Psychological Review
1981, Vol. 88, No. 6, 552-572

0033-295X/81/8806-0552$00.75

A Theory of Order Relations in Perceptual Matching

Roger Ratcliff
Yale University

A theory of order relations in the perceptual matching task relates order manipulations to research on retrieval processes and the representation of order information in memory. In experimental tests of the theory, presentation of a study string of letters to the subject was followed by a test string to which the subject responded *same* or *different*. The data of main interest concern the case where the test string is a permutation of the study string. When adjacent letters are switched, reaction time is long and accuracy low, suggesting that, in the comparison process, a test letter is not simply compared to the letter in the same position in the study string; rather, the comparison is distributed across positions. The memory model assumes that the representation of a letter is distributed (spread) over position and that the comparison process assesses the amount of overlap between the test string and the memory representation. The amount of overlap is transformed by a power function into the drift rate in a diffusion (random walk) comparison process. The diffusion retrieval model and overlap memory model are fitted to the data and goodness-of-fit is assessed. Shortcomings of alternative models are considered and applications of the model to related matching tasks are described.

One of the main problems with models in cognitive psychology at the present time is that they tend to be developed in isolation. For example, a sophisticated model of memory structure may have associated with it only the simplest retrieval mechanism, or alternatively, a detailed model of retrieval processes may have little to say about memory structure. The aim of this paper is to attack this problem by integrating two classes of models: a perturbation or overlap model of memory for order information and a random-walk model of retrieval processes. These models will be applied to the area of research known as perceptual matching. This paper will first review the perceptual matching literature, emphasizing retrieval models, and then the order information literature. A new model for the representation of order information and retrieval processes in the perceptual matching task will then be presented.

This research was supported by Grant HD 13318 from the National Institute of Child Health and Human Development. I would like to thank Gail McKoon for her help throughout this research. I would also like to thank Mike Hacker, Bill Estes, and Cathy Lee for their very useful comments at all stages of the research and Lester Krueger and Gary Dell for useful comments on the manuscript.

Requests for reprints should be sent to Roger Ratcliff, Department of Psychology, Yale University, Box 11A Yale Station, New Haven, Connecticut 06520.

Perceptual Matching

The perceptual matching task requires the matching of one pattern against another pattern. It has been argued many times that this process is a fundamental component of human information processing. In the world around us, we are continually matching patterns, whether it be faces, common objects, traffic situations, or whatever. With typical panache experimental psychologists have translated this rich matching process into the tightly experimentally controlled task of matching one letter string against another letter string.

One of the major findings of this area of research is that judgments are made faster when two letter strings are the same than when they are different. Several models have been developed to explain this result and these have been reviewed in a number of places (e.g., Krueger, 1978; Nickerson, 1978). It will be sufficient for this paper to discuss only the major current models.

Bamber (1969) developed a model with

a serial-matching process, that is, a process by which the letters from the study and test strings are compared one by one. This model gave good fits to mean reaction time for *different* judgments. For *same* judgments the model predicted reaction times longer than were found in the data and so a fast identity-matching mechanism was added to the model. The model did not make predictions about error rates; Bamber argued that errors were probably due to loss of information from the sensory information store and were therefore outside the scope of the model.

Taylor (1976) compared serial and parallel search models for perceptual matching and concluded that a parallel model with a limited-capacity parallel comparison process is superior to any serial model. The parallel model includes an identity matcher as in Bamber's model and incorporates a guessing process to account for errors.

There are several problems with models that, like those proposed by Bamber and Taylor, include scanning plus identity-match comparison processes. First, accuracy and reaction time are not tied together in the processing mechanism; rather, error rate and reaction time differences result from different mechanisms. Thus, these models are unable to deal with speed–accuracy trade-off relations, except by the manipulation of unrelated mechanisms (scanning and guessing). Second, it is not possible according to these models for *same* judgments to be slower than *different* judgments, even though a relatively simple manipulation (Experiment 1 below) removes the *same* advantage. The third problem is that there is no attempt to account for the shapes of the distributions of reaction times.

