, or a synonym of a presented word. Ten words were presented to the subject (Sternberg varied-set procedure), the judgment condition was cued, and the probe word was presented. Accuracy, as measured by d', and latency behaved in much the same way: d' was highest and latency shortest for the identical condition, d' was lower and latency longer for the homonym condition, and d' was lowest and latency longest for the synonym condition.

Some similarity effects can be quite unexpected. For example, Morin, DeRosa, and Stultz (1967) and Marcel (1977) have shown that the more numerically remote a negative probe is from the memory set, the faster the response (e.g., Probe 1 is faster than Probe 6 to Memory Set 798). Thus, numerical relatedness can enter the recognition process.

All of these results suggest that similarity between probe and memory-set items is a major determinant of recognition performance, and that performance is best conceived in terms of discriminability between memory-set items and distractors.

Further, the fact that so many factors influence item recognition performance raises the question, How do we conceptualize the structure of the memory trace? Tulving and Bower (1974) have suggested that the most generally accepted conception of the trace is that of a collection of features or a bundle of information. In addition, Tulving and Bower describe methods that have been used to study the features or combinations of information in the trace; some of the methods are similar to those described earlier in this section. However, they conclude by stressing that inferences about trace structure can only be drawn in the context of a process model. The theory presented here specifies a process model in great detail but does little more than describe the memory trace as a bundle of information while allowing some quantitative assessment of the extent to which certain (possibly featural) information is represented in the memory trace.

It seems that many qualitatively different types of information are contained in the memory trace. In order to determine how well a probe matches an item in the memory set, the interaction of the trace and probe information must be assessed. I will use the single term *relatedness* to describe the amount of match or size of resonance. Thus, in Figure 2, Processes 1 to 4 have relatedness varying from high to low. Relatedness will be assumed to vary over items because whenever a group of nominally equivalent items is memorized, some items are better remembered than others (cf. strength theory and attribute theory; Murdock, 1974). In quantification of the mathematical model, relatedness has variance (a normal distribution of relatedness is assumed), which allows the calculation of a d' measure of discriminability. Variability in relatedness is shown later to be necessary to ensure that asymptotic accuracy is not infinite. It is possible to assess the extent to which different kinds of information contribute to the memory trace, for example, by calculating d' discriminability values for different kinds of negative items in the experiment by Juola et al. (1971) described earlier.

So far, I have made two main points. First, the resonance metaphor is used in order to emphasize that we are dealing with a parallel interaction between the probe and the representation of the memory-set items. The retrieval scheme suggested by the resonance metaphor allows performance to be affected by items outside the memory set, although this feature will not be used here in any of the applications of the mathematical model and is included only for sufficiency. Second, all trace information is mapped onto a unidimensional variable, that of relatedness, and relatedness is the dimension on which discriminability between positive and negative items is assessed.

Comparison Process

Comparison of a probe to a memory-set item proceeds by the gradual accumulation of evidence, that is, information representing the goodness of match, over time. It is easiest to conceptualize the comparison process as a feature-matching process in which probe and memory-set item features are matched one by one. A count is kept of the combined sum of the number of feature matches and nonmatches, so that for a feature match, a counter is incremented, and for a feature nonmatch, the counter is decremented. The counter begins at some starting value Z, and if a total of A counts are reached, the probe is declared to match the memory-set item ($A - Z$ more feature matches than nonmatches). But if a total of zero counts are reached, an item nonmatch is declared. This process is called a *random walk* (see top panel of Figure 3).

Relatedness corresponds to the ratio of the number of matching features to the total number of features. Two probe–memory-set item comparisons involving identical relatedness may still have different comparison times. Suppose that in one process, the features are ordered so that the matching features are processed first, whereas in the second process, nonmatching comparisons are carried out first. Then, if the number of feature matches required for an item match is sufficiently less than the total number of features, the two comparison processes will give quite different comparison times (the first fast and the second slow) or perhaps will give different results (item match and nonmatch, respectively). In general, there is variation in comparison time, depending on order of feature comparisons and the distribution of matching and nonmatching features within that order. Therefore, two sources of variance have now been identified: variance in relatedness and variance in the comparison process.

Diffusion Process

In the mathematical formulation of the theory, the comparison process is represented by the diffusion process, the continuous version of the random walk in both temporal and spatial coordinates. The diffusion process has the advantage that the mean rate of accumulation of information and the variance in that rate are independent parameters. In the feature-matching model, the process is essentially binomial, and mean and variance of a binomial are correlated; whereas the diffusion process is essentially normal, and mean and variance are independent. The third panel of Figure 3 illustrates the diffusion process.

Diffusion and relatedness. The critical assumption made is that drift rate in the diffusion process (in Figure 3, average distance traveled vertically per unit time) is equal to

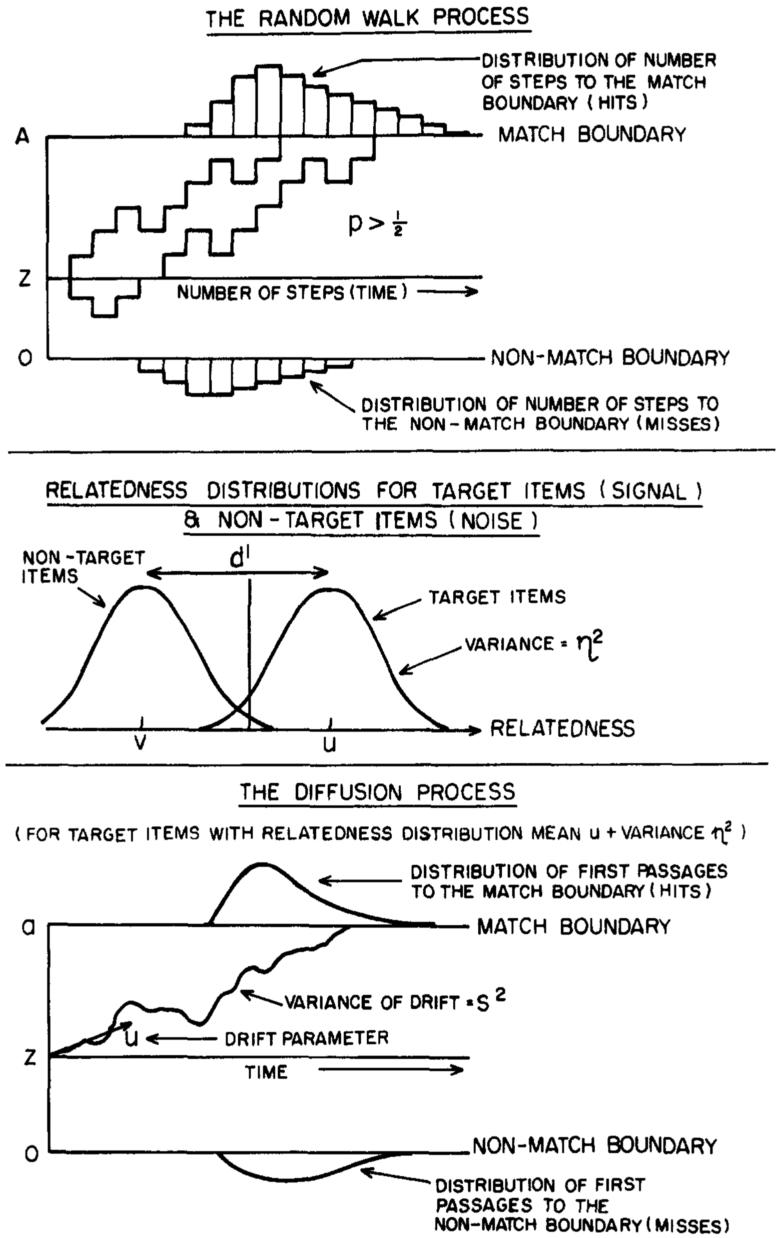

Figure 3. An illustration of the random walk and diffusion process, together with relatedness distributions that drive the diffusion process.

the relatedness value. Thus, any particular probe–memory-set item comparison has drift equal to the relatedness value of that comparison, so that, for example, the greater the probe–memory-set item relatedness, the faster the match boundary is reached. The second panel of Figure 3 shows two distributions of relatedness, one for probe–target item comparisons and one for probe–nontarget item comparisons.

Parameters for the comparison process. Figure 3 shows the six parameters of the comparison process. Over the five paradigms examined, variance in relatedness (η^2) and variance of drift in the diffusion process (s^2) were kept constant. Thus, there were four free comparison process parameters to be fitted to data: relatedness parameters u and v and diffusion process boundary parameters z and a. These, plus an encoding and response stage parameter T_{ER}, summarize performance in item recognition experiments.

Variable criteria. For any particular item recognition task, there are three variable criteria to be set by the subject: the zero point on the relatedness dimension (corresponding to β in signal detection theory) and the two boundary positions in the diffusion process. Note that an item with relatedness zero, that is, at the relatedness criterion, will have an average rate of accumulation of evidence that leads to no systematic drift toward either the match or nonmatch boundaries. Changes in boundary positions are the basis of speed–accuracy trade-off and are discussed in a later section.

Reaction time distributions. One feature that distinguishes the retrieval theory from most other models in this area of research is that reaction time distribution shapes are predicted and fitted. Figure 4 illustrates how the normal distribution of relatedness with mean u maps into a skewed reaction time distribution. Also shown in Figure 4 is an illustration of the way in which the distribution changes as u changes: When u decreases, the fastest responses slow a little, and there is a large elongation of the tail of the distribution. Thus, a prediction of the model is that if relatedness decreases, the mean and mode of the distribution will diverge. In general, the mathematical model will be manipulated to yield expressions for reaction time distributions and accuracy for both hits and correct rejections that can then be fitted to data.

Further aspects of the comparison process are discussed later in relation to speed–accuracy trade-off and in relation to other models of the comparison process.

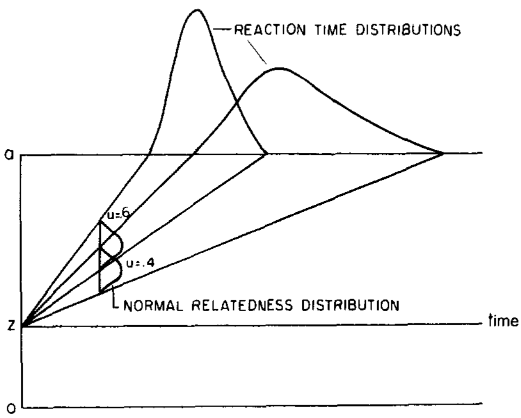

Figure 4. A geometrical illustration of the mapping from a normal relatedness distribution to a skewed reaction time distribution (with variance in drift $s^2 = 0$). (Note that as relatedness decreases, the distribution tail skews out. a represents the distance between the bottom and top boundaries of the diffusion process, z represents the distance between the bottom boundary and the starting point, and u represents the mean of the normal relatedness distribution.)

Decision Process

The decision process can be viewed as a combination stage in which the products of the many comparisons are combined to produce either a "yes" or "no" response. The decision process is self-terminating on a match, so that if enough evidence is accumulated in just one comparison process, a "yes" response is produced. In contrast, the decision process is exhaustive on nonmatch comparisons, so that a "no" response is produced only when all comparison processes terminate in nonmatches.

It has already been noted that the finishing time distribution for a single comparison is similar to observed reaction time distributions. It turns out that the distribution of finishing times of the maximum of a number of diffusion processes (each with a skewed finishing time distribution) is again shaped very much like observed reaction time distributions. Thus, the mathematical model produces acceptable shapes for both positive and negative response time distributions.

In most item recognition studies, "no" responses are almost as fast as "yes" responses, and sometimes, this finding has provided problems for certain models (see Collins & Quillian, 1970; Smith, Shoben, & Rips, 1974). This is not a problem for the retrieval theory, although it may be hard to see how "no" responses, which are based on exhaustive processing of the search set, can be fast enough. In fits of the mathematical model, it is found that the separation of the starting point and

Table 1
Accuracy and Latency as a Function of the Number of Extra Fast Finishing Processes for Correct Rejections, with $v = -.4$, $a = .12$, $z = .03$, *and Search Set Size 16*

No. extra processes *m* (evoked set = *m* + 16)	No. standard deviations *r* (of the *m* process distribution) below the nontarget distribution	Proportion correct	Mean reaction time (in msec)
0	0	.839	358
100	2η	.838	363
500	3η	.838	362
5,000	4η	.838	358
25,000	5η	.838	363
100,000	5η	.835	358

nonmatch boundary (z) is smaller than the separation of the starting point and match boundary ($a - z$). Therefore, it is even possible to have "no" responses much faster than "yes" responses (as found by Rips, Shoben, & Smith, 1973).

Resonance Metaphor

It was argued earlier that the search set could be made very large, but that reaction time and error rates would still be determined only by the slower comparison processes. Suppose that in a typical fit to data from the study–test paradigm, *m* extra processes are added to the 16 probe–memory-search-set comparison processes. Table 1 shows accuracy and reaction time as a function of *m* and *r*, where *r* is the number of standard deviations the extra process distribution is placed below the nontarget item distribution. Note that the number of extra processes is adjusted, so that the change in latency and accuracy is small for the particular value of *r* used; thus, more processes than *m* would produce significant changes in accuracy and latency. The results in Table 1 show that many probe–item comparisons with small relatedness may enter the decision process without affecting accuracy or latency. These results suggest that the resonance metaphor has the right sort of properties to be useful in theorizing about item recognition processes.

Speed–Accuracy Trade-off

In many models of item recognition, accuracy and latency are not explicitly related, even though it is well known that a subject can sacrifice accuracy to produce a faster response (e.g., Pachella, 1974). There are several methods for studying speed–accuracy trade-off, for example, Wickelgren (1977) lists six methods. His six methods fall into three basic classes: (a) methods that induce the subject, for example, by use of instructions or payoffs, to adjust the amount of information required for a response; (b) methods that force the subject, for example, by use of deadlines, response signals, or time bands, to respond within some time limit; and (c) methods that partition reaction times into groups (e.g., those between 420 msec and 440 msec, those between 440 msec and 460 msec, and so on) and compare accuracy values within each partition.

The method of partitioning reaction times has been discussed by Pachella (1974) and Wickelgren (1977). Pachella claims that the speed–accuracy function obtained from this procedure is probably independent of the speed–accuracy functions obtained from the other two methods. In fact, the speed–accuracy functions obtained from partitioning simply tell us about the latency density functions for error and correct responses and, in terms of the retrieval theory, are largely unrelated to the other speed–accuracy measures and so are not considered further.

In the retrieval theory, the relation between speed and accuracy is central, and the diffusion comparison process (plus the decision process) may be viewed as a transformation from a relatedness discriminability scale to observed speed and accuracy measures. The first speed–accuracy method, in which the amount of information required for a response is manipulated, and the second method, in which the time of response is manipulated, can be viewed as complementary in terms of the theory. When instructions or payoffs are manipulated, the process will be referred to as *information controlled*; when response signals or time windows are used, the process will be referred to as *time controlled.*

Information-controlled processing. To

demonstrate the way in which speed–accuracy trade-off is modeled when instructions or payoffs are manipulated, let us consider two probe–memory-set item comparisons: one match and one nonmatch. The assumption is that payoffs or instructions induce the subject

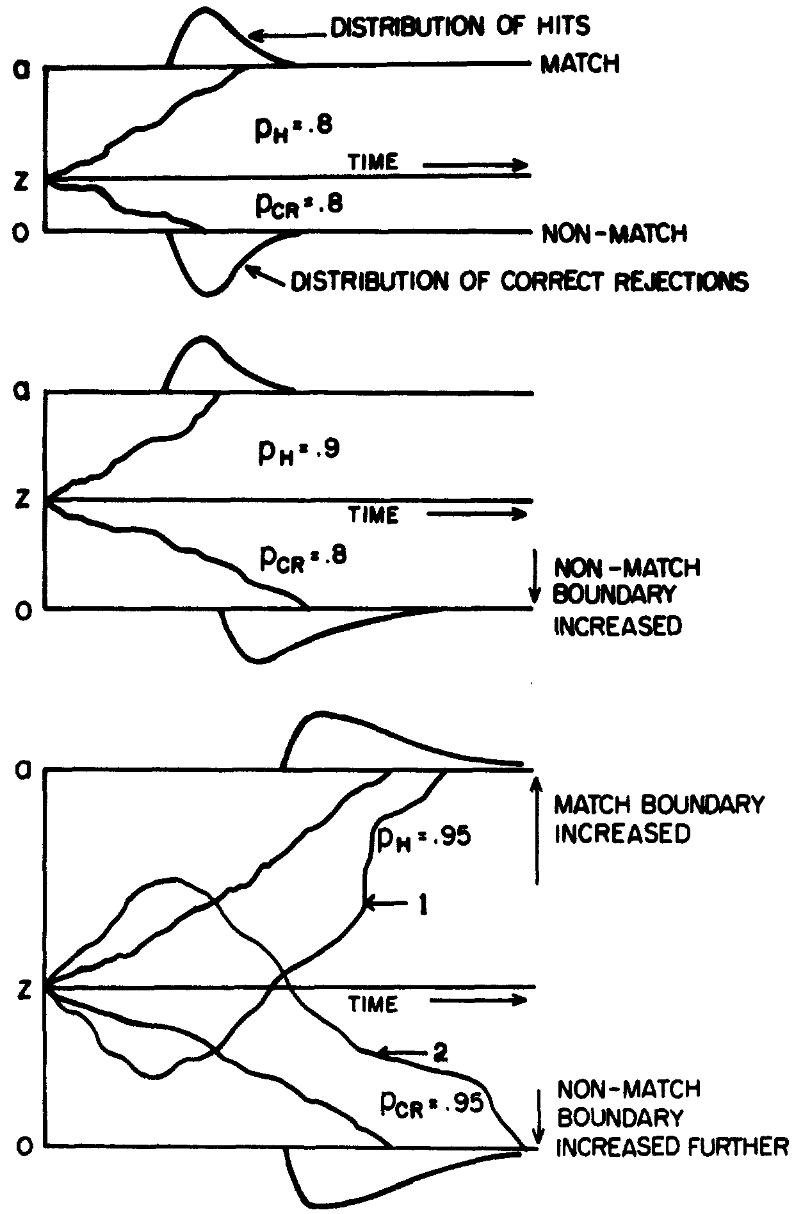

Figure 5. An example of the change in comparison time distributions and correct response rates (p_H is the probability of a hit and p_{CR} is the probability of a correct rejection) for changes in boundary positions for one match process and one nonmatch process, with all other parameters constant. (The numerical values and positions and shapes of the latency distributions illustrate the qualitative relations between the three speed–accuracy conditions but are not quantitatively exact. [Note that in the middle panel, there will be a slight decrease in p_{CR}.] Processes 1 and 2 illustrate comparisons that would have terminated at the wrong boundary if the boundaries had been close as in the top panel. a is the distance between the two boundaries in the random walk comparison process, and z is the bottom boundary to starting point distance.)

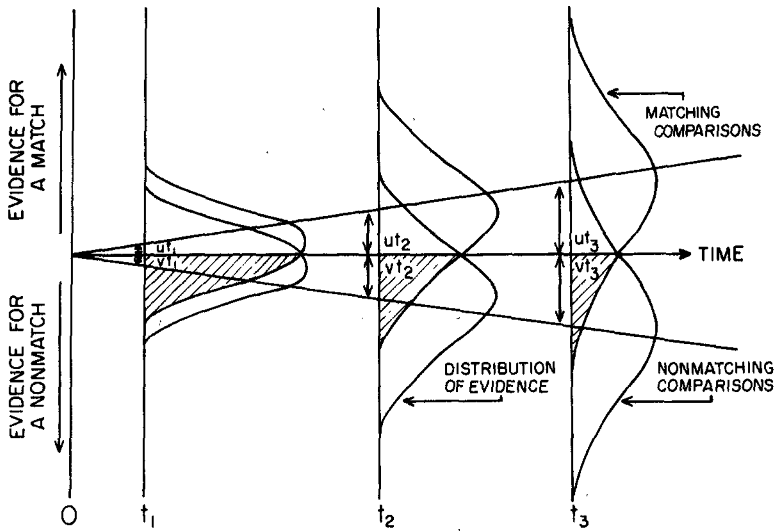

Figure 6. Spread of evidence as a function of time in the unrestricted diffusion process for one matching and one nonmatching process. (At time t_1, there is a large amount of overlap; at t_3, the overlap has reached asymptote, with asymptotic $d' = (u - v)/\eta$ [see middle panel of Figure 3]. u is the mean of the match relatedness distribution, and v is the mean of the nonmatch relatedness distribution.)

to adjust the positions of the match and nonmatch boundaries and so to adjust the amount of information required for a decision. For example, Figure 5 illustrates the effect of increasing the position of the nonmatch boundary (middle panel) and increasing the position of both boundaries (bottom panel). When the position of the nonmatch boundary is increased, "no" responses are made slower and the number of misses reduced. When both boundaries are moved apart, latency and accuracy of both hits and correct rejections increase.

As the boundaries move apart, the reaction time distributions both shift (the fastest responses or leading edge of the distribution become slower) and spread (the mode and mean of the distribution diverge), as shown in Figure 5. The bottom panel of Figure 5 illustrates the time course of evidence accumulation for four comparison processes. The two processes marked "1" and "2" are examples of processes that would terminate incorrectly with narrow boundaries but that terminate correctly with the relatively wide boundaries shown in the bottom panel. It is the correct termination of processes such as Processes 1 and 2 that is responsible for the increase in accuracy as the diffusion process boundaries are moved apart.

Time-controlled processing. Time-controlled processing is modeled in a different way than information-controlled processing. Figure 6 shows the way evidence accumulates from two distributions of relatedness: (a) a probe–memory-set item match and (b) a nonmatch (see Figure 3, middle panel) as a function of time. The decision assumption is that if a particular comparison is stopped at time t_i and the amount of evidence is greater than the starting value, then the comparison is called a *match*, and if the amount of evidence is less than the starting value, the comparison is called a *nonmatch*. (For mathematical tractability later, it is assumed that the match and nonmatch boundaries are far apart and that the random walk diffusion process can be considered unrestricted). Figure 6 shows the way evidence has accumulated at times t_1, t_2, and t_3. At early times (t_1), there is a large amount of overlap, and d' is rather small. At later times (t_2 and t_3), target and nontarget distributions have diverged and are approaching an asymptotic form that corresponds to the relatedness distributions shown in the middle panel of Figure 3. Thus d' as a function of time approaches an asymptote. (Mathematical expressions for d' as a function of time and fits to data are shown later.)

Thus, we see that the random walk comparison process can account for results from both information- and time-controlled speed–accuracy experiments.

Model Freedom

In the description of the retrieval theory, I have stressed the flexible nature of the processing system, and this flexibility shows up as apparent model freedom. I now discuss the problem of model freedom and indicate which functional relations provide reasonably constrained tests of the mathematical model.

When applied to an experiment, the mathematical model has five variable parameters, namely, T_{ER}, z, a, v, and u. The encoding and response parameter T_{ER} is fixed, so that in order to evaluate the effect of independent variables on performance, T_{ER} can be considered constant. In any situation where experimental conditions change between trials and the subject can know which condition is being tested (e.g., set size in the Sternberg paradigm), it is possible that criteria (z, a, v, and u) will change across conditions. So, there are four parameters free to vary in fitting the data. There is a rather weak constraint on the value of these parameters in that if performance exhibits regular behavior, then changes in z, a, v, and u should be regular. More useful tests of the mathematical model can be found in experiments where, on any particular trial, it is not possible for the subject to know which condition is being tested prior to retrieval (e.g., serial position in the Sternberg paradigm). In such situations, criteria cannot be adjusted to systematically vary with the independent variable. Thus, the only parameter free to vary is relatedness u, and with only u varying, accuracy and latency must be simultaneously fitted.

One major aim in developing the mathematical model is to fit reaction time distributions. In fact, reaction time distributions are fitted without the need to add any further parameters. This account contrasts sharply with that given by the serial scanning model for the Sternberg paradigm. It is generally agreed that the exhaustive serial scanning model is simple and easily testable, with only three parameters, one slope parameter, and two intercept parameters. It is not generally realized, however, that to fit reaction time distributions (using additive-factors logic and the Pearson system of frequency curves), the serial scanning model suddenly requires nine parameters (Sternberg, Note 1). If accuracy performance were to be fitted also, then still more parameters would be required. Thus, the number of parameters needed for the retrieval theory to fit results from the Sternberg paradigm may turn out to be much the same as the number of parameters needed by the serial scanning model. This discussion shows that there are certain constraints that are rather hard to evaluate, for example, how many degrees of freedom are taken up in representing reaction time distributions.

It was argued earlier that many kinds of information enter the comparison process, including recency information. The decay of recency information (or strength) has been modeled several times, and various simple models with few parameters have been developed (Atkinson & Shiffrin, 1968; Murdock, 1974; Norman & Rumelhart, 1970; Wickelgren & Norman, 1966). The retrieval theory describes performance as a function of recency with a series of relatedness values u. Thus, for example, there will be a separate value of relatedness u at each serial position, that is, there will be as many free parameters as there are serial positions. If some assumptions about decay in relatedness were made, then this series of relatedness values could possibly be fitted by a simple few-parameter model (as in the models mentioned above), thus reducing the number of free parameters.

In conclusion, it seems that the problem of model freedom is more complex than it would appear at first sight. Simple few-parameter models require the addition of many more parameters to account for anything more than crude summary statistics, and complex models may be quite tightly constrained in nonobvious ways.

Mathematical Model

In this section, some of the principal parameters, functions, and expressions for the mathematical model are introduced. A much more complete discussion of the random walk and diffusion processes with derivation combination rules for exhaustive and self-terminating parallel processing are presented in the Appendix. Later in the present section, the method of fitting the theory to data is described, checks on the fitting program are described, and relations to other models are discussed.

Parameters of the Mathematical Model

The parameters used in the mathematical model arise from different theoretical sources and may be summarized as follows: (a) for the comparison (diffusion) process: lower boundary zero, starting point in the diffusion process z, and upper boundary a; drift in the diffusion process (equal to relatedness) and variance in the drift s^2; (b) for probe–item relatedness: a normal distribution $N(u, \eta)$ for matching comparisons and a normal distribution $N(v, \eta)$ for nonmatching comparisons; and (c) an encoding and response output parameter T_{ER}. Note that s^2 and η^2 are kept constant throughout the fits presented in this article, so that the only variable parameters are a, z, u, v, and T_{ER}. Thus, u and v alone represent input from memory into the decision system.

Diffusion Process

The diffusion process is used to represent the accumulation of evidence for a single probe–memory-set item comparison. As discussed earlier, one has to select aspects of the data (summary statistics) as points of contact between theory and data. I have chosen to use mean error rate and reaction time distribution statistics, and so it is necessary to derive expressions for error rate and finishing time density functions for the diffusion process. In the Appendix, expressions for error rate and finishing time density functions are presented for the discrete random walk. Taking the limit as the number of steps becomes large and the size of each step becomes small allows us to obtain expressions for error rate and finishing time density functions for the diffusion process (as shown in the Appendix).

Consider a single probe–memory-set item comparison with relatedness ξ; as noted earlier, drift in the diffusion process is set equal to relatedness. Let the nonmatch boundary be at zero, the starting point of the process be at z, and the match boundary be at a. Then, if s^2 is the variance in the drift in the diffusion process, the probability of a nonmatch is given by

$$\gamma_-(\xi) = \frac{e^{-(2\xi a/s^2)} - e^{-(2\xi z/s^2)}}{e^{-(2\xi a/s^2)} - 1}. \tag{1}$$

(Throughout this section the subscript minus will refer to a nonmatch and the subscript plus will refer to a match.) The finishing time density function for a nonmatch is given by

$$g_-(t, \xi) = \frac{\pi s^2}{a^2} e^{-(z\xi/s^2)} \sum_{k=1}^{\infty} k \sin\left(\frac{\pi z k}{a}\right) \times e^{-\frac{1}{2}(\xi^2/s^2+\pi^2k^2s^2/a^2)t}. \tag{2}$$

The equivalent expressions for a match [$\gamma_+(\xi)$ and $g_+(t, \xi)$] can be found by setting $\xi = -\xi$ and $z = a - z$ in Equations 1 and 2. Note that $g_-(t, \xi)$ is not a probability density function because some proportion of comparisons end up as matches. However, $g_-(t, \xi)/\gamma_-(\xi)$ is a probability density function.

Decision Process

The decision process is self-terminating for matches and exhaustive for nonmatches. Thus, in order for a "no" response to be made, all the comparison processes must finish, and the decision time is the maximum of the individual comparison process times. Let $G_{i-}(t)$ be the finishing time distribution function

$$G_-(t, \xi) = \int_0^t g_-(t', \xi)dt'$$

for a nonmatch of process i, and $G_{\max}(t)$ be the finishing time distribution function for the maximum of n nonmatch processes. Then,

$$G_{\max}(t, \xi) = \prod_{i=1}^{n} G_{i-}(t, \xi)$$

and also

$$\gamma_-(\xi) = \prod_{i=1}^{n} \gamma_{i-}(\xi), \tag{3}$$

where $\gamma_{i-}(\xi)$ is the probability of a nonmatch of the ith comparison, and $\gamma_-(\xi)$ is the probability of obtaining a nonmatch from the n comparisons. The expressions for the match decision (minimum of the comparison processes that terminate in a match) are more complicated and are shown in Equations A19 to A22 in the Appendix.

Equations 1, 2, 3 and A20 form the basis for computation of predicted reaction time distributions and error rates.

Probe–Memory-Set Item Relatedness

Because relatedness ξ has a normal distribution [$N(u, \eta)$ and $N(v, \eta)$ for matching and nonmatching comparisons, respectively], it is necessary to average $G(t, \xi)$ and $\gamma(\xi)$ over relatedness before obtaining the final theoretical expressions for error rates and reaction time distributions.

A complete description of the mathematical model, together with all the equations necessary to obtain reaction time distributions and error rates, is presented in the Appendix.

Fitting the Mathematical Model to Data

Despite the recent popularity of reaction time research, reaction time distributions have been largely ignored as a source of information. This is surprising in view of the fact that distributional information can be useful in evaluating models (Ratcliff & Murdock, 1976; Sternberg, Note 2). However, one problem in using reaction time distributions concerns the choice of statistics used to describe a distribution. The traditional method that uses moments (Sternberg, Note 1) turns out to be of little practical use because many thousands of observations per experimental condition are required before stable estimates can be obtained. Furthermore, higher moments (third and fourth) describe the behavior of the extreme tail of the distribution (Ratcliff, Note 3) and, therefore, are of little practical interest. Ratcliff (Note 3) has suggested that the parameters of the probability density function of the convolution of a normal distribution [$N(\mu, \sigma)$] and an exponential distribution [$f(t) = (1/\tau)e^{-t/\tau}$] provide a good summary of empirical reaction time distributions (see also Ratcliff & Murdock, 1976). I have decided to use the convolution distribution

$$c(t) = \frac{e^{-[(t-\mu)/\tau]+\sigma^2/2\tau^2}}{\tau\sqrt{2\pi}} \times \int_{-\infty}^{[(t-\mu)/\sigma-\sigma/\tau]} e^{-y^2/2}dy \quad (4)$$

as a meeting point of theory and data. The convolution model is fitted to data, and parameter values obtained serve to summarize that data. Theoretical reaction time distributions are computed, and the convolution model is fitted to the theoretical distribution giving theoretical values of the parameters. Figure 7 shows some fits of the convolution model to the theoretical distributions, and it can be seen that the curves are similar.

In order to fit the theory to data, the equations derived earlier were used to give values of correct response rate p_t and reaction time distribution parameters μ_t and τ_t. (The parameter σ of the convolution model can be left out because generally σ does not theoretically and empirically change appreciably with experimental condition, and the values for both theoretical and empirical estimates are similar.) The corresponding three empirical parameters p_e, μ_e, and τ_e are estimated from the data (averaging across subjects), and the function

$$\sum_{\mathrm{H,CR}} \left[\frac{(p_e - p_t)^2}{\sigma_p^2} + \frac{(\mu_e - \mu_t)^2}{\sigma_\mu^2} + \frac{(\tau_e - \tau_t)^2}{\sigma_\tau^2} \right] \quad (5)$$

(summed over both hits [H] and correct rejections [CR]) is minimized as a function of the theoretical parameters. The variances σ_μ^2 and σ_τ^2 are asymptotic variance estimates for the convolution model (see Table 2 of Ratcliff & Murdock, 1976), and $\sigma_p^2 = p_e(1 - p_e)/N$, where N is the number of observations on which p_e is based.

Because the theoretical reaction time distributions are best fits to the average reaction time distributions across subjects (μ_e and τ_e are averages over subjects), it is difficult to get any direct idea of how well the theory accounts for the data. Ratcliff (Note 3) has developed a method of producing group reaction time distributions. Essentially, reaction

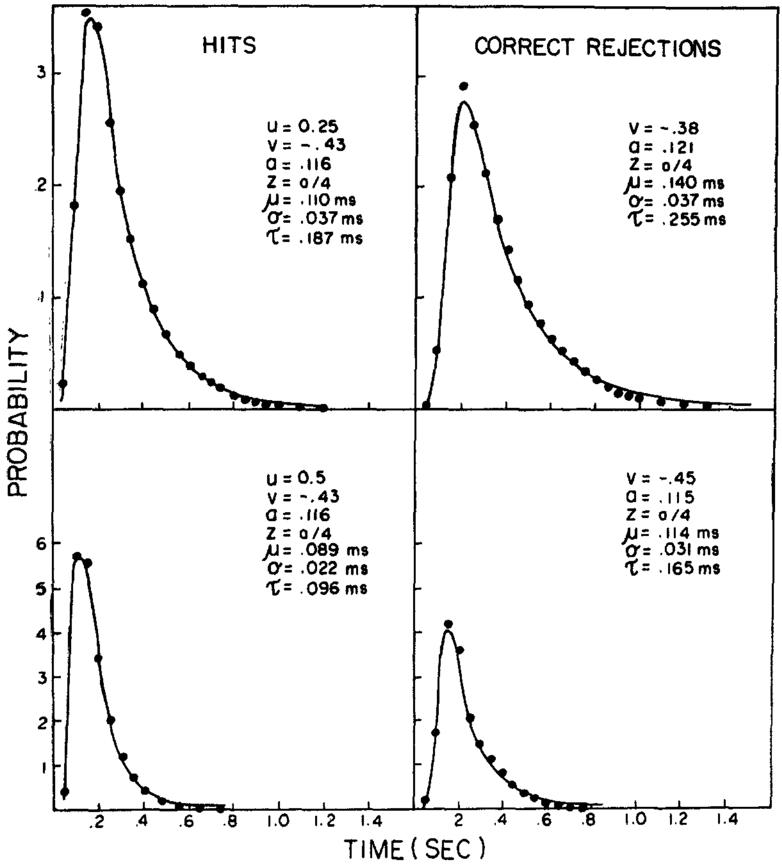

Figure 7. Some sample theoretical reaction time distributions (solid lines) for theoretical parameters u, a, z, and v and fits of the convolution model parameters μ, σ, and τ to theoretical distributions (dots). (These distributions illustrate the adequacy of the convolution approximation to the theoretical distributions. ms = msec.)

time quantiles are derived for each subject cell, and these quantiles are averaged over subjects to give group quantiles. The group quantiles are then plotted on the horizontal axis of a graph. If, for example, 5% quantiles were used, then 5% of the probability density would lie between adjacent quantiles. Therefore, a probability density function can be constructed by drawing equal area rectangles between adjacent quantiles. The theoretical probability density function can be compared with the empirical function by plotting the two functions on the same graph. This method is used to show how well the theoretical reaction time distributions fit the data. It should also be noted that the convolution parameters derived by fitting the convolution model to the reaction time quantiles are almost identical to the convolution parameters that are the average, across subjects, of convolution fits to individual subject cells (Ratcliff, Note 3).

Checks on the Estimation Method

In fitting moderately complex models, such as this diffusion process model, there are many places where errors might occur. Therefore, two kinds of checks should be carried out: internal and external checks. For example, an internal check that should be carried out when numerically integrating to infinity involves printing out the integrand over the range of integration and checking the values by hand calculation. External checks can be somewhat more involved. The idea here is to compare values obtained from the program with some known or established values. A program to compute accuracy and reaction time distributions was implemented on an IBM 370-165 computer in WATFIV FORTRAN and, without the integration over relatedness (Equations A24 and A25), was implemented on a PDP-12A computer in FOCAL language. Letting $\eta \rightarrow 0$ in the FORTRAN program, the two programs were found to give identical results. The FOCAL program was checked by calculating values of first passage time distribution for $a \gg z$. These results were then compared with values obtained from the one barrier diffusion process, Equation 6 (see Cox & Miller, 1965, p. 221; Feller, 1968, p. 368), with probability density function

$$g_{-}(t) = \frac{z}{\sqrt{2\pi s^2 t^3}} e^{-(z-ut)^2/(2s^2t)}. \quad (6)$$

These values were identical, and so the evaluation of the model can be considered satisfactory.

Relations to Other Models

In this section, I compare and contrast the present theory with three classes of models and theories. The first class of models assumes exponentially distributed processing stages; some of these models attempt to deal with the same experimental domain as the retrieval theory. The second and third classes are random walk and counter models; these have been developed to account for both accuracy and latency in choice reaction time and research in perception.

Models with exponential processing stages. In several models of memory search and memory retrieval, exponential distributions have been used to represent the finishing time distributions of processing stages (Anderson & Bower, 1973; Atkinson, Holmgren, & Juola, 1969; Townsend, 1971, 1972). The exponential distribution is useful because it is a one-parameter distribution and has a skewed positive tail just as observed reaction time distributions. Furthermore, the exponential is easy to work with and quite often allows exact solutions to be obtained for complicated expressions (e.g., Anderson, 1974). However, there are several problems in using the exponential distribution. First, Sternberg (1975) argues that the Markov (no memory) property of exponential distributions is inconsistent with what is meant by processing over time. For an exponential process, the expected time to termination is independent of the time already elapsed; whereas a more reasonable view of cognitive processing would suppose that the longer a process has been running, the greater the probability of its termination. Second, the use of exponentials to represent stage distributions may produce good fits to mean reaction time but may not lead to theoretical distributions that match the shape of observed reaction time distributions. Furthermore, when the additive-factors method has been applied and component stage distributions extracted using Pearson's method (Sternberg, Note 1), the underlying distributions are far from exponential. Thus, a good strategy may be to develop a model using exponentials and then test whether in fact the stage distributions are exponential. Third, simply assuming that processing stages have certain finishing time distributions avoids the crucial fact that both latency distributions and errors arise from common stages or mechanisms, and that perhaps the major theoretical effort should focus on relating accuracy and latency.

Exponential distributions are an extremely useful tool for beginning an investigation but can be somewhat misleading theoretically if they represent processing stage distributions in a mature model.

Random walk models. Several models of choice reaction time based on the random walk comparison process have been developed recently (Laming, 1968; Link, 1975; Link & Heath, 1975; Stone, 1960). Here, I consider a recent formulation termed ***relative judgment theory*** (Link, 1975; Link & Heath, 1975) and note some of the similarities and differences to the retrieval theory.

In relative judgment theory, a physiologically transduced value of the stimulus is compared to an internal referent or standard. During a unit of time, the difference obtained is added to an accumulator of differences. When the accumulated difference exceeds one of two subject-controlled thresholds, a response is made.

Relative judgment theory has several important similarities to the theory developed in this article. First and foremost, the decision process is modeled by a random walk, and information discriminating the two stimuli is mapped onto a single dimension. Second, there are three criteria to be set by the subject: the value of the internal referent, corresponding to the zero point of relatedness in the retrieval theory, and the two boundary positions. Thus, there are the same number of variable criteria in the two mathematical models. Third, the two theories are both concerned with the joint behavior of accuracy and latency as a function of independent variables.

It is the differences between the theories that provide most insight into their relation. First, the random walk in relative judgment theory has discrete steps, and the difference accumulated in a particular interval has some distribution; whereas the retrieval theory assumes continuous processing. Mathematical expressions for the first passage time (latency) distributions have not yet been obtained for relative judgment theory, except in the case of binomial (or trinomial) and normal distributions of differences, as they have for the continuous random walk. For choice reaction time, error responses are faster than correct responses when accuracy is high, and relative judgment theory accounts for this by assuming that the moment-generating function for differences is nonsymmetric. This assumption is equivalent to supposing that the individual step distribution is positively skewed in the direction of an incorrect response. In contrast, in the retrieval theory, the continuous random walk has the normal distribution underlying performance [essentially drift is distributed $N(u, s)$]. Second, the internal referent in relative judgment theory is assumed to have no systematic drift over the course of testing. Thus, there is only one source of noise or variance and that is in the comparison process. In the retrieval theory, there is a further source of variance and that is in memory-set item–probe relatedness. Therefore, relative judgment theory predicts that if the response boundaries are set at infinity, discriminability d' increases as a function of the square root of the number of steps. This prediction is quite reasonable if, for example, the choice is between two well-discriminated tones. In contrast, variance in relatedness in the retrieval theory ensures that d' asymptotes as a function of retrieval time. Third, relative judgment theory has been developed to deal with the simplest case, namely, a two-stimulus, two-response discrimination procedure. On the other hand, the retrieval theory has been developed to deal with the problem of matching one item against many items in memory and discriminating one match against many nonmatches. Fourth, relative judgment theory has provided somewhat cleaner experimental tests than will be shown in the next section. However, it is somewhat difficult to compare the power of the two sets of tests of the two theories.

From this discussion, it can be seen that there is potential for making explicit a continuity between memory retrieval and perceptual discrimination.

Counter models. The class of counter models provides a serious competitor to random walk models. Perhaps the major difference between counter and random walk models is that random walk models assume that positive and negative differences cancel, whereas counter models assume separate counters for positive and negative evidence. Thus, for counter models, termination occurs when one of the two counters exceeds a criterial count.

Counter models have been examined in some detail by Pike (1973) with specific application to signal detection. In Pike's article, many of the main properties of counter models are listed and discussed. Anderson (1973) combines a counter model with a neural network model to account for results from choice reaction time studies and the Sternberg paradigm. Huesmann and Woocher (1976) have applied a counter model, together with parallel processing assumptions, to the Sternberg paradigm.