Krueger (1978) has developed a model that deals with many of these problems. According to this model a letter is composed of a number of features (Krueger assumed 100 features in fitting the data). It is assumed that if a study letter matches a test letter, there are no nonmatching features; if the study and test letters do not match, then there is a small number of nonmatching features (e.g., 6 to 12 out of 100). The comparison process is executed in a series of passes. On each pass the number of feature nonmatches out of the total number of features is counted. If this number is above one criterion, a *different* response is initiated; if this number is below another criterion, a *same* response is initiated. If the number of nonmatches is between the two criteria, another pass is performed and the number of nonmatches again checked against two (adjusted) criteria. Passes continue to be executed until a criterion is met or a deadline is reached. During the comparison process a feature match may be misperceived as a feature nonmatch or a feature nonmatch may be misperceived as a feature match. The probability of such misperceptions is low (e.g., less than .1). Such misperceptions represent noise in the comparison process.

For multielement strings it could be assumed that there are just more features to be compared. However, this assumption does not provide good fits of the model to data. Instead, it is assumed that the probability of obtaining a feature mismatch in the comparison process decreases as the number of mismatching letters increases. This assumption will be discussed later when Krueger's model is compared to the model presented in this paper.

Krueger's model is of the class of sequential-sampling models and as such relates accuracy and reaction time to each other (and provides an account of the shape of reaction time distributions). Models of this class have been used in several different areas of research to relate accuracy and reaction time, including recognition memory (Ratcliff, 1978), choice reaction time (Laming, 1968; Link, 1975; Stone, 1960), and semantic memory (McCloskey & Glucksberg, 1979; Collins & Loftus, 1975).

Order Information

Models of the representation of order information in short-term (or primary) memory have been presented by Lee and Estes (1977, 1981) and by Shiffrin and Cook (1978).

Lee and Estes's model is based on the perturbation model of Estes (1972) and is designed to explain data from recall experiments. In experiments presented by Lee and Estes (1977), subjects studied four letters interspersed among eight digits. The subjects were asked to recall the letters in correct position. The data of main interest were the position gradients of recall for the studied letters (i.e., the frequency with which item i was recalled at position i, $i + 1$, $i - 1$, etc.). The perturbation model assumes that an item has some probability of perturbing to adjacent positions in the string. This model provided a good account of transposition data (i.e., items i and j recalled in incorrect order) when study letters were separated but overpredicted the number of transpositions when study letters were adjacent. It was proposed that adjacent study letters form a chunk (Estes, 1972), so that recall of adjacent letters could be mediated through the chunk, giving better order recall than would otherwise be predicted by the perturbation model.

Shiffrin and Cook (1978) have presented a model with a different flavor. Subjects are assumed to encode a study string of letters both with item-to-item links, which provide the basis for ordered recall, and item-to-context links, which provide information about an item's occurrence but not its position. They applied the model to data from several experiments using short-term recall of letter strings. The model provided good fits to the data, and they concluded that this model was a viable alternative to the perturbation model.

Perceptual Matching and Order

An Overview of the Model

The model proposed here provides both a submodel for the representation of information in memory and a submodel for the process of matching the study and test representations to produce a *same* or *different* response. The memory model assumes that the letters of the study string have distributions over position so that the representation of one letter will extend into adjacent letter positions; thus, the model is closely related to Lee and Estes's (1977) model. This assumption leads to interesting predictions about the relative difficulty of various negative conditions in the matching task. For example, if a test string contains the same letters as the study string but with two adjacent letters switched in position, then the test string will appear to match the study representation very well. This condition will be more difficult to respond *different* to than the condition in which two letters of the study string are replaced by new letters in the test string.

The retrieval model is concerned with the assessment of the amount of match or overlap between the study and test strings; the result of this assessment is a *same* or *different* judgment as to the identity of the two strings. It is assumed that the amount of overlap (or relatedness) between the two strings is not constant during the course of the comparison process but varies randomly about some mean value. This mean value determines the mean drift rate in the random-walk comparison process: The greater the relatedness, the more quickly evidence is accumulated toward a positive response; the smaller the relatedness, the more quickly evidence is accumulated toward a negative response. The particular version of the random-walk model used in this paper is the continuous diffusion model (Ratcliff, 1978). An overview of the whole scheme is shown in Figure 1.

Item and order information. The memory model represents both item and order information in the same way, that is, in the distribution of information that is spread over position. Loss of order information is represented by spreading of the study letters over positions so that positional certainty is lost. Loss of item information about a letter is represented by assuming that the study distribution of that letter has an area less than one (for further discussion of item and order information, see Murdock, 1976).