One major problem with counter models concerns the behavior of latency distributions as a function of count criteria. When count criteria are low (e.g., 1 or 2), latency distributions are reasonably skewed. However, as the count criteria increase (e.g., 4 to 10), the distributions become more normal (Pike, 1973, p. 61). This contrasts with the prediction of random walk models that the mode and mean of the latency distribution should diverge as the random walk boundaries (criteria) are moved apart.

At the moment, counter models as well as random walk models have not been investigated in enough detail, and further research may provide useful competitive theories for both memory retrieval and perceptual discrimination.

Experimental Paradigms

Up to this point, I have set up the basic framework of the theory and derived the necessary mathematical expressions. In the following sections, the theory is applied to five experimental paradigms.

Study–Test Paradigm

In a typical study–test experiment, 16 stimulus words are presented for study, and then 32 words (16 old and 16 new in random order) are tested for recognition, one at a time. The study phase is paced at about 1 sec per item, and the test phase is self-paced, with each test item staying in view until a response is made. Responses are made on a 6-point confidence scale, and accuracy and latency are recorded.

The study–test paradigm has been studied recently in some detail (Murdock, 1974; Murdock & Anderson, 1975; Ratcliff & Murdock, 1976), primarily as the empirical basis for the conveyor belt model (Murdock, 1974) for item recognition. The conveyor belt model assumes that a record of the temporal order of both study and test items is stored in memory. In order to decide whether a test item was in the study list or not, a high speed self-terminating backward serial scan is carried out over the temporal record of both test and study items. The conveyor belt model deals with retrieval from supraspan lists only and brings coherence to a number of empirical results. However, the model has some problems and limitations, and these are discussed in Ratcliff and Murdock (1976). Because of the large amount of data collected in Murdock's experiments and because of the stability of the empirical effects, the study–test paradigm provides a very strong empirical foundation for attempts to model recognition memory processes. Ratcliff and Murdock (1976) provided a summary of empirical effects, and this summary (see Figure 8) is the basis for much of the following discussion. It should be noted that the data actually come from a confidence judgment procedure but are modeled as though they came from a yes–no procedure. It turns out that results from the two procedures are similar (Murdock, Hockley, & Muter, 1977), so this approximation seems reasonable.

To model the study–test paradigm, I will assume that a test (or probe) item is encoded, and parallel comparisons are made between the encoded test item and members of the search set. The search set is approximated by the 16-item study set, with the same normal probability density function $N(v, \eta)$ for each nonmatching comparison (i.e., it is assumed that only study items are accessed by the test item and all nonmatching comparisons have equal relatedness). The probability density function of relatedness for a matching comparison is $N(u, \eta)$.

Serial and Test Position Functions

The main empirical effects to be modeled are the behavior of latency and accuracy as a function of study (or input) and test (or output) position for old items and of test position for new items (see Figures 8b and 8c). Data are taken from Ratcliff and Murdock's (1976) Experiment 1.

Parameter variation. As noted earlier, when variation in some experimental variable can be known to the subject prior to retrieval, there is the possibility of changes in criteria as a function of that experimental variable. Test position is one such variable. As test position increases, forgetting of study items increases, and the subject may set more lenient criteria (increases a and z) in an effort to improve accuracy. Thus, relatedness of probe–nontarget item comparisons (v) and diffusion process boundary parameters (z and a) can change as a function of test position.

Relatedness of probe–target item comparisons (u) varies as a function of test position, as do v, a, and z. However, unlike v, a, and z, u varies as a function of study (or serial) position. The subject cannot alter criteria as a function of serial position because knowledge of the

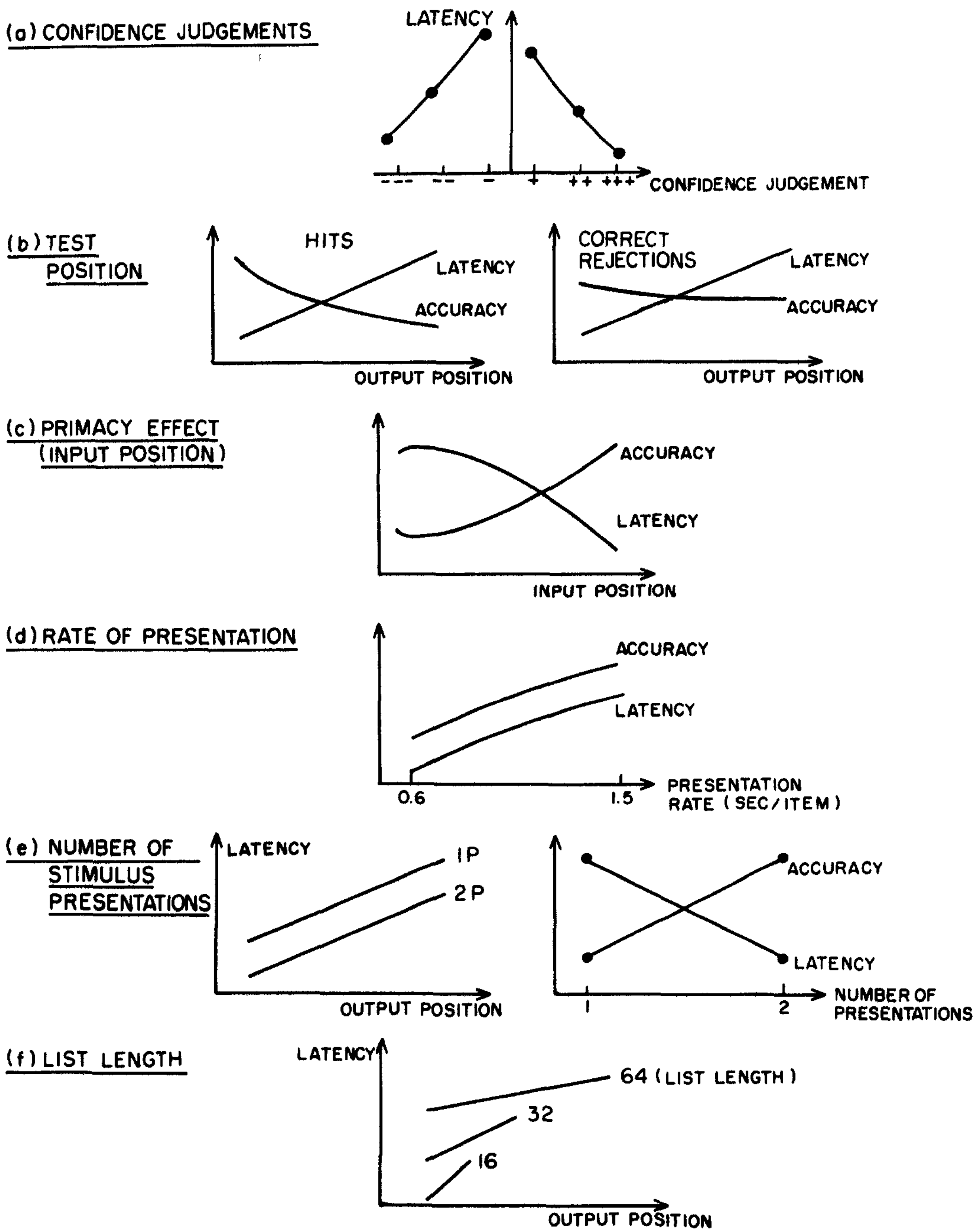

Figure 8. Schematic representation of functional relations obtained in the study–test paradigm (from Ratcliff & Murdock, 1976). (1P and 2P represent once-presented and twice-presented stimuli, respectively.)

study position of a probe item can only result from the retrieval of that item. Thus, changes in accuracy and latency as a function of serial position must both be predicted by changes in the single parameter u.

The variance parameters s^2 and η^2 are kept constant, as mentioned earlier, at $(.08)^2$ and $(.18)^2$, respectively. A single constant time T_{ER} is assumed for probe encoding, preparation for comparison and decision processes, execution of the decision process, and response output processes. These processes could be assumed to have some distribution of processing times with small variance, without altering the fits to distributions shown later.

Correct rejections. Figure 9 shows accuracy, mean reaction time, and reaction time distribution parameters μ and τ for correct rejections as a function of test position, together with theoretical fits and theoretical parameters v, a, and z. The fits are quite acceptable, and it can be seen that, indeed, criteria change as a function of test position to compensate for the overall drop in discriminability (forgetting).

Hits. Figure 10 shows accuracy and mean reaction time as a function of study (or serial) position, together with theoretical fits and values of relatedness u. From these fits, it can be seen that the one-parameter u can fit both accuracy and latency and that changes in accuracy and latency can be modeled by changes in the single parameter u. Inspection of Figure 10 shows that u behaves in a very regular manner as a function of study and test

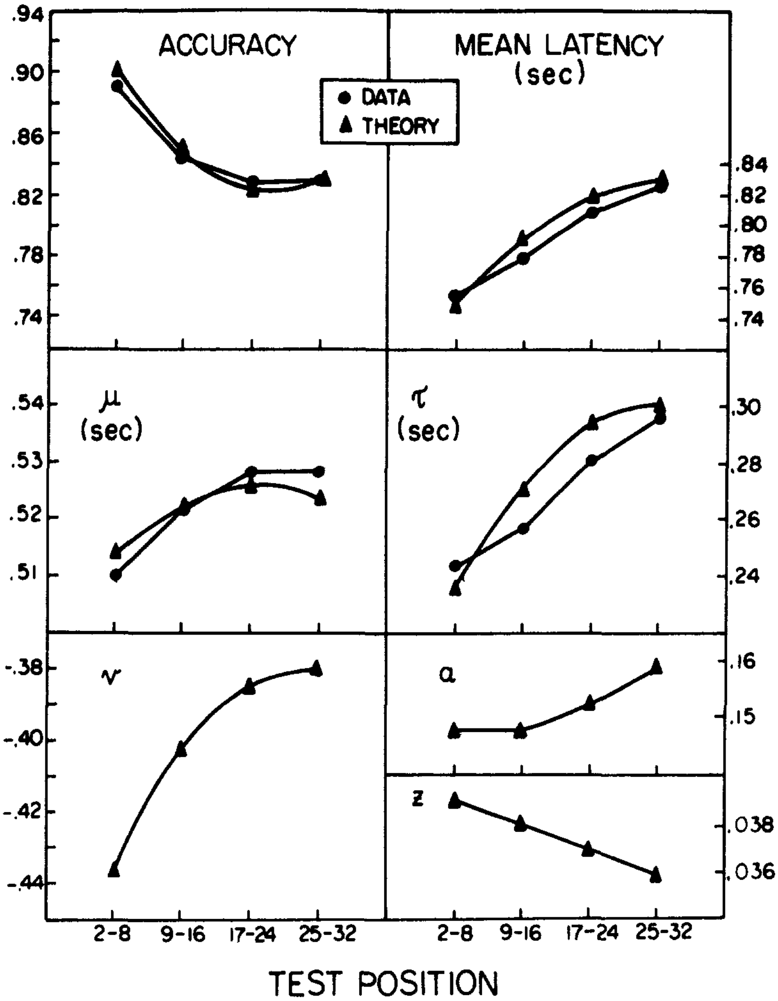

Figure 9. Accuracy, mean reaction time, and reaction time distribution parameters (μ and τ) as a function of test position for the study–test paradigm, together with the theoretical parameters v, a, and z, which are used to generate the theoretical fits. (Data are from Ratcliff and Murdock's, 1976, Experiment 1.)

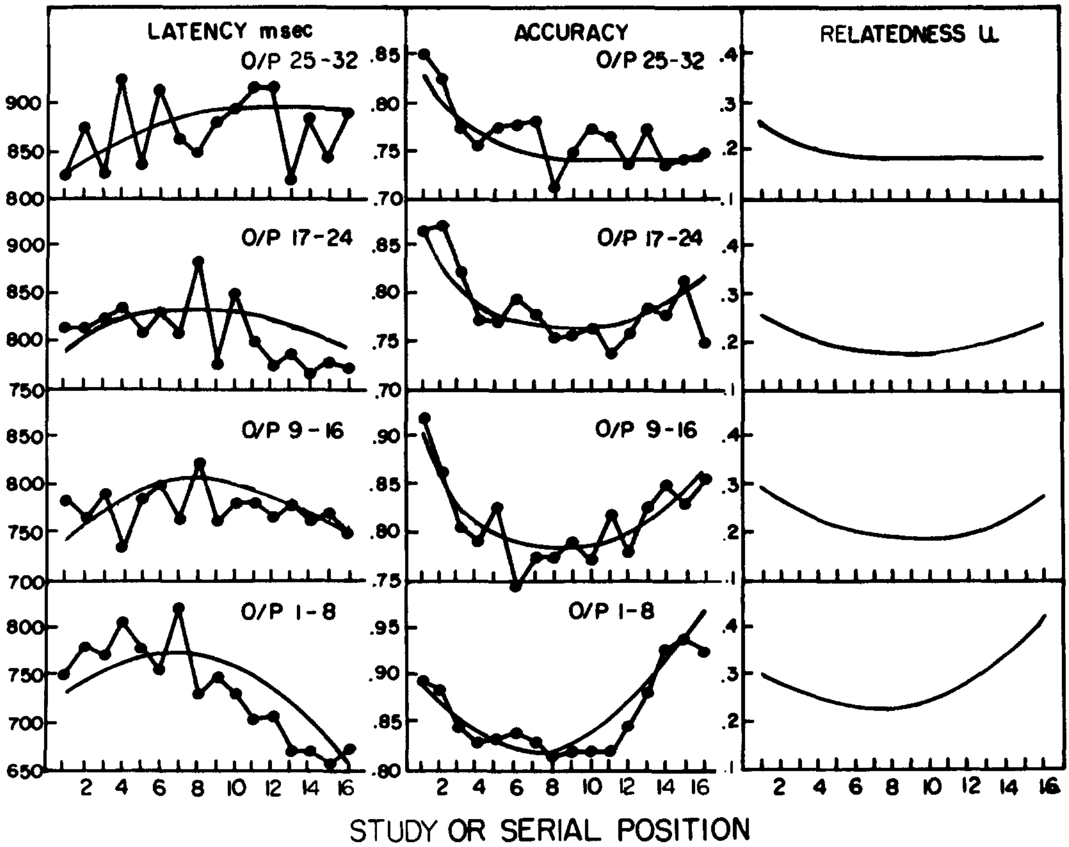

Figure 10. Accuracy and mean latency as a function of study position blocked by test, or output, position (O/P), which is used to produce the fits. (Other theoretical parameters used in the derivations are shown in Figure 9.)

position. There is a consistent primacy effect across test (or output) position, and the strong recency seen at early test positions has dissipated by later test positions. Therefore, it seems likely that one could reduce the number of degrees of freedom by modeling the pattern of behavior of u just as the pattern of accuracy performance in free recall has been modeled as a function of serial position and delay (i.e., test position; Murdock, 1974, chap. 8). I have not attempted to fit such a functional relation because it adds little theoretically. Changes in reaction time distributions as a function of serial position are mainly changes in spread of the distribution or the τ parameter in the convolution model (Ratcliff & Murdock, 1976), and this pattern of results is predicted by the theory.

Reaction time distributions. In the section on fitting the mathematical model to data, I argued that the most satisfactory way to fit the theoretical reaction time distributions to data (in the case of this theoretical model) is through the use of the convolution model as an empirical summary of shape. Further, it was also argued that the best way to display fits of the theory to data averaged over subjects is by the use of group reaction time distributions (see also Ratcliff, Note 3). Figure 9 shows both the theoretical and empirical values of μ (leading edge) and τ (tail), parameters of the convolution model, as a function of test position for correct rejections. It is these data that posed problems for the backward serial scanning (conveyor belt) model (Ratcliff & Murdock, 1976). The conveyor belt model predicts that most of the increase in mean latency will be accounted for by an increase in latency of the whole distribution (μ parameter) rather than the tail (τ parameter). Although the theoretical distributions are fitted to empirical data by using the μ and τ parameters (i.e., minimizing a function of μ and τ) of the convolution model, the retrieval theory makes the strong prediction that changes in mean latency will be manifest mainly as changes in τ (i.e., slower responses getting slower; see Figure 4). In fact, patterns of results with changes in μ greater than changes in τ as a function of test position would be impossible to accommodate without large changes in random walk boundary positions a and z. Large changes in a and z would change error rates and hit latencies in such a way that the set of data could not be fitted simultaneously.

Figure 11 shows fits of the theoretical distributions to the group reaction time distributions for both hits and correct rejections as a function of test position. From these graphs, it can be seen that the theory does a good job of fitting the reaction time distributions.

Error reaction times. It was noted earlier that the data used in fitting the theory came from a confidence judgment procedure rather than a yes–no procedure. For correct reaction times, it was reasonable to use the data from high-confidence correct responses because results were not changed much by addition of lower confidence reaction times. Another reason to use only the high-confidence responses is that many lower confidence responses may be contaminated by other processes such as more than one comparison or a slow guess. In contrast, there are relatively few error responses, and high-confidence responses are typically 150 msec to 600 msec faster than low-confidence responses. The theory makes predictions (using parameter values obtained earlier in this section for Ratcliff & Murdock's, 1976, Experiment 1) about error reaction times that are not too unreasonable. For example, the theory predicts that miss reaction times increase with output position at about 5 msec per item (by linear regression), with an intercept of about 860 msec. The data for high-confidence responses show a slope of about 11 msec per item, with an intercept of 820 msec. All miss responses combined (high, medium, and low confidence) give reaction times about 200 msec longer. For false alarms, predicted slope is about 5 msec per item, and the intercept is 1,200 msec. High-confidence

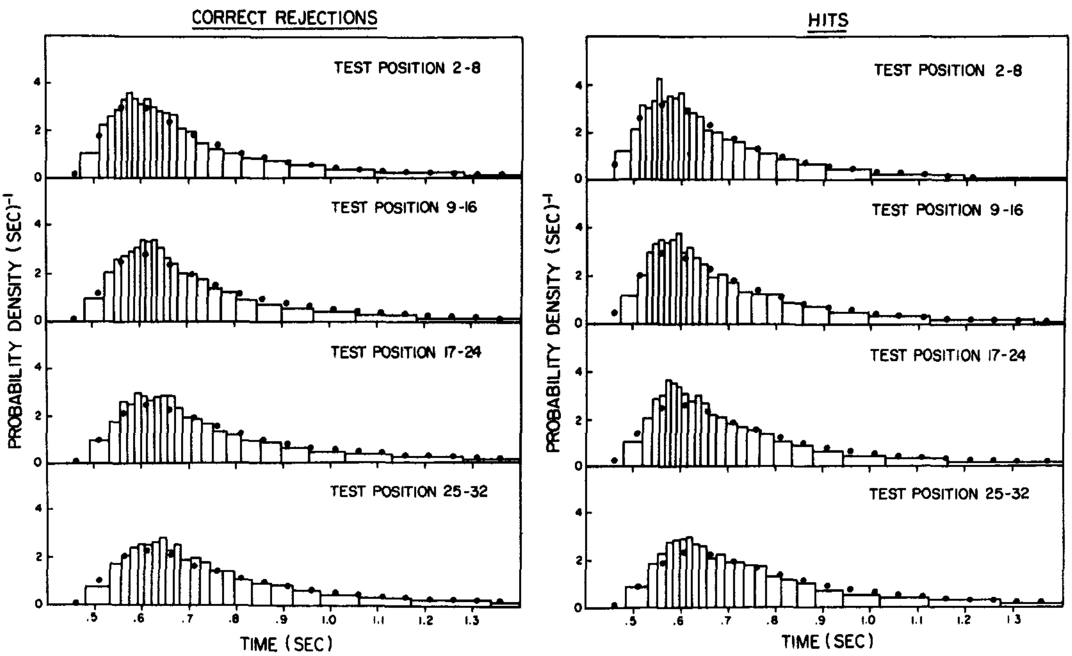

Figure 11. Group reaction time distributions and theoretical fits to those distributions from Ratcliff and Murdock's (1976) Experiment 1.

Table 2
Fits to Accuracy and Latency for Repeated and Nonrepeated Study Items in the Study–Test Paradigm (Data are from Ratcliff & Murdock's, 1976, Experiment 4)

	Theoretical							Experimental			
Response	u	v	a	Accuracy	μ	σ	τ	Accuracy	μ	σ	τ
Correct rejection	—	−.38	.1	.74	472	37	216	.74	470	40	220
Once-presented hit	.18	−.38	.1	.69	455	34	190	.70	450	40	190
Twice-presented hit	.18	−.38	.1	.87	445	26	170	.84	450	40	160

Note. μ, σ, and τ are measured in msec; T_{ER} = 350 msec and z = .02.

false alarms have slopes of 8 msec per item and intercepts of 830 msec, so the theory overpredicts high-confidence false alarms. However, adding lower confidence false alarms increases reaction time by about 600 msec. Thus, the theoretical error reaction times come near to the data in some cases but are not so near in other cases.