Matching process. The process of matching is assumed to proceed in real time in the following way: A particular study string is encoded and a representation is laid down in memory, the letters in the representation having distributions over position. Then the test string is encoded; because the comparison of the study and test strings begins almost immediately, the letters in the test string are not assumed to be distributed over position. The comparison process measures the amount of overlap between the study and test strings as that amount varies continuously over a period of time about some mean

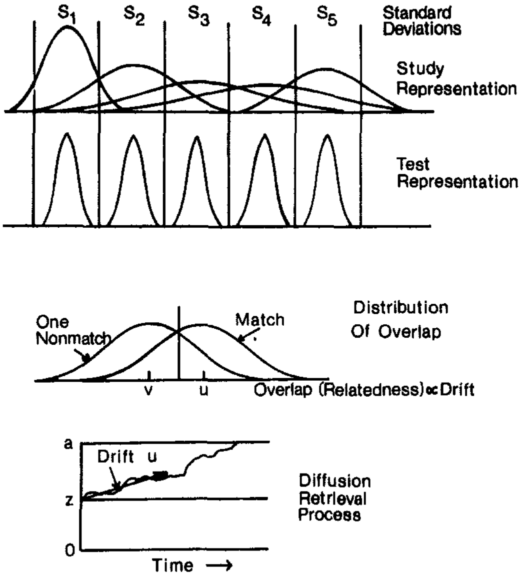

Figure 1. Overview of the model. (The top panel depicts the memory representation of the study and test strings. The middle panel shows the distributions of overlap for a matching and nonmatching comparison. The bottom panel depicts the diffusion comparison process. s_1 through s_5 are the standard deviations of the study distributions, u and v are mean relatedness values for matching and nonmatching comparisons respectively [u is also the mean drift rate denoted by the arrow in the bottom panel], and z and a are the starting point and match boundary positions in the diffusion process.)

value. The quantity representing the amount of overlap is transformed so that small differences between the amount of overlap of matching strings and the amount of overlap of nonmatching strings are magnified. This transformed value is the drift rate that drives the random-walk comparison process. This quantity is accumulated (integrated) over time so that the larger the average drift, the faster the process proceeds to a match, and the smaller the average drift, the faster the process proceeds to a nonmatch.

Memory Model

The overlap memory model assumes that the representation of letters in memory is distributed. Mathematically, these distributions are assumed to be normal. The distributions for the different letters of a string are centered at equal distances on a spatial axis. The amount that a particular letter spreads into adjacent positions depends on the variance of its distribution. It is assumed that the distributions for different positions have different variances and these variances are the free parameters of the model. The distributions of letters in a test string are assumed to have small variance so that they do not spread into adjacent positions (see Figure 1). The amount of overlap between the representations of the study and test strings is assessed in the following way. First, for each position, the area of the test distribution of the letter in that position (always equal to one) is multiplied by the area of the study distribution of the same letter in that position (i.e., the area within the slot). Then the results of this multiplication are summed across all positions. Mathematically this corresponds to:

$$\sum_{i=1}^{5} \int_{i-½}^{i+½} f_T(x)dx \int_{i-½}^{i+½} f_S(x)dx, \quad (1)$$

where i is the center of the slot, $f_T(x)$ is the distribution of a test item centered on i, $f_S(x)$ is the distribution of a study item, and x is position along the spatial axis. There are many other possible choices for the assessment of overlap between the study and test strings; for example, a correlation of the two strings can be performed. However, most other schemes will introduce more parameters and require further assumptions about the nature of the distributions of items in the test string. Until there are experimental data that can give us some information about the test distributions, the simple form presented above is used.

If a letter in a particular position in the study list is replaced by a new letter in the test list, then there is zero overlap between the study and test letters in that position (in fact, similarity or similar letters in the same position on earlier trials could provide some overlap, but for our purposes here we will assume zero overlap). When two letters are switched, the tails of the distributions of the letters in the memory representation overlap with the switched letters in the test string. When two adjacent letters are switched, there is considerable overlap, sometimes nearly as much as in the *same* condition.

Figure 2 shows two examples of the assessment of overlap. For study condition *i* the study string is AXCDE and the test string is ABCDE. The amount of overlap between study and test strings works out to be the shaded areas in the top panel. In study condition *ii* ABCED is studied and ABCDE is tested. The shaded areas show the considerable overlap between the tails of E and D in the memory representation of the studied string and D and E in the test string.

The possibility that item information can be lost from the memory representation is not included in the model as presented here because the delay between study and test presentation is short and the number of letters in the strings is small.