Number of Stimulus Presentations

When items are repeated in the study list, I assume that separate representations are established. The assumption of separate representations is made because of the problems that a single-representation strength theory appears to have (Anderson & Bower, 1972; Wells, 1974). With two representations, there are two comparison processes with high relatedness values, and these two processes race to match. Thus, reaction time will be faster and accuracy higher than for singly presented items, as is seen in the data (see Figure 8e). To model these data, representative values of accuracy and latency for once-presented items are chosen, model parameters for those values are deduced, and then accuracy and latency for the two racing processes are computed. To compute fits for twice-presented items, it was assumed that the two racing processes had equal u values, and Equation A22 was modified appropriately. This approximation is reasonable, taking into account the input or serial position functions shown in Figure 10, because in general, u values are not too different at different serial positions. Table 2 shows theoretical fits and experimental results for accuracy and latency distributions. The only discrepant result is the slight theoretical overestimation of accuracy of twice-presented items.

Rate of Presentation: Dangers Inherent in "Between" Designs in Reaction Time Experiments

A major problem for strength- or familiarity-based theories of retrieval is the result shown in Figure 8d. As rate of presentation of study items decreases, accuracy increases; but reaction time increases instead of decreasing, as a strength theory would predict. In the experiment from which Figure 8d was derived (Ratcliff & Murdock's, 1976, Experiment 2), rate of presentation was a between-sessions variable. The retrieval theory would predict the same result as a strength theory if all criteria were constant across conditions. However, there is a way out for the retrieval theory, and that is to suppose that criteria change with conditions.

As discussed earlier, one of the stronger tests of the retrieval theory is made when criteria cannot be adjusted to experimental conditions. Thus, a strong test of the theory can be made if presentation time can be made a within-trials variable. Then, only relatedness u can vary as a function of presentation time per item, and variation in both accuracy and latency must be predicted by variation in u alone. If the results shown in Figure 8d hold up when presentation time per item is made a within-trials variable, then the theory in its present form must be rejected. In Experiment 1, presentation time per item is made a within-trials variable.

Experiment 1

Method. A standard recognition memory procedure was employed. A list of 16 study items was presented, followed by a test list containing all the study items and an equal number of new items in random order. The lists were random samples from the University of Toronto word pool, a collection of 1,024 two-syllable common English words not more than eight letters long, with homophones, contractions, and proper nouns excluded. Each trial consisted of a random selection from the word pool. Each session had 32 lists, and there were no repetitions within a session. List generation, display, and response recording were controlled by a PDP-12A laboratory computer. In each study list, there were four different study times (.5, .8, 1.2, and 1.8 sec per item), and four study items (in random presentation order) were assigned to each of four study time conditions per list. The study list was terminated by an instruction asking the subject to press a response key to start the test phase. The test phase was self-paced, and items stayed in view until a response was made. A confidence judgment procedure was employed, and the subjects had to respond on a 6-point scale from − − − ("sure new") to +++ ("sure old") by pressing the appropriate response key. For each item tested, input and output position (output position only for new items), confidence judgment (i.e., key pressed), and latency (stimulus onset to response key depression) were recorded. A 5-msec time base was used. The four subjects were undergraduates in psychology at the University of Toronto and were paid $30 for the 12 sessions plus 1 practice session.

Results. Figure 12 shows the main result: When presentation time per item is a within-trials variable, then the longer an item is presented, the more accurate and faster is the response to that item. This result contrasts with the result found in Ratcliff and Murdock's (1976) Experiment 2 (shown in Figure 8d), where responses became more accurate but slower as duration of presentation increased. Also shown in Figure 12 are values of u as a function of presentation time per item and predicted values of accuracy and reaction time. Thus, the theory makes the correct prediction and is supported by the experimental results.

Implications for reaction time research. Besides providing a critical test of the retrieval theory, these results on presentation time per item have important consequences for the methodology of reaction time research. It has often been argued that subjects in reaction time experiments can adopt different speed–accuracy criteria under different experimental conditions, and subjects have been induced to do so in many experiments (e.g., Banks & Atkinson, 1974; Lively, 1972). However, in most reaction time experiments, there does not seem to be much evidence of criteria being

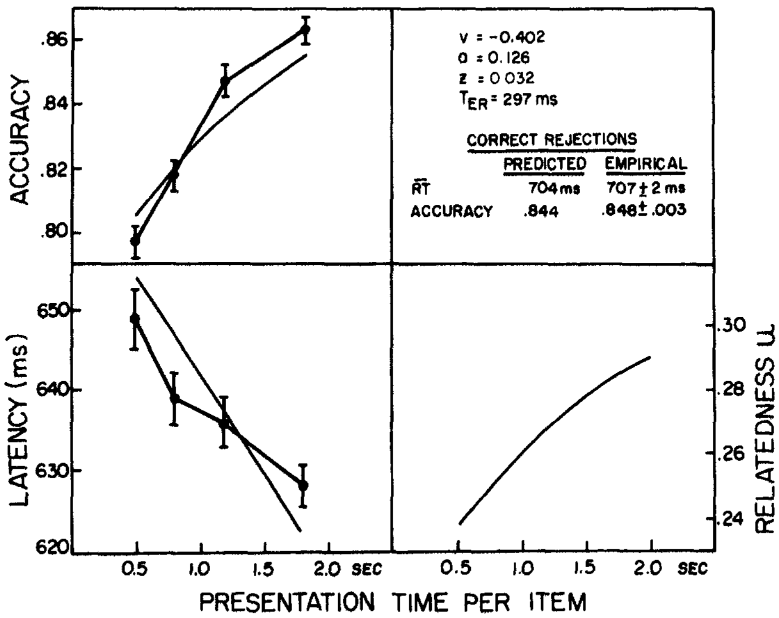

Figure 12. Accuracy, mean latency, and relatedness u as a function of presentation time per item for Experiment 1. (Theoretical parameters v, a, z, and T_{ER} are shown in the figure. $\overline{RT}$ represents mean latency; ms = msec.)

adjusted as a function of experimental condition. For example, Wickelgren (1977) states,

> To be completely fair, we do not know for certain that subjects' capability to achieve any degree of speed–accuracy tradeoff (from chance to asymptotic accuracy) under speed–accuracy tradeoff instructions implies that this capacity is used under reaction time instructions. Perhaps in reaction time experiments, all subjects use the same speed–accuracy criterion and use it under all conditions. (p. 81)

The above results on presentation time show that a theoretically important functional relation (latency as a function of presentation time per item) contradicts theories when studied in a between design yet supports the same theories when studied in a within design. There are two important consequences: (a) In any reaction time experiment where criteria can be changed as a function of experimental condition, any differences in reaction time as a function of condition may be interpretable only as a change in speed–accuracy criteria and nothing more. (b) Theories of reaction time should be flexible enough to deal with criteria changes such as those demonstrated above.

Study–Test Paradigm: Summary

The study–test paradigm provides a firm empirical basis for modeling item recognition processes. The retrieval theory models results by assuming that parallel comparisons between the probe (test item) and members of the study list are made. The decision process is self-terminating on matches and exhaustive on nonmatches. Test or output position functions are well fitted by assuming that v, relatedness of nontarget comparisons, decreases and diffusion process boundaries increase a little with increasing test position. Latency and accuracy as a function of study position show serial position functions with both primacy and recency. For fixed output position, only the relatedness parameter u can vary with serial position, and fits to both accuracy and latency with just this one parameter varying are good. Latency distributions and error latencies are also well described by theoretical predictions.

Ratcliff and Murdock (1976) presented an experiment in which number of stimulus presentations was varied. The theory was able to account for results from this experiment by assuming separate representations, so that there were two processes with high relatedness racing to the match boundary. Both latency and accuracy were well fitted by results derived from this assumption.

An investigation into rate-of-presentation effects exposed important methodological dangers in between designs involving the use of latency measurements. Ratcliff and Murdock (1976) showed that both accuracy and latency increased as rate of presentation was decreased in a between-sessions design. The latency effect is contrary to many theories, including the theory developed here, if it is assumed that speed–accuracy criteria are kept constant. The present Experiment 1 varied presentation time per item within trials. Accuracy increased as presentation time increased, as before, but reaction time decreased. Therefore, in between designs, it is possible for the subject to change speed–accuracy criteria and thus to make between latency comparisons subject to misinterpretation.

Sternberg Paradigm

The Sternberg paradigm is the most studied paradigm in memory research that employs latency measures. The method is simple: A small number of items (digits, words, pictures, and so on) within memory span is presented to the subject one at a time. A probe item is then presented, and the subject has to decide if that item was in the set of study items or not. The subject pushes one button for a "yes" response, another for a "no" response, and reaction time and accuracy are recorded. The result that has been of most interest theoretically has been the reaction-time–set-size function. Often, reaction time has been a linear function of set size, with equal slopes for "yes" and "no" responses. This result led Sternberg (1966) to develop a serial exhaustive scanning model. The model is easily testable, and several serious problems have been found; for example, the model is unable to deal with serial position effects (Corballis, 1967; Corballis, Kirby, & Miller, 1972), repetition effects (Baddeley & Ecob, 1973), stimulus and response probability effects (Theios et al., 1973), and nonlinear set-size effects (Briggs, 1974). There are alternative simple models that have been set up to account for subsets of the empirical findings, but it seems that each of these is inconsistent with some data (Sternberg, 1975). Thus, we find an area of research in which experiments are relatively easy to perform and in which each of the competing models (in an unelaborated form at least) is falsified.

There are more serious problems though. First, it is generally found that error rate and mean reaction time covary (Banks & Atkinson, 1974; Forrin & Cunningham, 1973; Miller & Pachella, 1973), but little attempt has been made to relate accuracy and latency theoretically (Wickelgren, 1977). Second, it seems likely that each of the models could be elaborated to account for most of the empirical effects (Townsend, 1974), as long as mean reaction time is the only statistic considered, so that no resolution at this level of analysis can be expected. Third, the main statistic used to describe reaction time is mean reaction time. Because of this emphasis, few of the models are able to account for properties of reaction time distributions (Sternberg, 1975). This is important because distributional properties may prove to be decisive in evaluating models (Ratcliff & Murdock, 1976; Sternberg, 1975, Note 1, Note 2). For reviews of this area of research, see Corballis (1975), Nickerson (1972), and Sternberg (1969a, 1969b, 1975).

To model the Sternberg paradigm, I will use the same assumptions used in modeling the study–test paradigm. The probe is encoded, and parallel comparisons are made between the encoded probe and members of the search set. The search set is approximated by the memory set, each item having the same normal probability density function $N(v, \eta)$ of relatedness for nonmatching comparisons. The probability density function of relatedness for a match comparison is $N(u, \eta)$, where u varies with serial position and set size. v, a, and z may vary with set size, and η and s are kept constant at the same values used in the study–test paradigm. As before, a single constant time is assumed for probe encoding, response output, and so on.

Experiment 2 was performed to demonstrate the way in which the theory is applied to the Sternberg paradigm. The experiment was designed to replicate typical patterns of results found in the varied-set procedure while maximizing conditions for serial position effects (fast presentation rate and short probe delay). Serial position effects were sought because changes in accuracy and latency as a function of serial position can only be the result of changes in the single relatedness parameter u, as all criteria are fixed. Two subjects were each tested for eight experimental sessions in order to obtain reasonable estimates of error rates and reaction time distributions.

Experiment 2

Method. Two paid student volunteers served as subjects. Sequence generation, display, and response recording were controlled by a PDP-12A laboratory computer. The memory set was either three, four, or five randomly selected digits, from the set zero to nine, with no repetitions. The memory set was presented sequentially at .5 sec per digit. One-tenth second following the presentation of the last digit, a fixation point was displayed, which remained on for .4 sec. The probe digit immediately followed and remained on until a response was made. Each list length was presented equally often, and each serial position within each list was probed equally often. Also, there were equal numbers of old and new probe items. There was a 3-sec delay between trials, with a fixation point appearing for the last .5 sec of that period. Responses were made by depressing one of two buttons on a response box, right hand for "yes" responses and left hand for "no" responses. The subjects served in 8 sessions preceded by 1 practice session. Each session consisted of 10 blocks of 48 trials. Subjects were instructed to be as fast as possible while maintaining high accuracy. Feedback on accuracy performance was presented at the end of each session.

Results. Figure 13 shows mean latency as a function of set size for hits and correct rejections, together with linear least-squares regression lines. These show that the basic finding, namely, parallel, linear set-size–latency functions, is replicated. Figure 13 also shows serial-position–latency functions, which have a large recency effect and replicate Burrows and Okada (1971), Corballis (1967), Corballis et al. (1972), and Forrin and Cunningham (1973).

Figure 14 shows detailed fits of the mathematical model to the data. As noted earlier, the mathematical model was fitted to the summary statistics (proportion correct and

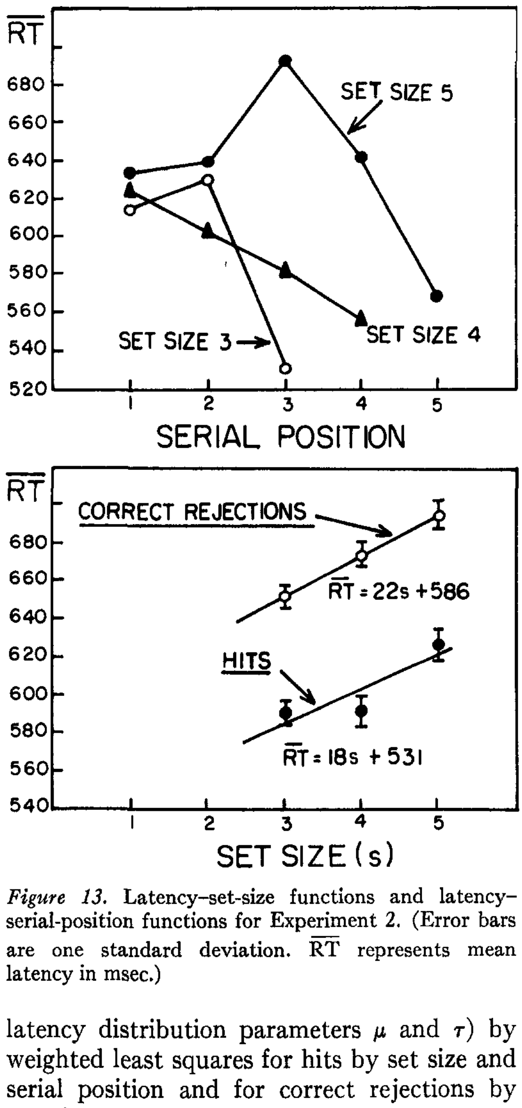

Figure 13. Latency–set-size functions and latency–serial-position functions for Experiment 2. (Error bars are one standard deviation. $\overline{RT}$ represents mean latency in msec.)

latency distribution parameters μ and τ) by weighted least squares for hits by set size and serial position and for correct rejections by set size. The encoding and response output parameter T_{ER} was kept constant across all these conditions, but it also turned out that to a good approximation, random walk boundary criteria a and z and relatedness of probe–nontarget item comparisons v were constant across set size. Thus, latency and accuracy for correct rejections changed as a function of number of items in the memory set (no additional degrees of freedom once a, z, and v were fitted). Latency and accuracy for hits varied only as a function of probe–target item relatedness u (and memory set size); thus, across serial position, the single parameter u varied to produce fits to the three parameters determined from the data, that is, latency distribution parameters μ and τ and accuracy. The fits to the data, although not perfect, show no more than four deviations of theory from data greater than about two standard deviations out of 45 fitted statistics.

Figure 15 shows fits of the theoretical distributions to the group reaction time distributions. The theoretical distributions capture the main features of the empirical distributions, and changes in distribution shape as a function of set size and serial position are well modeled by changes in the size of memory set and relatedness u.

Model Freedom

At this point, it may be argued that the Sternberg exhaustive serial scanning model has far fewer parameters than the random walk model presented above. If mean reaction time as a function of set size is to be fitted, then the Sternberg model has 3 parameters and the retrieval theory (for the above data) has 16 parameters (T_{ER}, v, a, z, and 12 values of u) or 7 if serial position effects are averaged. However, if distributions are to be modeled, the number of parameters in the Sternberg model jumps to 9 (Sternberg, Note 1): 4 parameters for the comparison stage distribution and 4 parameters for the base (encoding, decision, and response output) distribution (plus one parameter for the yes–no intercept difference). Further, if error rates as a function of set size are to be fitted, then perhaps as many as 6 more parameters have to be added. The number of parameters in the two models is now comparable. If serial position functions are considered, then as noted earlier, the Sternberg model is falsified, but the retrieval theory does a good job. Thus, the problems of model freedom and model comparison are not as simple and straightforward as might be imagined.

Perhaps the major difference between the retrieval theory and the Sternberg model can be summarized as follows: The retrieval theory (unlike the Sternberg model) provides an intrinsic tie-up between reaction time and accuracy, whereas the Sternberg model (unlike the retrieval theory) can make the advance prediction that the reaction-time–set-size function is linear.

Further reduction in the number of parameters in the retrieval theory might be accomplished if the relation between relatedness u and serial position could be described by a few-parameter functional relation, for example, exponential decay. I have not attempted to fit such a functional relation because at present it adds nothing theoretically. However, if a theory of recency effects were developed, then the prediction of relatedness as a function of serial position would be an important test of the conjunction of that theory with the retrieval theory.

Error Reaction Times

In typical reports of Sternberg varied-set experiments, error reaction times are rarely

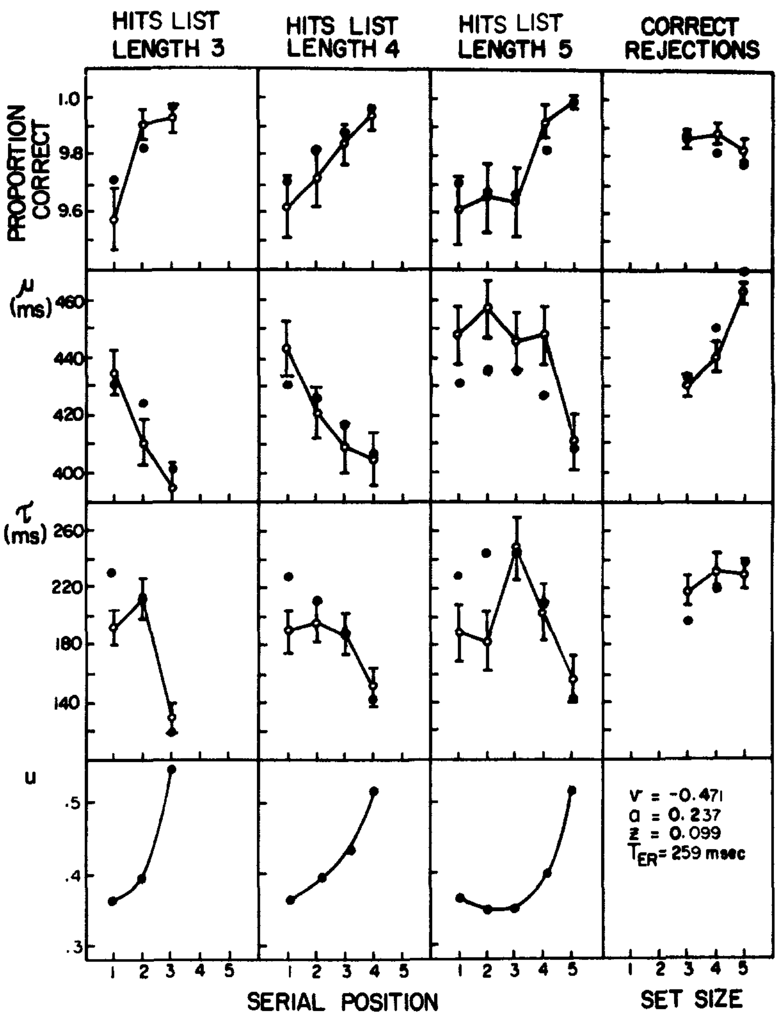

Figure 14. Accuracy, latency distribution parameters (from the convolution model) μ and τ, and relatedness u as a function of serial position and list length for hits, together with accuracy and latency distribution parameters μ and τ as a function of set size for correct rejections in Experiment 2. (The filled circles represent the theoretical values. For correct rejections, only number of processes in the search set changes with set size; for hits, only relatedness u changes with serial position. All other theoretical parameters [v, a, z, and T_{ER}] are fixed at the values shown in the bottom right-hand panel. Error bars are one standard deviation. ms = msec.)

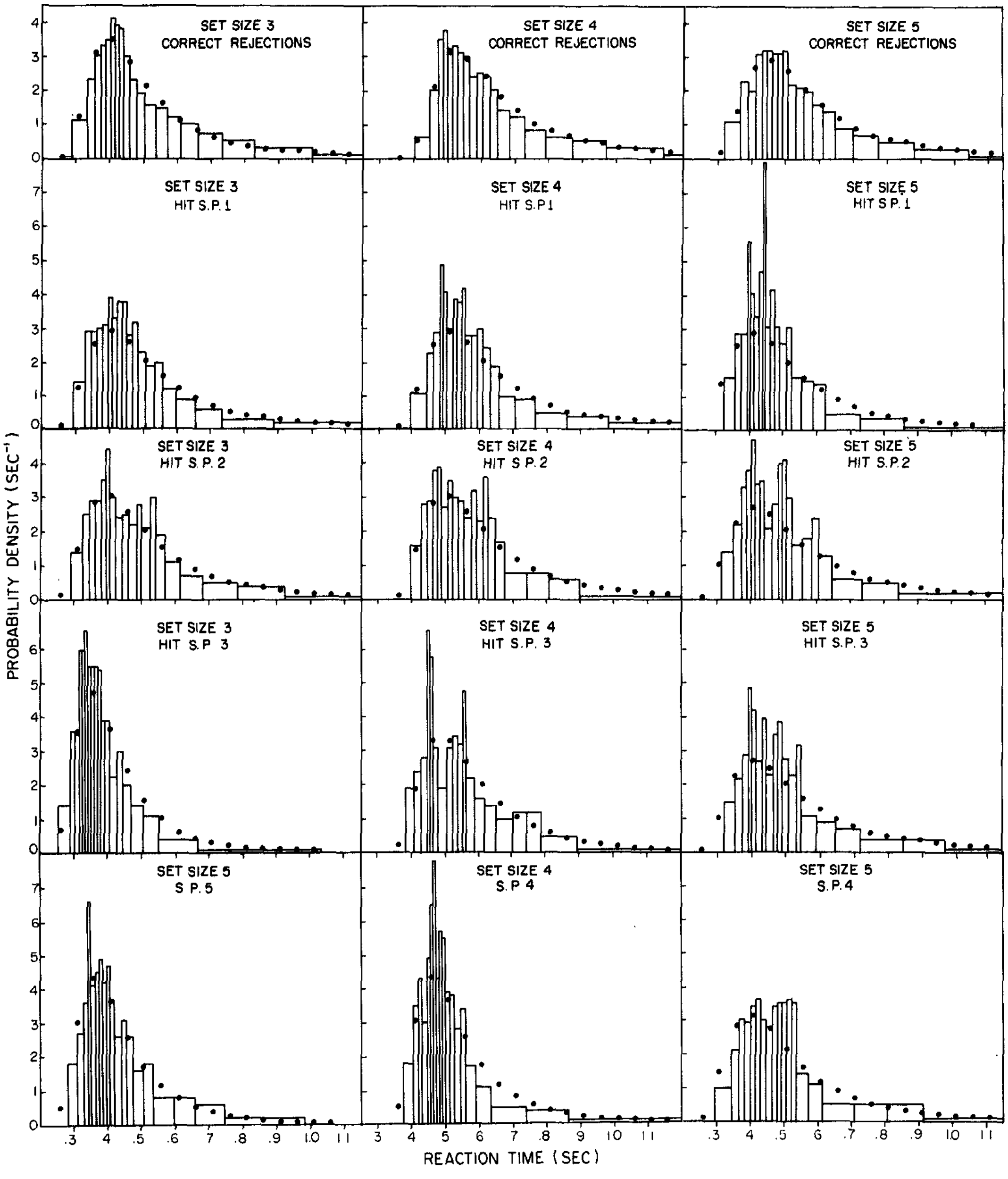

Figure 15. Group reaction time distributions (bar graphs) and theoretical fits (dots) for correct rejections by set size and for hits by set size and serial position (S.P.) for Experiment 2. (These data and fits should be compared with the distribution parameters μ and τ in Figure 14.)

presented because error rates are usually low, so that error latencies are unreliable, and also because error latencies are not of theoretical interest. However, error latencies can be useful in deciding between theories (Corballis, 1975; Murdock & Dufty, 1972). Corballis (1975) reported that in his experiments using the Sternberg procedure, error latencies were generally smaller than latencies for correct responses. In Experiment 2 above, error latencies for Set Sizes 3, 4, and 5 were 940 msec, 760 msec, and 1,005 msec for misses (average of 902 msec) and 1,307 msec, 909 msec, and 1,055 msec for false alarms (average of 1,090 msec), respectively. Predictions of mean reaction time from the theory are far too large. For typical values of parameters used to fit Experiment 2 (u = .435, v = −.437, a = .237, z = .099, and Set Size 4), reaction time for false alarms is 2.05 sec; for misses, it is 1.59 sec.

Error latency depends to a large extent on the tails of the relatedness distributions. Therefore, if the choice of normal distributions were inappropriate, everything else in the theory might still be tenable. For example, suppose the relatedness distributions have extremely long tails, so that about 1% to 2% of the total probability density resides in these tails. Then, theoretical values of latency and accuracy of correct responses may not be altered much, but error latency may change quite significantly.

Figure 16 demonstrates the effect of the shape of the distribution and tails of the distribution on latency and accuracy. For reference, the bottom panel of Figure 16 shows the normal distribution of relatedness for a nonmatching comparison with correct and error latencies and error rate. The middle panel shows the same normal distribution, with 3% of the probability density removed and added as a rectangular distribution over the range of relatedness −2 to +1. The error rate is increased a little; correct latency is the same as correct latency in the bottom panel, but error latency is reduced to 1.3 sec, which is near the experimental result.

The top panel of Figure 16 shows just how far it is possible to change the assumption of normal relatedness distributions without affecting correct reaction time or accuracy. The large rectangle contains 97% of the probability density; and as in the middle panel, the small rectangle, between −2 and +1 relatedness, contains 3% of the probability density. With this assumed relatedness function, mean reaction time for correct responses decreases to .62 sec, with the leading edge (μ parameter in the convolution model) of the reaction time distribution increasing by 30 msec and the τ parameter in the convolution model decreasing by 50 msec. The reaction time distribution is still very similar in shape to those shown in

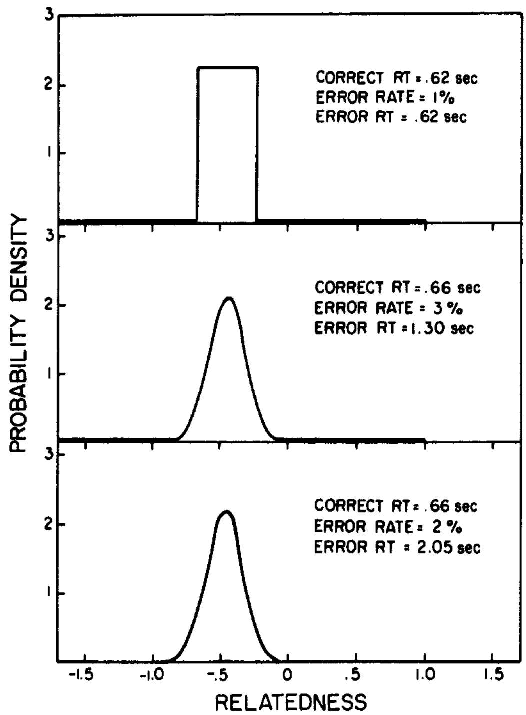

Figure 16. Effect of the shape of the relatedness distribution on correct and error reaction time (RT) and error rate. (The bottom panel shows a normal distribution of relatedness; the middle panel shows a normal distribution, with 3% of the probability density forming a rectangular distribution between −2 and +1 relatedness; the top panel shows a rectangular distribution of relatedness with the same 3% tail as in the middle panel).

Figure 15. The error rate drops to 1%, all errors coming from the tail of the small rectangular distribution that extends into positive relatedness. But now error latency has decreased to 620 msec. The assumption of rectangular distributions has some precedent in that Luce (1963) and Krantz (1969) have both used rectangular distributions in high-threshold models to account for signal detection.

There are three conclusions that can be drawn from this exercise. First, reaction time distributions for correct responses are rather insensitive to the shape of the relatedness distributions so long as there is a lump of probability density around the appropriate relatedness value. Second, error rates depend on the amount of probability density near and greater than the zero point of relatedness. Third, error latency depends critically on the

Table 3
Theoretical Predictions for Latency and Accuracy of Hits for Repeated Items for Parameter Values in Experiment 2

Response type	u	Accuracy	Mean latency	μ	σ	τ
Once-presented hit	.35	.966	686	433	47	253
Twice-presented hit	.35	.999	566	429	44	137
Once-presented hit	.435	.988	605	420	42	185
Twice-presented hit	.435	1.0	514	415	38	99
Once-presented hit	.55	.998	592	404	36	125
Twice-presented hit	.55	1.0	465	399	32	66

Note. Mean latency, μ, σ, and τ are in msec. Other parameter values are as follows: T_{ER} = 259 msec, $v = -.471$, $a = .237$, $z = .099$, and memory set size = 4.

shape of the tail of the relatedness distribution near the zero point of relatedness. Thus, reaction time distributions and error rates are reasonably independent of relatedness shape so long as there is a lump of probability density near the mean relatedness value of the normal distribution used to fit the experimental data and so long as there is a reasonable amount of overlap. This shape-independent property suggests that the correct focus of the model is on error rates and correct response latencies rather than on error latencies.

Reed (Note 4) has some evidence that the signal distribution (in a signal detection analysis of bilingual recognition) changes shape over the time course of recognition. The signal distribution starts off as multimodal, with a signal peak in the middle of the noise distribution, and then quite quickly becomes unimodal. This kind of change in relatedness distributions (the theoretical equivalents of signal and noise distributions in signal detection analysis) would give rise to fast errors in the retrieval theory and so may be an alternative to non-normal unimodal relatedness distributions. Although this finding is a little tentative, little is known about the time course of recognition, and we should at least keep such possibilities in mind.

One further point is that just as very long outlier reaction times are often assumed to be spurious, a certain proportion of errors may also be spurious. For example, subjects often report that they simply hit the wrong button. Therefore, to get meaningful and stable error latency data, it is better to work with a paradigm with reasonably large error rates, such as the study–test paradigm, rather than a paradigm with error rates around 1% of the total responses.

Repetition Effects

Baddeley and Ecob (1973) studied performance on memory sets containing repetitions using a varied-set Sternberg procedure and found faster recognition for repeated items. The effects on reaction time (50 msec to 100 msec) were larger than those observed in the study–test paradigm (30 msec to 40 msec; see Table 2). Changes in accuracy were not reported, probably because accuracy was near 100%, and there were not enough observations to allow the detection of significant differences.

There are two possible ways to model the effect of repetitions. For repetition effects in the study–test paradigm, I assumed two matching racing processes. The alternative is to assume one representation twice as strong, that is, a compressed memory set (presentation of Digits 3238 would lead to Memory Set 238 with 3 stronger than 2 or 8). However, Baddeley and Ecob (1973) ruled out this possibility by noting that nonrepeated items from sequences with repetitions were no faster than items from sequences (of the same size) with nonrepetitions.

To model the repetition result, I took parameter values used in Experiment 2 and computed accuracy and latency for one and then two racing matching processes. Results are presented in Table 3 and show that twice-presented items are between 64 msec and 120 msec faster than once-presented items. Thus, the theory gives quite good quantitative agreement with the result presented by Baddeley and Ecob (1973).

Discussion and Summary of the Sternberg Paradigm

There are many variations on the Sternberg paradigm published in the literature, but none provide qualitative problems for the theory. For example, stimulus probability effects in the fixed-set procedure have provided problems for the serial exhaustive scanning model (Miller & Pachella, 1973; Sternberg, 1975; Theios et al., 1973; Theios & Walter, 1974). In Link's (1975) random walk model for choice reaction time, stimulus probability effects provide the basis of a successful test of the theory. In effect, stimulus probability affects the random walk boundary positions. Therefore, it seems likely that the memory retrieval theory could account for stimulus probability effects in much the same way.

There have been several speed–accuracy studies using the Sternberg paradigm, but because the question of speed–accuracy tradeoff is central to the retrieval theory, the topic is discussed later in a separate section.

To summarize, the Sternberg paradigm is modeled in much the same way as the study–test paradigm. Parallel comparisons between the probe and members of the search set are made. The decision process is self-terminating on matches and exhaustive on nonmatches. The search set is approximated by the memory set, even though representations of items from earlier trials (with probe–item relatedness lower than probe–memory-set item relatedness) may have to be included to account for effects of recency of negative probes (Atkinson et al., 1974). An experiment was performed to demonstrate fits of the mathematical model to data. All criteria were fixed as a function of set size, and changes in accuracy and latency of correct rejections as a function of set size were well fitted as a result of search set size changing (no free parameters to produce those changes). The covariation of accuracy and the shape of latency distributions as a function of serial position was well accounted for by changes in the single parameter u (relatedness for a probe–memory-set item comparison).

The fits to error latency were poor, and this led to an investigation of the dependence of fits of the model on the form of the relatedness distributions. Results showed that mean latency and the shape of latency distributions for correct responses are rather insensitive to the shape of the relatedness distribution so long as there is a lump of probability density around the appropriate relatedness value; further, error rate does not depend on the shape of the relatedness distribution but on the amount of probability density around the zero point of relatedness; finally, error latency is dependent on the tail of the relatedness distribution near the criterion. Although the fits to error latency were poor, it was shown that better fits could be obtained by changing the shape of the relatedness distribution, and such changes could be made without affecting predictions about correct latency or error rate. Therefore, until a theory of the structure of the memory trace can be developed that will specify the relatedness distribution, the most useful data for testing the mathematical model appear to be error rates and latency distributions for correct responses.

Supraspan Prememorized List Paradigm

It has been argued that in order to compare reaction time results from experiments using long lists (above memory span) with results from experiments using short lists (subspan), error rates should be low and approximately the same. To produce low error rates with long lists, prememorized lists have been employed. An experiment that shows the way in which the theory applies to the prememorized list paradigm is the Burrows and Okada (1975, Experiment 1) study. This experiment used the Sternberg fixed-set procedure, with prememorized memory-set items and list length varying from 2 to 20. Each word in the positive set was tested once in the test sequence for that list, together with an equal number of new words. Subjects had to press one button to indicate that the test word was in the study list, another button to indicate a new word.

The results showed parallel slopes for positive and negative responses as a function of set size and low error rates. Reaction time as a function of set size showed an apparently discontinuous function. Subspan reaction times were fitted by a linear function with a slope approximating that typically found in the Sternberg paradigm (35 msec to 55 msec per item). Supraspan reaction times were fitted by a linear function with a much lower slope (13 msec per item). It was argued that the discontinuity was evidence for separate long-term and short-term retrieval processes. However, in the same article, it was shown that a continuous function (logarithmic) did almost as good a job of fitting the data as the two linear functions.

I now demonstrate that the data from the Burrows and Okada (1975) study can be well fitted by the random walk retrieval theory, which assumes only one kind of retrieval process. It should be noted that in terms of the retrieval theory, it is not necessary to have low error rates in studying retrieval from long lists. The study–test paradigm uses supraspan lists, and error rates can be as high as 25%.

In order to fit the mathematical model to the data, it is necessary to estimate reaction time distribution parameters using the empirical convolution model, as was done with other paradigms. There were between 70 and 90 observations per set size per subject, which is barely enough to allow stable estimations of distribution parameters. The fits of the convolution model were performed, and it was immediately apparent that the σ parameter (the standard deviation in the normal component of the convolution) was far larger than in any fits to any other paradigm: The average value of σ across subjects and set size was 104 msec. The large value of σ (and inspection of the distributions) showed that the reaction time distributions had slowly rising leading edges, sometimes with short outlier reaction times.

I was at a loss to explain why these reaction time distributions were much less skewed than usual until I noted that the test words were presented verbally. A recent article by Morton, Marcus, and Frankish (1976) suggests dangers in this procedure. They demonstrated that if the physical onset of a series of words is fixed, the time for perception of the words can be extremely variable. For example, the delay between onset and perceptual center of the spoken digits "seven" and "eight" differs by 80 msec. Burrows and Okada used two-syllable words, and it is likely that the perceptual centers of those words could vary even more than 80 msec. Of course, if the main reaction time statistic considered is mean reaction time, then there is no problem—just an addition of a little noise—but if distributions are considered, then the use of verbal probes can lead to fast (as well as slow) outliers.

To obtain better estimates of distribution parameters, I trimmed off 5% of the fast reaction times and refitted the convolution model. The average value of σ was reduced by 47 msec to 57 msec. The main effect on the other distribution parameters was to increase τ, leaving μ about the same as before trimming. The retrieval theory was applied as in the study–test and Sternberg paradigms. Parallel comparisons are made between the encoded probe and members of the search set. The search set is approximated by the experimental memory set, and each item has a normal distribution of relatedness $N(v, \eta)$ for non-matching comparisons. Each memory-set item has a normal distribution of relatedness $N(u, \eta)$ for matching comparisons, and the variance parameters η and s are kept constant at the values assigned earlier.

The fits of the theory to the data are shown in Figures 17 and 18. Figure 17 shows fits to the empirical reaction time distribution parameters μ and τ for hits and correct rejections. The error bars represent the maximum likelihood standard deviation estimates in parameters but do not take into account any subject-to-subject variability. Figure 18 shows fits of the theoretical error rates to data, together with behavior of theoretical parameters u, v, a, and z as a function of set size. The fits of the theory to data are by no means perfect. The fact that the theory cannot be made to give better fits demonstrates quite clearly that there are fairly powerful nonobvious constraints on the theory (before the fast reaction times were trimmed, the best fit was much worse; predicted miss rates were half the size of those shown in Figure 18).

Figure 18 shows values of u and v that allow a d' measure of discriminability to be calculated from $d' = (u - v)/\eta$. d' starts at a value of 5.3 at Set Size 2 and decreases to 4.8 by Set Size 10 and remains at that value to Set Size 20. The constancy of d' for longer lists can be interpreted as showing that the list prememorization criterion produced equivalent relatedness levels for the larger set sizes. The continuous change in d' as a function of set size shows quantitative continuity between subspan and supraspan retrieval in item recognition.

Much of the research done using this paradigm has been performed within the framework of the Atkinson and Juola (1973) model. In the experimental procedure used by Atkinson and Juola (1973), a study list was memorized prior to the experimental session. In any block of trials, test items consisted of both new positive items and new negative items, together with all previously presented items. Results showed that repetition of positive items made them faster and more accurate, whereas repetition of negative items made them slower and less accurate. The way in which the theory would deal with these results can be seen by referring to Figure 2. First presentations of positives and negatives would correspond to Processes 2 and 4, respectively, and repeated presentations would correspond to Processes 1 and 3, respectively, thus leading to the observed pattern of results.

Several studies have investigated the situation in which there are different sets of memory-set items stored in long- and short-term memory (Forrin & Morin, 1969; Scheirer & Hanley; 1974; Sternberg, 1969b; Wescourt & Atkinson, 1973). The general consensus seems to favor models that postulate parallel access to short- and long-term sets (Wescourt & Atkinson, 1976). For example, Wescourt and Atkinson (1973) used a procedure in which a 30-word long-term set was memorized preexperimentally. In an experimental trial, subjects were presented with a short-term set of one to four words (or zero in a control condition) before each test word. Test words came from either the short- or long-term set, with each having a probability of .25, or the test word was a negative. The set-size–latency function for probes from the short-term set

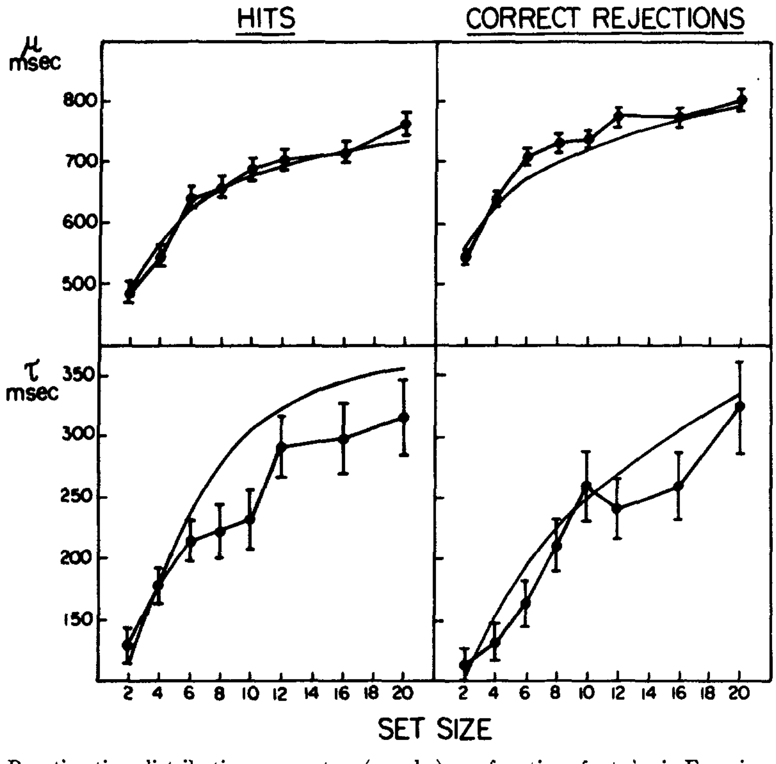

Figure 17. Reaction time distribution parameters (μ and τ) as a function of set size in Experiment 2 (Burrows & Okada, 1975), together with theoretical fits (continuous lines). (Note that T_{ER} = 420 msec. Error bars are one standard deviation.)