Numerical examples. There are two extremes in the parameter space of the memory model: First, when the letters in the study distributions have small variances and thus little spread into adjacent positions, and second, when the letters in the study distributions have large variances and hence large spread into adjacent positions. Numerical examples of the amount of overlap for *same*, single- and double-replace *different*, and switch-*different* conditions derived from Equation 1 are given in Table 1.

In order to interpret these numbers it is necessary to note that the subject is performing a discrimination between the *same* and the various *different* conditions. In other words, the subject must set a criterion to separate the amounts of overlap for strings that match the study string from the amounts of overlap for strings that do not match the study string. The largest amount of overlap between study and test strings is in the *same* condition. The next largest amount of overlap is in the most difficult *different* condition and so on through to the smallest amount of overlap in the easiest *different* condition. When the standard deviation for each item is .5, single replacements are the most difficult negatives (overlap is highest in these negative conditions). There is little spread so that even adjacent switches (i, $i+1$) are quite accurate. In contrast, when the standard deviation for each item is 4, there is considerable spread of items into adjacent positions, and the amount of overlap for the adjacent switch condition is almost as large as the amount of overlap for the *same* condition. In this case even switches between items four positions apart are less accurate than single replacements. From these two examples we can make the following observation: If the amount of overlap is lower for adjacent switch conditions than for single replace conditions, then the letter distribution spreads are relatively narrow and the subject has relatively good representation of the order of the study letters. On the other hand, if the adjacent switch conditions have high overlap relative to the single replace conditions, then the letter spread distributions are relatively wide and the subject has

Table 1
Overlap Values for Various Parameter Values and Conditions

	s_i		
Condition	.5	1.0	4.0
Same	3.41	1.92	.50
Single replace	2.73	1.53	.40
Double replace	2.05	1.15	.30
Switch i, $i+1$	2.36	1.63	.49
Switch i, $i+2$	2.05	1.27	.47
Switch i, $i+3$	2.05	1.16	.45
Switch i, $i+4$	2.05	1.15	.42

Note. s_i stands for the standard deviation of the item distributions (all distributions are set to have the same standard deviation).

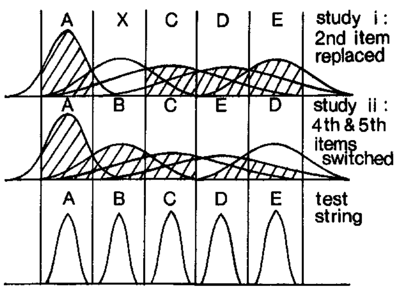

Figure 2. Examples of evaluation of overlap between study and test strings. (Study i [top panel] represents the case where AXCDE is studied and ABCDE is tested. Total overlap is given by the shaded area. Study ii [middle panel] represents the case where ABCED is studied and ABCDE is tested. Overlap is given by the shaded area.)

poor representation of order of the study letters. It ought to be noted that, in practice, the item distributions do not have equal standard deviations as in the examples presented above, but this does not change the point of the examples.

Interfacing the Memory and Retrieval Models

It is necessary at this point to describe how the memory and retrieval models interface or connect. The perceptual matching task is essentially a discrimination task; that is, subjects have to discriminate between *same* and *different* trials. The connection between the memory and retrieval models is made through a series of discriminability statistics of d's, one for each *different* condition, scaled against the *same* condition. Thus, the retrieval model is fitted to accuracy and reaction time data to produce a set of d' scores to which d' scores derived from the memory model can be compared. The memory model produces d' scores by calculation of the difference between the amount of overlap for *same* and *different* and transformation of this difference so that small differences are magnified. This transformed value is labeled d'_m. Details of the transformation will be taken up later.

Raw accuracy scores cannot be used to generate the d' statistics to which the memory model is fitted, for the following reason: The estimated value of d' derived from raw accuracy scores will depend on the particular speed–accuracy criterion set by the subject; d' would be higher the more accurate and slower the subject attempted to be. But an important property of the memory model is that it should be independent of the particular retrieval parameters set by a subject. Thus, because the diffusion model allows computation of asymptotic discriminability d's (d'_a) that are independent of speed–accuracy criteria (Ratcliff, 1978), d'_a values computed from the retrieval model should be used as the meeting point of the memory and retrieval models.

Further discussion of the retrieval model will come after Experiment 1 because mathematical details of the model have been presented in a previous paper (Ratcliff, 1978) and a similar model has been presented by Krueger (1978).