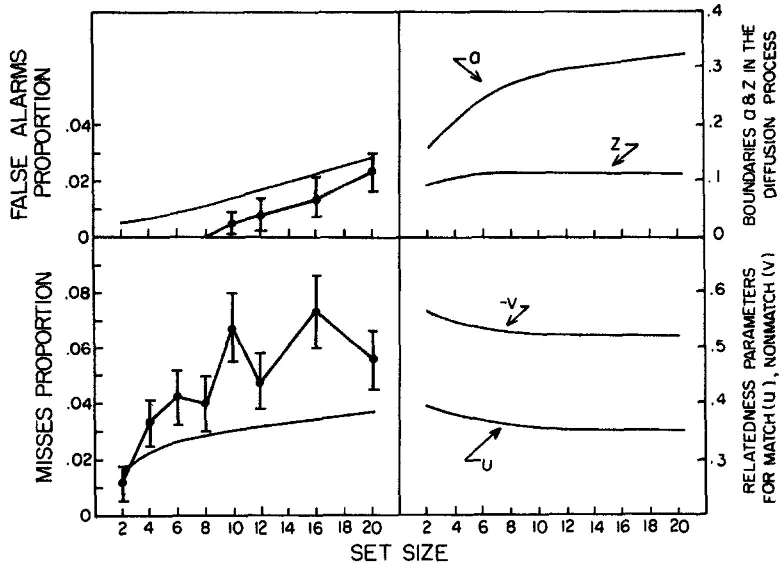

Figure 18. Error rates as a function of set size in Experiment 2 (Burrows & Okada, 1975), together with theoretical fits (continuous lines). (Also shown are theoretical parameters *u*, *v*, *a*, and *z* as a function of set size. Note that T_{ER} = 420 msec. Error bars are one standard deviation.)

was roughly linear with a slope of 21 msec per item. In contrast, the latency function for probes from the long-term set (as a function of short-term set size and excluding Set Size 0) was nearly flat, 1 msec per item slope, and about 100 msec slower than response to short-term probes. Negatives showed the same pattern of results with a slope of about 5 msec per item and responses 40 msec slower than responses to long-term probes. The specific model proposed by Wescourt and Atkinson (1973) is similar to the decision component of the retrieval theory, namely, parallel independent searches of the short-term and long-term sets, self-terminating on positives and exhaustive on negatives. To make the two models identical, it is only necessary to assume independent and parallel probe–memory-set item comparisons within both the long-term and short-term stores. Therefore, because processing is the same for both sets of items, there is no basis for separating short- and long-term stores in terms of processing.

Another finding that is consistent with the theory is the result that when material in the short-term set is discriminable from material in the long-term set, there is no effect of long-term set size on reaction time to short-term items and vice versa (Scheirer & Hanley, 1974). In contrast, when material in the two sets is not highly discriminable, latency is a joint function of short- and long-term set size.

Summary

Results from prememorized list procedures are modeled in the same way as results from the Sternberg and study–test procedures. Application of the theory to the Burrows and Okada (1975) study, in which list length was varied from subspan to supraspan, showed there is no need to assume separate retrieval processes for subspan and supraspan item recognition. Furthermore, in experiments where short-term and long-term sets are searched simultaneously, it seems that the best model for the data assumes parallel search of both sets in line with the retrieval schemes developed in this article.

Continuous Recognition Memory Paradigm

In the continuous recognition memory paradigm, single words are presented, and the subject has to press one button if the test word is new and another if the test word was presented earlier in the test list. This task is somewhat different from the three paradigms discussed up to this point in that items change from negative to positive on presentation, and the search set is continually increasing in size. The principal independent variable is lag between successive presentations of an item, and accuracy and latency are measured.

Okada (1971) performed two experiments that investigated performance on this task. In both experiments, mean reaction time as a function of lag was well approximated by a negatively accelerating exponential. I have fitted the mathematical model to Okada's (1971) Experiment 2 to show how the theory is applied. In his Experiment 2, items were presented once, twice, or three times, but only once- and twice-presented items were fitted because there were not sufficient data to perform distributional analyses on thrice-presented items. Even though the paradigm is different in many respects from the study–test and Sternberg paradigms, the same processing assumptions are made. The memory set changes from 0 to 150 in the course of a block of trials, but in order to approximate over the whole block, the search set size was set at 40 (though values between 30 and 60 would not change the fits very much). At any test position in the sequence, the test item can be either new or old with lag values of 0, 1, 2, 4, or 6. Therefore, in the theory, all criteria are fixed and only relatedness u may vary with lag. Thus, any single value of u must fit both accuracy and reaction time distributions (through the parameters μ and τ). Empirical results are presented in Figure 19 for proportion correct and reaction time distribution parameters μ and τ, together with theoretical fits, as a function of lag. In these fits, s and η are fixed at .08 and .18 as before. The three experimental statistics (accuracy and latency distribution parameters μ and τ) are adequately fit with just one theoretical parameter u, which varies as a function of lag.

Thus, the pattern of results found in the continuous recognition memory task can be modeled in the same way as results from the study–test and Sternberg paradigms.

Speed–Accuracy Trade-off Paradigms

Experiments to study speed–accuracy trade-off are usually performed within the context

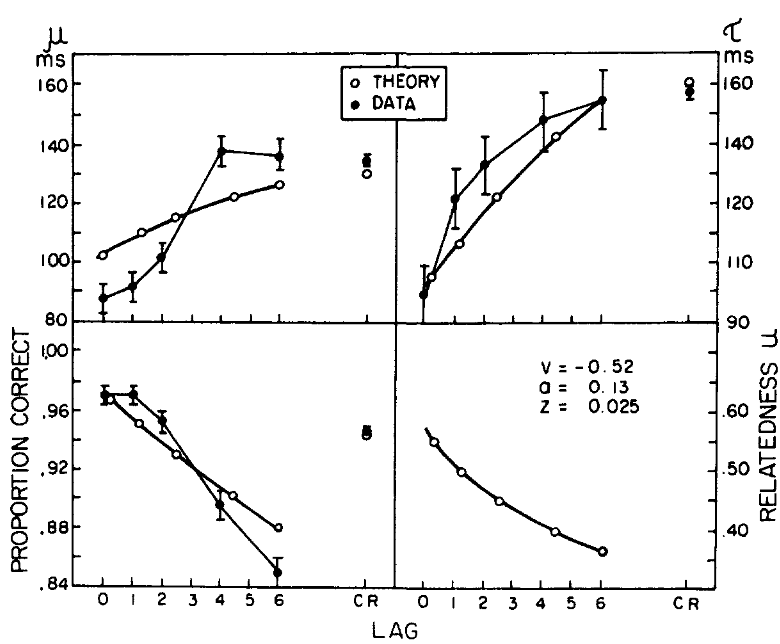

Figure 19. Proportion correct, reaction time distribution parameters μ and τ, and relatedness u for hits as a function of lag for Okada's (1971) Experiment 2. (Theoretical fits are shown by open circles; also shown are correct rejection data [CR]. The number of processes assumed to be in the search set was 40, and T_{ER} = 520 msec. Theoretical parameters v, a, and z are as shown in the figure.)

of a particular paradigm, for example, choice reaction time and the Sternberg paradigm (Pachella, 1974), the Brown–Peterson paradigm (Reed, 1973), sentence processing (Dosher, 1976), cued recall (Murdock, 1968), paired associate recall, and recognition (Wickelgren & Corbett, 1977). Because the topic is central to the theory developed in this article, it is dealt with in this separate section.

In the introduction, speed–accuracy tradeoff methods were partitioned into two classes (excluding latency partitioning). Information-controlled processing refers to the usual reaction time method in which the subject sets some limit on the amount of information required for a response, and time of response is not limited. Experimentally, the amount of evidence required for a response can be manipulated using instructions (e.g., speed vs. accuracy) or payoffs. Time-controlled processing refers to the class of experimental methods where the time of response is experimentally controlled using methods such as response signals, deadlines, or time windows.

Information-Controlled Processing

Each of the preceding item recognition paradigms is an example of information-controlled processing. To check that the retrieval theory is capable of accounting for results obtained when speed–accuracy criteria are changed, I shall show that changes in boundary criteria are sufficient to account for the kinds of changes observed in experimental data when speed–accuracy conditions are changed. Banks and Atkinson (1974) used the method of payoffs to investigate speed–accuracy effects in the Sternberg varied-set procedure. They found that with payoffs stressing accuracy, error rates were about 1%; whereas with payoffs stressing speed, error rates were between 10% and 30%, the slope of the latency–set-size function was halved, and the intercept was reduced by about 200 msec. To model this pattern of results, it is assumed that under speed conditions, the random walk boundaries are moved close together; whereas under accuracy conditions, the boundaries are moved relatively far apart (see Figure 5 for an illustration).

To check that the numerical values generated from the retrieval theory are compatible with the size of the effects obtained by Banks and Atkinson (1974), typical parameter values used in Experiment 2 were taken, and random walk boundary position parameters a and z were reduced to .04 and .015, respectively. Latency for both hits and correct rejections decreased by 300 msec, and average error rates increased from 2% to 18%. Thus, the theory produces effects that are quite consistent with results from the payoff method.

Payoff or instruction is usually a between-sessions variable; therefore, criteria can change between conditions, and such experiments do not provide a strong test of the theory. Nevertheless, the demonstration that speed and accuracy can covary over a wide range of values is important in evaluating models of retrieval processes. For further discussion of these methods, see Pachella (1974) and Wickelgren (1975, 1977).

Time-Controlled Processing

In this section, it is shown that time-controlled processing methods produce data that provide a strong test of the retrieval theory. One example of time-controlled processing is the response signal method that was developed to avoid the problem that criteria may change between conditions. In essence, the subject is given the probe followed by a signal to respond presented a variable amount of time after probe onset (Reed, 1973, 1976; Schouten & Bekker, 1967). Reed (1976) used the response signal method to investigate the time course of recognition in the Sternberg varied-set procedure. One, two, or four letters comprising the memory set were presented sequentially. Following probe onset, a signal to respond was presented at a lag varying from 7 msec to 4,104 msec. Subjects were required to respond "yes" or "no" as quickly as possible after the signal and to make a confidence judgment on how confident they felt at the time of response. There were two analyses of major interest: d' as a function of response signal lag and response latency as a function of signal lag. It was argued that the first of these indicated the time course of accumulation of evidence on whether the probe was a member of the memory set, and the second gave evidence that the decision is based on the termination of a discrete process. Before discussing the results, I will derive an expression for the increase in accuracy as a function of signal lag from the theory.

In order to make the problem tractable, let us assume that the match and nonmatch boundaries are placed far from the starting point, so that in the range of signal lags considered here, the diffusion process can be considered unrestricted. Also, let us consider only Memory Set Size 1. In an unrestricted diffusion process, the probability density function $h(x, t)$ at position x and time t with drift ξ and variance in the drift s^2 is given by

$$h(x, t) = \frac{1}{\sqrt{2\pi s^2 t}} e^{-\frac{1}{2}(x-\xi t)^2/s^2 t} \quad (7)$$

(Feller, 1968, p. 357). In the theory, however, relatedness ξ has a normal distribution (see Equation A24). Therefore, in order to calculate the distribution of evidence as a function of time, we must integrate over the distribution of relatedness $N(u, \eta)$ and find

$$y(x, t) = \int_{-\infty}^{\infty} h(x, t) N(u, \eta) d\xi$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi s^2 t}} e^{-\frac{1}{2}(x-\xi t)^2/s^2 t} \frac{1}{\sqrt{2\pi\eta^2}} \times e^{-\frac{1}{2}(u-\xi)^2/\eta^2} d\xi;$$

therefore,

$$y(x, t) = \frac{1}{\sqrt{2\pi}\sqrt{t(\eta^2 t + s^2)}} e^{-\frac{1}{2}(x-ut)^2/[t(\eta^2 t + s^2)]}. \quad (8)$$

Thus, the distribution of evidence is given by a normal distribution $N[ut, \sqrt{t(\eta^2 t + s^2)}]$. For two distributions of relatedness, $N(u, \eta)$ for matches and $N(v, \eta)$ for nonmatches, d' is given by

$$d' = \frac{u - v}{(\text{variance})^{\frac{1}{2}}} = \frac{u - v}{\eta\sqrt{[1 + s^2/(\eta^2 t)]}}. \quad (9)$$

Now as $t \to \infty$, the value of d' asymptotes (ASY) at $d'_{\text{ASY}} = (u - v)/\eta$, that is, d' of the original relatedness (as expected, given a large amount of time). Thus,

$$d' = \frac{d'_{\text{ASY}}}{\sqrt{1 + s^2/(\eta^2 t)}}. \quad (10)$$

Equation 10 gives an expression for the growth of d' as a function of time for Memory Set Size 1. For Set Sizes 2 and 4, the decision rule must be taken into account: If at the time of the response signal, there is evidence from at least one process greater than the starting value z, then a "yes" response is produced; otherwise, a "no" response is produced (see Equation A23). Therefore, to calculate d' values for Set Sizes 2 and 4, it is necessary (a) to calculate d' values from Equation 10, (b) to use the d' values to give values of hit rate and false alarm rate, (c) to combine the hit rate and false alarm rate according to the decision rule, and (d) to thus produce a hit rate and false alarm rate for the combined process from which a d' value can then be calculated.

Let p be the hit rate and q the false alarm rate for Set Size 1 at some time t (derived from the d' calculated from Equation 10). Then, for n processes, the probability of a nonmatch w is given by

$$w = 1 - (1 - q)^n, \quad (11)$$

and the probability of a hit r is given by

$$r = 1 - (1 - p)(1 - q)^{n-1}, \quad (12)$$

by analogy with Equation A19. Thus, r and w depend on the particular values of p and q, so that for a particular discriminability value $d' = (u - v)/\eta$, the discriminability for n processes depends on the criterion chosen, and subjects could maximize the d' value by adjusting the criterion. The dependence of discriminability d' on criterion value is rather an undesirable property; however, the dependence is rather small. Consider $d' = 1.85$; values of p and q that can give this value are $p = .86$, .75, and .54 and $q = .22$, .12, and .04, respectively. Using Equations 11 and 12, values of d' for $n = 2$ are 1.51, 1.52, and 1.54, respectively, and for $n = 4$ are 1.14, 1.20, and 1.27. Thus d' is reasonably robust to the choice of criterion. In Reed's study, hit rates were usually high (80% to 90%), so that these larger p values were used in the following fits.

Figure 20 shows a plot of the accuracy–response-signal-lag data obtained by Reed (1976), together with fits from Equation 10 and d' values for Set Sizes 2 and 4, calculated using Equations 11 and 12. The encoding and response output parameter was 220 msec, and

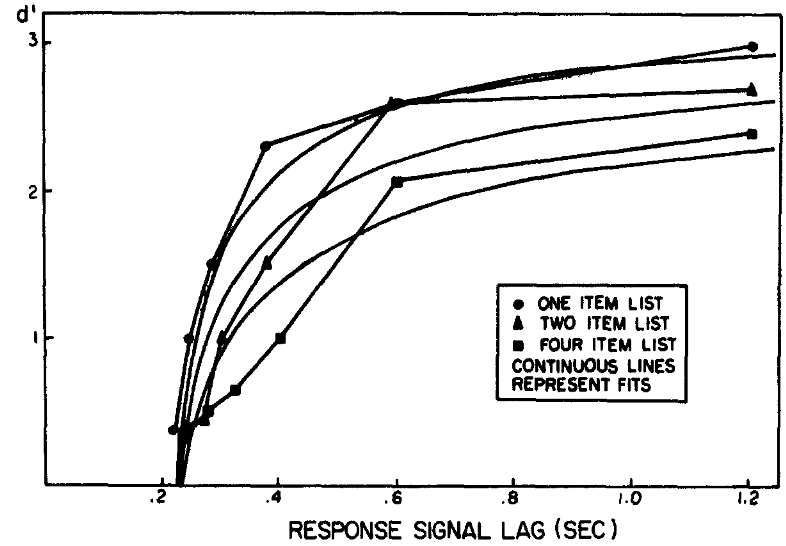

Figure 20. d' as a function of signal lag for a speed–accuracy experiment by Reed (1976) using response signals in the Sternberg varied-set procedure. (The broken line represents data, and the smooth curves represent fits [derived from Equations 10, 11, and 12].)

d'_{ASY} for Set Size 1 was set at 3.2. It should be noted that each of the three theoretical curves was fixed in shape by the initial choice of s^2 and η^2 made at the start of fitting the study–test paradigm, so that shape is described with no free parameters.

Reed (1976) has argued that the response-latency–signal-lag function is particularly useful in discriminating models. These curves show a minimum at some particular lag, depending on type of response (yes–no) and list length. It was argued that these provide evidence for the termination of a discrete process that leads to an accelerated response latency for a short period. The difference between the minimum for "yes" and "no" responses was used to argue for differential termination of the "yes" and "no" decision process. These two results are exactly what the theory predicts if it is now assumed that the diffusion process boundaries are close to the starting point of the process but not close enough to significantly affect the shape of the theoretical accuracy–signal-lag curves. It should be noted that the latency minima correspond to a significant increase in response latency rejections ($100 \text{ msec} < T < 500 \text{ msec}$ only accepted; see Table 1 of Reed, 1976), and this may be a problem for interpretation of these results.

Confidence Judgment Procedures

The study–test paradigm employs a 6-point confidence judgment procedure for response recording. Up to this point, the task has been modeled as though it used a yes–no procedure because results from a yes–no procedure show the same pattern of data. In order to model (qualitatively) the confidence judgment procedure, it is necessary to assume a variable internal temporal deadline. In terms of the theory presented here, the only information available about the "strength" of the item during the recognition process is the comparison time. Thus, if a comparison is taking a long time, the subject may reduce his confidence and respond on a lower confidence key. It should be noted that subjects are quite good at responding to external deadlines (Reed, 1976) and so may be able to set internal deadlines. Thus, confidence judgment results may be explained in terms of subject-set internal deadlines on the comparison process.

Summary

Speed–accuracy trade-off is a phenomenon that can probably be induced in all reaction time experiments. Therefore, any retrieval model that deals adequately with both latency and accuracy must have mechanisms at the heart of the model to deal with this trade-off. The theory developed here was designed with this in mind, and so speed–accuracy relations are fundamental. The theory is quite consistent with results from several speed–accuracy procedures and fits results from the response signal procedure (Reed, 1976) especially well. In the theory, two variance parameters (η^2 and s^2) were fixed throughout fits to previous paradigms. The shape of the accuracy–signal-lag function was determined solely by the ratio of those two parameters (s^2/η^2), and the fit was excellent.

Discussion

Comparison and Summary of Paradigms

Each of the experimental paradigms considered so far has been discussed more or less in isolation from the others. Therefore, it is appropriate to give a brief summary and comparison of experimental parameters between paradigms. The study–test, Sternberg, prememorized list, and continuous recognition paradigms are all assumed to have the same processing stages (see Figure 1). The encoded representation of the probe is compared in parallel with members of the memory search set. The search set is defined as those items determining processing rate and is to be conceptualized as the set of resonating elements in the framework of a resonance metaphor. Each comparison process is formalized as a diffusion process. In the diffusion process, evidence for a probe–memory-set item match is accumulated over time: the greater the evidence (greater the resonant amplitude), the more likely a match; the lower the evidence (smaller the resonant amplitude), the more likely a nonmatch (see Figure 2). The rate of accumulation of evidence is determined by the relatedness of the probe to a memory-set item (see Figure 3), where relatedness is a single-dimension variable onto which all structural, semantic, and phonemic (etc.) information is mapped. The decision process is self-terminating on a comparison process match, leading to a "yes" response. For a "no" response, all comparison processes must have terminated with a nonmatch.

The search set was assumed to be the memory set for the study–test, Sternberg, and prememorized list paradigms, though it was noted that this is only an approximation because there is evidence that items from earlier lists enter the comparison process. For the continuous recognition memory paradigm, a search set size of 40 was assumed.

Table 4 shows parameter values for the four item recognition paradigms for the two conditions in each experiment in which the largest value of d' and the smallest value of d' were obtained. The theoretical value of d' represents the difference in the size of resonance between an item in memory and the probe when the probe matches and when the

Table 4
Parameter Values for Four Item Recognition Paradigms for the Two Conditions Giving the Largest and Smallest d′ *Values*

	Largest d' value					Smallest d' value					
Paradigm	a	z	u	v	d'	a	z	u	v	d'	T_{ER}[a]
Study–test (Experiment 1)	.15	.04	.45	−.44	5.0	.16	.04	.19	−.38	3.2	360 msec
Sternberg (Experiment 2)	.24	.1	.55	−.47	5.7	.24	.1	.35	−.47	4.6	260 msec
Prememorized lists	.16	.09	.4	−.55	5.3	.33	.11	.36	−.52	4.9	420 msec[b]
Continuous recognition	.13	.03	.6	−.52	6.2	.13	.03	.35	−.52	4.8	520 msec[c]

[a] In the response signal study by Reed (1976), T_{ER} = 230 msec and d' = 3.2.
[b] The test words were spoken, and spoken words have a larger stimulus onset to perception than visually presented words.
[c] Stimuli were presented on an IBM typewriter advanced by a solenoid, the pulse triggering the solenoid, which started a timer. There is a 250-msec delay (approximately) from timer start to stimulus onset with this apparatus (compare Experiments 1 and 2 of Murdock et al., 1977, with stimuli presented on a computer and stimuli presented on a typewriter).

probe does not match the item. Therefore, to calculate asymptotic d' for any experiment, it is necessary to obtain a hit and false alarm rate from the d' value, to use Equations 11 and 12 to derive hit and false alarm rates for the combination of n processes in the memory set, and then to use the derived hit and false alarm rates to estimate asymptotic d'. It turns out that in all four experiments, asymptotic accuracy is not much more than the accuracy obtained in the experiment, which suggests that subjects are working with an emphasis on accuracy.

Parameters show no more variation within a paradigm than a factor of two; and between paradigms, variations are no more than a factor of three. Values of v, that is, probe–memory-set item nonmatch relatedness, are relatively invariant across paradigms, showing that subjects tend to set a relatively constant relatedness criterion with respect to the nontarget item relatedness (noise) distribution.

The encoding and response output parameter (T_{ER}) seems to fall in the range of 250 msec to 350 msec, so that the other 250 msec to 500 msec is comparison and decision time. This division of processing time seems much more reasonable than that proposed by other models, for example, 35-msec comparison and 400-msec encoding, decision, and response output (Sternberg, 1966) or 5-msec comparison and 600-msec encoding, decision, and response output (Murdock & Anderson, 1975).

It must be apparent that such cross-paradigm comparisons may prove most interesting when the same group of subjects is used in each of the paradigms. Furthermore, such comparisons may prove useful in studying the locus of individual differences.

Relation to Neural Network Models

A large part of this article has been concerned with the development of a theory that integrates a number of experimental paradigms. In this and the following section, the complementary problem is treated, that of the relation between this theory and other more global theories. I consider neural network models in this section and semantic network models in the next section.

Perhaps one of the simpler neural network models is the model developed by J. A. Anderson (1973), which deals with retrieval from short memorized lists. It should be stated at the outset that this model is not to be viewed as definitive but rather as the first step of a continuing and developing research program (Anderson, 1972, 1976; Anderson, Silverstein, Ritz, & Jones, 1977; Borsellino & Poggio, 1973; Cavanagh, 1976; Kohonen, 1976; Kohonen & Oja, 1973). The basic structure of the model is fairly simple. It is assumed that what is of importance is the simultaneous pattern of individual neuron activities in a large group of neurons. The patterns interact in the process of storage, so that memory is simply the sum of past activities in the system. In the recognition process, the probe input trace is matched to the memory. If a certain level of positive evidence is accumulated, a "yes" response is made, and if a certain level of negative evidence is accumulated, a "no" response is produced. The memory is formally represented by a vector of N elements, and a memory trace $\underset{\sim}{f}_i$ is represented as a specific vector of length N. If there are K traces, then the memory vector is constructed by

$$\underset{\sim}{s} = \sum_{k=1}^{K} \underset{\sim}{f}_k. \qquad (13)$$

Then, the match between the input probe $\underset{\sim}{f}$ and the memory vector is given by the dot product $\underset{\sim}{f} \cdot \underset{\sim}{s}$ (or matched filter), that is, each element of the input vector is multiplied with the corresponding element of the memory vector, and the sum of these values represents the amount of match.

If this description is rephrased slightly, then the recognition model shows remarkable similarity to the retrieval model developed in this article. For example, the comparison process can be viewed as resonance, and if each element of the dot product is accumulated successively through one filter (not two), then the resulting process is a random walk. Thus, the two theories are closely related.

However, there are some problems with Anderson's (1973) model. The model suffers many of the problems of strength theory in that there are no separate records of each memorized item, and activities simply sum, as in strength theory. Therefore, the model is not capable of performance levels shown by subjects in judgments of frequency (Wells, 1974) and list discrimination (Anderson & Bower, 1972).

I noted earlier that this model is part of a continuing research program, which can be seen in J. A. Anderson (1976) and Anderson et al. (1977). In these articles, it is shown how network models can deal with associations, probability learning, feature analysis and decomposition, and categorical perception. Therefore, it may be a useful future exercise to attempt to relate in some detail these two theories: neural network theory and the retrieval theory presented here.

Relation to Semantic Network Models and Propositional Models

Semantic processing models. There is no doubt that the human memory system is highly structured. One approach to modeling the structural complexity of the system has involved the use of semantic networks. The use of a network representation has been in part a consequence of the appealing digital computer metaphor for the human information processor. However, just because this form of representation is appealing and most useful in setting out structural relations, it does not mean that process models suggested by a network representation are true or useful. This point can be illustrated as follows.

Rips, Shoben, and Smith (1973) developed a set theoretic model of semantic representation. Later, Hollan (1975) pointed out that set theoretic representations can be translated into networks, that is, there exists an isomorphism between the two representations. However, Rips, Shoben, and Smith (1975) in a rejoinder made the important point that the choice of representation can lead to substantive processing differences. For example, network models explain reaction time effects in terms of the time for retrieval of pathways between nodes; whereas for set theoretic models, reaction time effects are determined by comparisons of semantic elements. Thus, for any network model, there is at least one other isomorphic representation that suggests alternative processing assumptions.

Much of the data used to support semantic processing models makes use of reaction time and, in particular, the statistic mean reaction time. One of the main thrusts of this article has been to demonstrate that mean reaction time is of limited use as a statistic. Also mean reaction time can often be misleading because models based on mean latency may be contradicted by further analyses of reaction time data (e.g., Ratcliff & Murdock, 1976). Furthermore, in many item recognition studies, it is found that accuracy and latency covary, and this seems to be true in many semantic memory studies (e.g., Collins & Quillian, 1969; Meyer, 1970; Rips et al., 1973). Therefore, any model that purports to account for latency effects should also account for accuracy effects and be capable of dealing with speed–accuracy trade-off.

Process models of retrieval latency that are derived from network models assume that latency is a function of number of links traversed. Latency for "yes" responses is well modeled using this assumption, but latency for "no" responses poses serious problems. For example, in verification of statements such as "a robin is a mammal," latency can be as much as 200 msec slower than for statements in which the two nouns are more unrelated semantically, for example, "a robin is a car" (Rips et al., 1973). In attempting to verify the latter, one might expect that more links would be searched than in the former (e.g., Collins & Quillian, 1969), but as shown above, the derived prediction is incorrect. It is interesting to note that the relation between latency and accuracy as a function of "semantic relatedness" is similar to the relation between latency and accuracy as a function of relatedness in the retrieval theory developed here, namely, extreme relatedness leads to fast reaction times and high accuracy.

Collins and Loftus (1975) have proposed a comparison process for semantic processing, wherein evidence is accumulated until either a positive or negative criterion is reached, whereupon a response is initiated. This comparison process belongs to the class of random walk processes (as Collins and Loftus point out) that forms the basis of the retrieval theory developed here. Collins and Loftus (1975, Table 1) listed qualitatively different kinds of evidence that can contribute to a decision. I argued earlier that qualitatively different kinds of information contribute to relatedness in item recognition. Therefore, it seems that the two points of view converge, and the theory developed earlier at least qualitatively extends into the domain of semantic processing.

Propositional models. Propositional models (J. R. Anderson, 1976; Anderson & Bower, 1973; Kintsch, 1974; Rumelhart, Lindsay, & Norman, 1972; Schank, 1976; see also Cofer, 1976) are close cousins of semantic network models but have slightly different aims in that the central task is to model sentence or text representation. The ancestry of these models is in linguistic theory, automata theory, and the theory of formal languages (Chomsky, 1963). Therefore, it is not surprising that a major goal of these theories has been sufficiency, that is, the ability of the theory (at least in principle) to deal with and represent the complexities of human language and knowledge. Thus, there has been less attention paid to processes underlying comprehension and retrieval, although the work of J. R. Anderson (1976), Anderson and Bower (1973), and Kintsch (1974, 1975) are important exceptions. Even so, Anderson and Bower (1973, p. 6), in their formulation, selected processing assumptions on the basis of their ease of implementation on a digital computer. In particular, serial search of network structures was assumed for matching or comparison processes.

J. R. Anderson (1974) developed two mathematical models for retrieval of propositional information. One model assumed independence of processing stages (the usual search model), whereas the other assumed "complete dependence," in which search time was exponentially distributed with the time constant equal to the sum of time constants for individual link search times. Anderson claimed that the two models gave similar predictions for mean latency, but it is easy to see that they give opposite predictions for the behavior of skewness of the latency distributions. As the number of links increases, for complete independence, skewness decreases; whereas for complete dependence, skewness increases. In general, in word recognition experiments, as mean latency increases, skewness of the distribution increases. Therefore, the complete-dependence assumption seems most consistent with data. The assumption of complete dependence certainly does not represent serial processing. Rather, it is more consonant with the view that the greater the number of links, the less evidence there is for the required relation holding (for in a network theory, eventually a pathway can be found between any pair of nodes). Thus, the assumption of complete dependence in processing seems halfway toward adoption of a decision process similar to the random walk class discussed earlier.

J. R. Anderson (1976) has gone a stage beyond earlier models in that the production system is the first attempt to model control processes in a model that makes close contact with experimental data. However, the processing assumptions underlying comparison processes still reflect the influence of the computer metaphor, and predictions for latency and accuracy come from separate sources (e.g., fan size and deadline, respectively) and are not closely interrelated. This is an important problem because most of the data the model makes contact with are reaction time data, and it can be seen that accuracy and latency covary in many of the data presented (J. R. Anderson, 1976).

This discussion suggests that the retrieval theory presented in this article may be capable of representing the retrieval and decision processes involved in processing semantic and more highly structured information (propositional and text). Thus, a task for the future is the integration of the retrieval theory and theories of semantic and text memory structure.

General Conclusions

I have presented a theory of memory retrieval that not only applies over a range of paradigms but also deals with experimental data in greater depth and more detail than competing models. The theory provides a rationale for relating accuracy, mean reaction time, error latency, and reaction time distributions. Such relations have not been dealt with explicitly before. Because the theory applies over a range of experimental paradigms, it provides a basis for the claim that the same processes are involved in each of these paradigms, allowing theoretical comparisons to be made between paradigms, and thus sets the stage for experimental comparison of paradigms. Also, the theory makes contact with several other more general theories, namely, neural network, semantic memory, and propositional models. The success of the theory so far, and the potential shown for integration with other general models, leads to the hope that some measure of theoretical unification may soon be achieved.

Reference Notes

1. Sternberg, S. *Estimating the distribution of additive reaction-time components*. Paper presented at the meeting of the Psychometric Society, Niagara Falls, Ontario, Canada, October 1964.
2. Sternberg, S. *Evidence against self-terminating memory search from properties of RT distributions*. Paper presented at the meeting of the Psychonomic Society, St. Louis, November 1973.
3. Ratcliff, R. *Group reaction time distributions and an analysis of distribution statistics*. Paper presented at the Mathematical Psychology meeting, Los Angeles, August 1977.
4. Reed, A. V. *Discriminability spectra in bilingual recognition*. Paper presented at the Mathematical Psychology meeting, Los Angeles, August 1977.

References

Allport, D. A. Critical notice: The state of cognitive psychology. *Quarterly Journal of Experimental Psychology*, 1975, *27*, 141–152.

Anderson, J. A. A simple neural network generating an interactive memory. *Mathematical Biosciences*, 1972, *14*, 197–220.

Anderson, J. A. A theory for the recognition of items from short memorized lists. *Psychological Review*, 1973, *80*, 417–438.

Anderson, J. A. Neural models with cognitive implications. In D. La Berge & S. J. Samuels (Eds.), *Basic processes in reading: Perception and comprehension*. Hillsdale, N.J.: Erlbaum, 1976.

Anderson, J. A., Silverstein, J. W., Ritz, S. A., & Jones, R. S. Distinctive features, categorical perception, and probability learning: Some applications of a neural model. *Psychological Review*, 1977, *84*, 413–451.

Anderson, J. R. Retrieval of propositional information from long-term memory. *Cognitive Psychology*, 1974, *6*, 451–474.

Anderson, J. R. *Language, memory and thought*. Hillsdale, N.J.: Erlbaum, 1976.

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

Anderson, J. R., & Bower, G. H. *Human associative memory*. Washington, D.C.: Winston, 1973.

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*. Hillsdale, N.J.: Erlbaum, 1974.

Atkinson, R. C., Holmgren, J. E., & Juola, J. F. Processing time is influenced by the number of elements in a visual display. *Perception & Psychophysics*, 1969, *6*, 321–327.

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

Atkinson, R. C., & Shiffrin, R. M. Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2). New York: Academic Press, 1968.

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. P., & Atkinson, R. C. Accuracy and speed strategies in scanning active memory. *Memory & Cognition*, 1974, *2*, 629–636.

Bobrow, D. G. Dimensions of representation. In D. G. Bobrow & A. M. Collins (Eds.), *Representation and understanding: Studies in cognitive science*. New York: Academic Press, 1975.

Borsellino, A., & Poggio, T. Convolution and correlation algebras. *Kybernetik*, 1973, *13*, 113–122.

Briggs, G. E. On the predictor variable for choice reaction time. *Memory & Cognition*, 1974, *2*, 575–580.

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

Burrows, D., & Okada, R. Memory retrieval from long and short lists. *Science*, 1975, *188*, 1031–1033.

Cavanagh, J. P. Holographic and trace strength models of rehearsal effects in the item recognition task. *Memory & Cognition*, 1976, *4*, 186–199.

Chomsky, N. Formal properties of grammars. In R. D. Luce, R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 2). New York: Wiley, 1963.

Churchill, R. V. *Fourier series and boundary value problems* (2nd. ed.). New York: McGraw-Hill, 1963.

Cofer, C. N. *The structure of human memory*. San Francisco: Freeman, 1976.

Collins, A. M., & Loftus, E. F. A spreading-activation theory of semantic processing. *Psychological Review*, 1975, *82*, 407–428.

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

Collins, A. M., & Quillian, M. R. Does category size affect categorization time? *Journal of Verbal Learning and Verbal Behavior*, 1970, *9*, 432–438.

Corballis, M. C. Serial order in recognition and recall. *Journal of Experimental Psychology*, 1967, *74*, 99–105.

Corballis, M. C. Access to memory: An analysis of recognition times. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance V*. New York: Academic Press, 1975.

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

Cox, D. R., & Miller, H. D. *The theory of stochastic processes*. London: Methuen, 1965.

Dosher, B. A. The retrieval of sentences from memory: A speed–accuracy study. *Cognitive Psychology*, 1976, *8*, 291–310.

Feller, W. *An introduction to probability theory and its applications* (Vol. 1, 3rd. ed.). New York: Wiley, 1968.

Forrin, B., & Cunningham, K. Recognition time and serial position of probed item in short-term memory. *Journal of Experimental Psychology*, 1973, *99*, 272–279.

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

Hollan, J. D. Features and semantic memory: Set-theoretic or network model? *Psychological Review*, 1975, *82*, 154–155.

Huesmann, L. R., & Woocher, F. D. Probe similarity and recognition of set membership: A parallel-processing serial-feature-matching model. *Cognitive Psychology*, 1976, *8*, 124–162.

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

Kintsch, W. *The representation of meaning in memory*. Hillsdale, N.J.: Erlbaum, 1974.

Kintsch, W. Memory representations of text. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium*. Hillsdale, N.J.: Erlbaum, 1975.

Kohonen, T. *Associative memory: A system–theoretical approach*. Berlin, West Germany: Springer-Verlag, 1976.

Kohonen, T., & Oja, E. Fast adaptive formation of orthogonalizing filters and associative memory in recurrent networks of neuron-like elements. *Biological Cybernetics*, 1973, *21*, 85–95.

Krantz, D. H. Threshold theories of signal detection. *Psychological Review*, 1969, *76*, 308–324.

Laming, D. R. *Information theory of choice reaction time*. New York: Wiley, 1968.

Link, S. W. The relative judgment theory of choice reaction time. *Journal of Mathematical Psychology*, 1975, *12*, 114–135.

Link, S. W., & Heath, R. A. A sequential theory of psychological discrimination. *Psychometrika*, 1975, *40*, 77–105.

Lively, B. L. Speed/accuracy trade off and practice as determinants of stage durations in a memory-search task. *Journal of Experimental Psychology*, 1972, *96*, 97–103.

Luce, R. D. A threshold theory for simple detection experiments. *Psychological Review*, 1963, *70*, 61–79.

Marcel, A. J. Negative set effects in character classification: A response-retrieval view of reaction time. *Quarterly Journal of Experimental Psychology*, 1977, *29*, 31–48.

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

Miller, J. O., & Pachella, R. G. Locus of the stimulus probability effect. *Journal of Experimental Psychology*, 1973, *101*, 227–231.

Morin, R. E., DeRosa, D. V., & Stultz, V. Recognition memory and reaction time. In A. F. Sanders (Ed.), *Attention and performance I*. Amsterdam: North-Holland, 1967.

Morton, J., Marcus, S., & Frankish, C. Perceptual centers (P-centers). *Psychological Review*, 1976, *83*, 405–408.

Murdock, B. B., Jr. Response latencies in short-term memory. *Quarterly Journal of Experimental Psychology*, 1968, *20*, 79–82.

Murdock, B. B., Jr. *Human memory: Theory and data*. Potomac, Md.: Erlbaum, 1974.

Murdock, B. B., Jr., & Anderson, R. E. Encoding, storage and retrieval of item information. In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola Symposium*. Hillsdale, N.J.: Erlbaum, 1975.

Murdock, B. B., Jr., & Dufty, P. O. Strength theory and recognition memory. *Journal of Experimental Psychology*, 1972, *94*, 284–290.

Murdock, B. B., Jr., Hockley, W. E., & Muter, P. M. Two tests of the conveyor belt model for item recognition. *Canadian Journal of Psychology*, 1977, *31*, 71–89.

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

Newell, A. You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press, 1973.

Nickerson, R. S. Binary-classification reaction time: A review of some studies of human information-processing capabilities. *Psychonomic Monograph Supplements*, 1972, *4*(17, Whole No. 65).

Norman, D. A., & Rumelhart, D. E. A system for perception and memory. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970.

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

Pachella, R. G. An interpretation of reaction time in information processing research. In B. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition*. Hillsdale, N.J.: Erlbaum, 1974.

Pike, R. Response latency models for signal detection. *Psychological Review*, 1973, *80*, 53–68.

Ratcliff, R., & Murdock, B. B., Jr. Retrieval processes in recognition memory. *Psychological Review*, 1976, *83*, 190–214.

Reed, A. V. Speed–accuracy trade-off in recognition memory. *Science*, 1973, *181*, 574–576.

Reed, A. V. List length and the time course of recognition in immediate memory. *Memory & Cognition*, 1976, *4*, 16–30.

Rips, L. J., Shoben, E. J., & Smith, E. E. Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, 1973, *12*, 1–20.

Rips, L. J., Shoben, E. J., & Smith, E. E. Set-theoretic and network models reconsidered: A comment on

Hollan's "Features and semantic memory." *Psychological Review*, 1975, *82*, 156–157.
Rumelhart, D. E., Lindsay, P. H., & Norman, D. A. A process model for long-term memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972.
Schank, R. C. The role of memory in language processing. In C. Cofer (Ed.), *The structure of human memory*. San Francisco: Freeman, 1976.
Scheirer, C. J., & Hanley, M. J. Scanning for similar and different material in short- and long-term memory. *Journal of Experimental Psychology*, 1974, *102*, 343–346.
Schneider, W., & Shiffrin, R. M. Controlled and automatic human information processing. *Psychological Review*, 1977, *86*, 1–66.
Schouten, J. F., & Bekker, J. A. M. Reaction time and accuracy. *Acta Psychologica*, 1967, *27*, 143–153.
Schulman, H. G. Encoding and retention of semantic and phonemic information in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1970, *9*, 499–508.
Seward, G. H. Recognition time as a measure of confidence. *Archives of Psychology*, 1928, *99*, 723. (Cited in Woodworth, R. S. *Experimental psychology*. New York: Holt, 1938.)
Smith, E. E., Shoben, E. J., & Rips, L. J. Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, 1974, *81*, 214–241.
Sternberg, S. High-speed scanning in human memory. *Science*, 1966, *153*, 652–654.
Sternberg, S. The discovery of processing stages: Extensions of Donders' method. In W. G. Koster (Ed.), *Attention and performance II*. Amsterdam: North-Holland, 1969. (a)
Sternberg, S. Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist*, 1969, *57*, 421–457. (b)
Sternberg, S. Memory scanning: New findings and current controversies. *Quarterly Journal of Experimental Psychology*, 1975, *27*, 1–32.
Stone, M. Models for choice reaction time. *Psychometrika*, 1960, *25*, 251–260.
Theios, J., Smith, P. G., Haviland, S. E., Traupmann, J., & Moy, M. C. Memory scanning as a serial self-terminating process. *Journal of Experimental Psychology*, 1973, *97*, 323–336.
Theios, J., & Walter, D. G. Stimulus and response frequency and sequential effects in memory scanning reaction times. *Journal of Experimental Psychology*, 1974, *102*, 1092–1099.
Tikhonov, A., & Samarskii, A. A. *Equations of mathematical physics*. New York: Macmillan, 1963.
Townsend, J. T. A note on the identifiability of parallel and serial processes. *Perception & Psychophysics*, 1971, *10*, 161–163.
Townsend, J. T. Some results concerning the "identifiability" of parallel and serial processes. *British Journal of Mathematical and Statistical Psychology*, 1972, *25*, 168–199.
Townsend, J. T. Issues and models concerning the processing of a finite number of inputs. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition*. Hillsdale, N.J.: Erlbaum, 1974.
Tulving, E., & Bower, G. H. The logic of memory representations. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8). New York: Academic Press, 1974.
Wells, J. E. Strength theory and judgments of recency and frequency. *Journal of Verbal Learning and Verbal Behavior*, 1974, *13*, 378–392.
Wescourt, K. T., & Atkinson, R. C. Scanning for information in short-term and long-term memory. *Journal of Experimental Psychology*, 1973, *98*, 95–101.
Wescourt, K. T., & Atkinson, R. C. Fact retrieval processes in human memory. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 4). Hillsdale, N.J.: Erlbaum, 1976.
Wickelgren, W. A. Dynamics of retrieval. In D. Deutsch & J. A. Deutsch (Eds.), *Short-term memory*. New York: Academic Press, 1975.
Wickelgren, W. A. Speed–accuracy tradeoff and information processing dynamics. *Acta Psychologica*, 1977, *41*, 67–85.
Wickelgren, W. A., & Corbett, A. T. Associative interference and retrieval dynamics in yes–no recall and recognition. *Journal of Experimental Psychology: Human Learning and Memory*, 1977, *3*, 189–202.
Wickelgren, W. A., & Norman, D. A. Strength models and serial position in short-term recognition memory. *Journal of Mathematical Psychology*, 1966, *3*, 316–347.

Appendix

In this section, the mathematical model used to relate the theory to data is formulated. The organization is as follows: First, the theory of the random walk (gambler's ruin problem) is described as an introduction to the diffusion process. Second, the theory of the diffusion process is outlined, and two ways of deriving first-passage time distributions and error rates are indicated. Third, equations for the maximum of n processes (exhaustive processing) and the minimum of n processes (self-terminating processing) are derived.

In the development of any model, one has to select aspects of the data, that is, summary statistics, as points of contact between the model and data. In item recognition, the statistic generally used is mean reaction time, but as argued earlier, mean error rate and reaction time distribution statistics are more useful. Thus, the following development is aimed primarily at deriving expressions for distributions and error rates.

Random Walk

Traditionally, the theory of this process is centered around the example of a gambler who wins or loses a dollar with probabilities p and q, respectively. His initial capital is Z and his opponent's capital is $A - Z$. The game terminates when his capital becomes A or 0 dollars (Feller, 1968, p. 342).

It is simple to rephrase the problem in terms of a feature-matching, memory comparison process (cf. Huesmann & Woocher, 1976). Suppose the features of a probe item are compared with the features of a memory item, and the probability of a match is p and of a nonmatch is q $(= 1 - p)$. Then, if Z more nonmatches than matches are required for a probe–memory item nonmatch, and $A - Z$ more matches than nonmatches are required for a probe–memory item match, this process is isomorphic to the gambler's ruin problem. I shall discuss the random walk within the feature match framework.

One of the main methods for finding solutions for stochastic processes consists of deriving difference equations for the process and solving these equations with respect to certain boundary conditions. In our feature comparison process, let q_Z be the probability of a probe–item nonmatch, given starting point Z, match at A, and nonmatch at 0. Then, after the first trial, either $Z - 1$ or $Z + 1$ steps are required for termination with a nonmatch. Thus,

$$q_Z = pq_{Z+1} + qq_{Z-1}, \tag{A1}$$

provided $1 < Z < A - 1$. For $Z = 1$ and $Z = A - 1$, the first trial may lead to termination, so we define

$$q_0 = 1 \quad \text{and} \quad q_A = 0 \tag{A2}$$

to be boundary conditions on the random walk. Thus, the probability of nonmatch q_Z satisfies Equations A1 and A2.

Using the method of particular solutions, the solution to Equations A1 and A2 for the probability of a nonmatch is given by

$$q_Z = \begin{cases} \dfrac{\left(\dfrac{q}{p}\right)^A - \left(\dfrac{q}{p}\right)^Z}{\left(\dfrac{q}{p}\right)^A - 1}, & \text{when } q \neq p \\ 1 - \dfrac{Z}{A}, & \text{when } q = p. \end{cases} \tag{A3}$$

This is the required expression for error rate. The next result needed is the first-passage time distribution for a nonmatch (the distribution of number of steps to nonmatch).

Let $g_{Z,n}$ denote the probability that the process ends with the nth step at the barrier 0 (nonmatch at the nth feature comparison). After the first step, the position is $Z + 1$ or $Z - 1$ for $1 < Z < A - 1$ and $n \geqslant 1$; thus,

$$g_{Z,n+1} = pg_{Z+1,n} + qg_{Z-1,n}. \tag{A4}$$

Boundary conditions for this difference equation are given by

$$g_{0,n} = g_{A,n} = 0$$

and (A5)

$$g_{0,0} = 1,\ g_{Z,0} = 0, \quad \text{when } Z > 0.$$

Then, the difference Equation A4, with boundaries given by Equation A5, holds for all $0 < Z < A$ and all $n \geqslant 0$.

The solution to Equations A4 and A5 is given by

$$g_{Z,n} = \frac{2^{n+1}}{A} p^{(n-Z)/2} q^{(n+Z)/2} \times \sum_{k<A/2} \cos^{n-1}\frac{\pi k}{A} \sin\frac{\pi k}{A} \sin\frac{\pi Z k}{A}. \tag{A6}$$

The derivation of Equation A6 is shown in Feller (1968, chap. 14) and involves deriving the probability generating function for first-passage times, then performing a partial fraction expansion on that expression.

In order to obtain the expressions equivalent to Equations A3 and A6 for a probe–memory item match, q and p are interchanged and Z becomes $A - Z$. It should also be noted that $g_{Z,n}$ is not a probability density function, but that the sum of the first-passage time probabilities for matches and nonmatches is a probability density function.

Diffusion Process

It was noted earlier that the discrete random walk could be used to represent a feature comparison process. However, I have chosen to use the continuous random walk (Wiener diffusion) process for several reasons. First, my conceptual bias is to think of information or evidence accumulation as a continuous process rather than a discrete process. Second, for the continuous random walk, drift and variance in drift are independent, and because these and other parameters of the model are relatively decoupled, manipulation is simpler. Third, it is faster and cheaper to integrate numerically over continuous functions than to sum over discrete functions.

Before proceeding, I will introduce the following convention: F and G will represent first-passage time distribution functions, and f and g will represent first-passage time density functions. These are not normalized (i.e., are not probability distributions or density functions), but they may be normalized by division by the appropriate error or correct probability γ. On the other hand, H and h will represent proper probability distribution and density functions. Subscripts $+$ and $-$ will refer to the boundaries giving match and nonmatch, respectively.

It was noted earlier that the Wiener diffusion process (Brownian motion) is the result of taking the limit as step size tends toward zero in the random walk process. This limiting process will now be described.

In order to apply the random walk in the case where the number of events per unit time is large, but the effect of each event is small (e.g., molecular collisions in Brownian motion), one can take the discrete expressions and find the limit directly. Let the length of the individual steps δ be small, the number of steps per unit time r be large, and $(p - q)$ be small, with the constraint that the following expressions approach finite limits:

$$(p - q)r\delta \rightarrow \xi, \quad 4pqr\delta^2 \rightarrow s^2. \tag{A7}$$

It is necessary for the two expressions to have finite limits. For example, if $s^2 \rightarrow 0$, then the process is deterministic; if $s^2 \rightarrow \infty$, then the process is infinitely variable, and the probability of terminating at each boundary is 1/2 with zero time delay. If $\xi \rightarrow 0$, there is no drift; if $\xi \rightarrow \pm\infty$, there is no time delay before the boundary is reached. The probability of a nonmatch is obtained by taking the limit in Equation A3 and is given by

$$\gamma_-(\xi) = \frac{e^{-(2\xi a/s^2)} - e^{-(2\xi z/s^2)}}{e^{-(2\xi a/s^2)} - 1}, \tag{A8}$$

where $Z \rightarrow z/\delta$ and $A \rightarrow a/\delta$. $\gamma_+(\xi)$, the probability of a match, can be obtained by replacing z with $a - z$ and ξ with $-\xi$. Then, $\gamma_+ + \gamma_- = 1$.

The limiting form of the first-passage time distribution is given by

$$g_-(t, \xi) = \frac{\pi s^2}{a^2} e^{-(z\xi/s^2)} \times \sum_{k=1}^{\infty} k \sin\left(\frac{\pi z k}{a}\right) e^{-\frac{1}{2}(\xi^2/s^2 + \pi^2 k^2 s^2/a^2)t}. \tag{A9}$$

$g_-(t, \xi)$ is a function of t, a, z, ξ, and s^2, but it is only necessary to express it as a function of t and ξ for further calculations in this section. The first-passage time distribution for a match $g_+(t, \xi)$ can be found by setting $\xi = -\xi$ and $z = a - z$ in Equation A9. Again, note that $g_-(t, \xi)$ and $g_+(t, \xi)$ are not probability density functions, but $g_+(t, \xi) + g_-(t, \xi)$, $g_-(t, \xi)/\gamma_-(\xi)$, and $g_+(t, \xi)/\gamma_+(\xi)$ are probability density functions.

It was stated earlier that Equation A9, the first-passage time distribution, can be derived by solving the diffusion equation. The diffusion equation can be derived from Equation A4, using the limiting expressions in Equation A7 for ξ and s^2. Equation A4 becomes

$$\frac{\partial g_-(t, \xi)}{\partial t} = \xi \frac{\partial g_-(t, \xi)}{\partial z} + \frac{s^2}{2} \frac{\partial^2 g_-(t, \xi)}{\partial z^2}. \tag{A10}$$

Methods for the solution of this equation (Cox & Miller, 1965) lead into the enormous literature on the solution of partial differential equations (Churchill, 1963; Tikhonov & Samarskii, 1963). By substituting $g_-(t, \xi)$ from Equation A9 in Equation A10, it can be seen that $g_-(t, \xi)$ is a solution of Equation A10.

The first-passage time density function $g_-(t, \xi)$ turns out not to be the best expression to use in evaluating the maximum and minimum of a number of similar processes. Instead, the first-passage time distribution function

$$G_-(t, \xi) = \int_0^t g_-(t', \xi) dt' \tag{A11}$$

proves most useful. Now,

$$G_-(t, \xi) = \frac{\pi s^2}{a^2} e^{-(z\xi/s^2)} \sum_{k=1}^{\infty} \frac{2k \sin\left(\frac{k\pi z}{a}\right)\left[1 - e^{-\frac{1}{2}(\xi^2/s^2 + \pi^2 k^2 s^2/a^2)t}\right]}{\left(\frac{\xi^2}{s^2} + \frac{\pi^2 k^2 s^2}{a^2}\right)};$$

however, this series converges slowly, and a more useful alternative is given by

$$G_-(t, \xi) = \gamma_-(\xi) - \frac{\pi s^2}{a^2} e^{-(z\xi/s^2)} \times \sum_{k=1}^{\infty} \frac{2k \sin\left(\frac{k\pi z}{a}\right) e^{-\frac{1}{2}(\xi^2/s^2 + \pi^2 k^2 s^2/a^2)t}}{\left(\frac{\xi^2}{s^2} + \frac{\pi^2 k^2 s^2}{a^2}\right)}. \quad \text{(A12)}$$

From this expression, it can be seen that

$$\lim_{t \to \infty} G_-(t, \xi) = \gamma_-(\xi).$$

Relatedness and Diffusion

The main assumption made to relate diffusion and relatedness is that relatedness is represented by drift ξ in the diffusion process. Thus, the greater the relatedness, the greater is the value of drift; the less the relatedness, the less the drift. The scale on which relatedness is measured is not absolute; rather, it is only sensible to talk about the degree of discriminability. The subject sets the zero point (or criterion) of relatedness, so that the average drift for a matching comparison is toward the match boundary, and average drift for a nonmatching comparison is toward the nonmatch boundary. The scale of relatedness is linear in drift, and therefore, the scale is defined in terms of the comparison process. There are many other possible scales and relations between relatedness and drift, but this one seems the most natural. I will use ξ to represent relatedness of probe–memory-set item match comparisons and ν to represent relatedness of probe–memory-set item nonmatch comparisons, where ν usually has the opposite sign to ξ.

Decision Process

The decision process is self-terminating on comparison process matches and exhaustive on nonmatches. The distributions of "yes" and "no" responses therefore correspond to the distributions of minimum and maximum completion times of the n comparison processes.

Maximum of n *processes*. Let $T_1, \ldots, T_n$ be independent observations on n processes, with probability density functions $h_1(t), \ldots, h_n(t)$ and distribution functions $H_1(t), \ldots, H_n(t)$. Let us assume each process terminates with probability one. Then, the probability distribution function of $T = \max_{1 \leq i \leq n} T_i$ is

$$\begin{aligned} H_{\max}(t) &= P_r(T \leq t) \\ &= P_r(\max_{1 \leq i \leq n} T_i \leq t) \\ &= P_r(T_1 \leq t, \ldots, T_n \leq t) \\ &= P_r(T_1 \leq t) \ldots P_r(T_n \leq t) \\ &\quad \text{(assuming independence)} \\ &= \prod_{i=1}^{n} H_i(t). \end{aligned} \quad \text{(A13)}$$

Then, the probability density function of the maximum of n processes is given by

$$\begin{aligned} h_{\max}(t) &= H'_{\max}(t) \\ &= \prod_{i=1}^{n} H_i(t) \sum_{j=1}^{n} \frac{h_j(t)}{H_j(t)}. \end{aligned} \quad \text{(A14)}$$

For misses and correct rejections, we are not dealing now with probability distribution functions because sometimes a process may terminate at the wrong boundary. Therefore, it is necessary to find the maximum of n processes, given that there are different probabilities of termination associated with each process. Let $G_{i-}(t)$ be the first-passage time distribution function for a nonmatch of process i, and $G_{\max}(t)$ be the first-passage time distribution function for the maximum of the n nonmatch processes. Then,

$$\begin{aligned} G_{\max}(t) &= P_r(T \leq t) \\ &= \prod_{i=1}^{n} G_{i-}(t) \end{aligned} \quad \text{(A15)}$$

from Equation A13. Also,

$$\gamma_- = \prod_{i=1}^{n} \gamma_{i-}, \quad \text{(A16)}$$

where γ_{i-} is the probability of the ith process nonmatch, and γ_- is the probability that all of the comparison processes produce nonmatches.

For correct rejections, the model will be applied, with each comparison assumed to have the same first-passage time distribution $G_-(t, \nu)$. Thus, for correct rejections (CR), the first-passage time distribution function $F_{CR}(t)$ is given by

$$F_{CR}(t) = G_-^{\,n}(t, \nu). \quad \text{(A17)}$$

Also, for the single nonmatching process when a memory-set item is probed, let the first-passage time distribution function be $G_-(t, \xi)$; thus, for misses (M),

$$F_M(t) = G_-(t, \xi) G_-^{\,n-1}(t, \nu). \quad \text{(A18)}$$

Minimum of n *processes*. For "yes" responses (hits and false alarms), only one comparison process has to terminate with a match; therefore, all combinations of the number of match terminations have to be considered. For example, if only one match occurs, then the time for that comparison is the time for the decision; if two matches occur, then the decision time is the minimum of those two comparison times. Therefore, it is necessary to sum over all possible combinations of processes finishing. First, let us consider the minimum of n processes finishing. Then, the distribution function $H_{\min}(t)$ of $T = \min_{1\leqslant i\leqslant n} T_i$ is found as follows:

$$P_r(T>t)=1-P_r(T\leqslant t)$$
$$=1-H_{\min}(t).$$

Now

$$P_r(T>t)=P_r(T_1>t,\ \ldots,\ T_n>t)$$
$$=P_r(T_1>t)\ldots P_r(T_n>t)$$
$$=[1-P_r(T_1\leqslant t)]\ldots[1-P_r(T_n\leqslant t)]$$
$$=[1-H_1(t)]\ldots[1-H_n(t)].$$

Therefore,

$$H_{\min}(t)=1-\prod_{i=1}^{n}[1-H_i(t)]. \quad \text{(A19)}$$

Let $F_{\min}(t)$ be the first-passage time distribution function (not a probability distribution function) for the self-terminating process, and let p_i be the probability of a match. Then,

$$P_r(T>t)=\sum_{j=1}^{n}\sum_{\substack{\text{all subsets}\\ \text{of size } j}}[P_r(T>t\,|\,i_1,\ldots,i_j)\times\prod_{i=i_1}^{i_j}p_i\prod_{k=k_1}^{k_{n-j}}(1-p_k)]$$
$$=\sum_{j=1}^{n}\sum_{\substack{\text{all subsets}\\ \text{of size } j}}\left\{\left[1-\prod_{i=i_1}^{i_j}\left(1-\frac{G_i}{p_i}\right)\right]\times\prod_{i=i_1}^{i_j}p_i\prod_{k=k_1}^{k_{n-j}}(1-p_k)\right\}, \quad \text{(A20)}$$

using Equation A19. As before, when the probe was not a memory-set item, all nonmatch comparison first-passage time distribution functions $G_-(t,\nu)$ are assumed equal. Then, for false alarms (FA),

$$F_{\mathrm{FA}}(t)=\frac{nG_+(t,\nu)}{\gamma_+(\nu)}\gamma_+(\nu)\gamma_-^{\,n-1}(\nu)+\frac{n(n-1)}{2}\left[\frac{2G_+(t,\nu)}{\gamma_+(\nu)}-\frac{G_+^{\,2}(t,\nu)}{\gamma_+^{\,2}(\nu)}\right]\times\gamma_+^{\,2}(\nu)\gamma_-^{\,n-2}(\nu)+\ldots, \quad \text{(A21)}$$

where $G_+(t,\nu)$ can be found by setting $z=a-z$ and $\nu=-\nu$ in Equation A12.

Let the first-passage time distribution of a probe being compared with a memory-set item (i.e., for a match) be $G_+(t,\xi)$. Then, for hits (H), the first-passage time distribution function is

$$F_{\mathrm{H}}(t)=\frac{G_+(t,\xi)}{\gamma_+(\xi)}\gamma_+(\xi)\gamma_-^{\,n-1}(\nu)+(n-1)\frac{G_+(t,\nu)}{\gamma_+(\nu)}\gamma_+(\nu)\gamma_-^{\,n-2}(\nu)\gamma_-(\xi)$$
$$+(n-1)\left\{1-\left[1-\frac{G_+(t,\xi)}{\gamma_+(\xi)}\right]\left[1-\frac{G_+(t,\nu)}{\gamma_+(\nu)}\right]\right\}\times\gamma_+(\xi)\gamma_+(\nu)\gamma_-^{\,n-2}(\nu)$$
$$+\frac{(n-1)(n-2)}{2}\left\{1-\left[\frac{G_+(t,\nu)}{\gamma_+(\nu)}\right]^2\right\}\times\gamma_+^{\,2}(\nu)\gamma^{n-3}(\nu)\gamma_-(\xi)+\ldots. \quad \text{(A22)}$$

At this stage, we should note that both $\gamma_+(\nu)$ and $\gamma_-(\xi)$ are usually small (.1 or less) and $G_\pm(t)$ tends to $\gamma_\pm$ as t becomes large; therefore, Equations A21 and A22 are adequately approximated by the first few terms in the series.

The probabilities of a correct rejection and miss are given by

$$p_{\mathrm{CR}}=\gamma_-^{\,n}(\nu)=1-p_{\mathrm{FA}}$$

and

$$p_{\mathrm{M}}=\gamma_-^{\,n-1}(\nu)\gamma_-(\xi)=1-p_{\mathrm{H}}. \quad \text{(A23)}$$

The approximation in Equation A22 was checked by calculating $F_{\mathrm{H}}(\infty)$ and by making sure that the value is equal to p_{H}.

Relatedness Distributions

A critical assumption is that relatedness enters the diffusion process as the drift parameter ξ. In the last section, it was noted that there should be a distribution of relatedness and therefore a distribution over drift values (different from the variance of the diffusion process s^2). The simplest choice used with most precedence (Murdock, 1974, p. 28) is the normal distribution. This means that the distribution function $G(t)$ and accuracy γ for the diffusion process have to be averaged over the relatedness distribution before being used to calculate distributions and proportions of hit, miss, correct rejection, and false alarm responses.

Let relatedness ξ be distributed $N(u,\eta)$ and relatedness ν be distributed $N(v,\eta)$. Then, we can compute first-passage time distribution functions $G_{\pm}(t, u)$ and $G_{\pm}(t, v)$, using

$$G_{\pm}(t, u) = \int_{-\infty}^{\infty} G_{\pm}(t, \xi)\frac{1}{\sqrt{2\pi\eta^2}}e^{-\frac{1}{2}(u-\xi)^2/\eta^2}d\xi \quad (A24)$$

and

$$G_{\pm}(t, v) = \int_{-\infty}^{\infty} G_{\pm}(t, \nu)\frac{1}{\sqrt{2\pi\eta^2}}e^{-\frac{1}{2}(v-\nu)^2/\eta^2}d\nu. \quad (A25)$$

Similarly, $\gamma_{\pm}(u)$ and $\gamma_{\pm}(v)$ can be calculated using Equations A24 and A25, respectively, substituting $\gamma_{\pm}(\xi)$ and $\gamma_{\pm}(\nu)$ for $G_{\pm}(t, \xi)$ and $G_{\pm}(t, \nu)$, respectively.

The equations developed to this point are sufficient for the calculation of error rates and reaction time distributions.

Received December 17, 1976
Revision received July 12, 1977 ■

Erratum for Librarians and Secondary Services

For all six issues of Volume 84 (1977) and the first issue of Volume 85 (January 1978) of *Psychological Review,* the ISSN Number (8098-7970) on the front cover is incorrect. The correct ISSN Number is 0033-295X.