Empirical Applications

In order to provide tests of the model, the perceptual matching procedure will be used with the manipulation of the order of the letters in the test string as the variable of principal interest. There will be two experimental tests, one using the usual reaction time procedure and one using a response signal procedure that measures accuracy as a function of time. The procedures of both experiments begin with a study phase in which a string of letters is presented to the subject. This is followed by one of three types of test strings: (a) The letters are identical to the studied string, (b) one or two of the studied letters are replaced by new letters, or (c) the order of two of the studied letters is switched. The subject's task is to recognize that a test string in which letters are switched or replaced is different from the study string. In the first experiment subjects are allowed to respond in their own time, and reaction time and accuracy data are collected. In the second experiment subjects are required to respond at one of several time deadlines, and accuracy as a function of time provides the data of principal interest.

These experiments differ from the usual perceptual matching experiment in the inclusion of the switch condition. Previously, the major variables used in perceptual matching have been the lengths of the study and test strings and the number of letters replaced in the replacement condition. Previously proposed models will be examined in light of the data obtained from this procedure, with the result that the models for perceptual matching (Bamber, 1969; Krueger, 1978; Taylor, 1976) and Shiffrin and Cook's (1978) model for order information are shown to be inadequate. It will then be shown that a conjunction of a random-walk retrieval model and an overlap model for order information provides a good account of the data, including the relative difficulty of experimental conditions (both reaction time and accuracy) and the shapes of reaction time distributions for Experiment 1, and the relative difficulty of experimental conditions and the time course data for Experiment 2.

Experiment 1

Method

Subjects were presented with a study string of five letters followed by a mask and then a test string. Subjects were instructed to respond *same* if the study and test strings matched exactly. They were to respond *different* either if one or more letters in the study string were replaced in the test string by new letters or if two of the letters in the study string were interchanged in the test string. One third of the trials were positive (*same* correct) and two thirds were negative. There were 25 types of negative trials: 5 were replacements of one letter, 10 were replacements of two letters, and 10 were interchanges of two letters. The data of main interest are the reaction time and accuracy of *different* responses for the 25 types of negative trials.

In the perceptual matching paradigm, other researchers have used list length manipulations to obtain data for testing and developing models. Elsewhere (Ratcliff, 1978), I have argued that such between-trial variables allow the subject to alter speed–accuracy criteria before retrieval and therefore are not particularly useful for testing models. Thus the variables used in this study are within-trial variables so that the subject cannot anticipate the experimental condition before the test item is presented.

Subjects. Four undergraduate volunteers served as subjects in 12 50-min. sessions and were paid at a rate of $3 per 50-min. session.

Apparatus. The stimuli were presented on a Datamedia Elite 1520 video display terminal. The video terminal was controlled by a microcomputer that was interfaced to the main university time-sharing computer.

Stimuli. The 10 consonants *C, D, F, H, J, K, L, R, S,* and *T* were used in constructing the study and test strings. These consonants were used by Taylor (1976) in his perceptual matching experiments and were chosen for ease of identification and minimum confusability in the 5-by-7-dot-matrix character set produced by the display unit. The display letters were 5.1 mm high and 2.4 mm wide with one dot separating adjacent characters. Viewing distance was not controlled but was approximately 60 cm. For a string of 5 characters, the visual angle subtended was 1.3°.

For each trial the letters of the study string were randomly selected from the 10 letters without replacement (there were no repeated letters in either the study or test strings). In a block of 75 trials, there were 25 negative trial types, each represented twice, and 25 positive trials. Five of the negative trial types had one of the study letters (1 to 5, respectively) replaced by a new letter randomly selected from the remaining letters. Ten of the negative trial types had all combinations of two of the study letters replaced by new letters. The remaining 10 negative conditions represented all combinations of two study letters interchanged. The order of the negative trials in a block of 75 trials was randomized along with the 25 *same* trials. Subjects received 12 blocks of 75 trials in a 50-min. session.

Procedure. Each trial began with a fixation point that remained on for 1 sec. The study string then replaced the fixation point and subjects studied it for 1.5 sec. The study string then disappeared and was replaced after 50 msec by a five-asterisk mask that remained on for 200 msec. Following a 50-msec blank period, the test string was presented for 250 msec (this brief presentation was used to reduce the possibility that eye movements would take place), and then the screen remained blank until the subject responded. Following the response the fixation point for the next trial appeared. After each block of 75 trials, subjects took a self-paced break.

Results

Response times and error rates for the positive condition and the 25 negative conditions, averaged over subjects, are shown in Table 2. There are three results of immediate interest: First, responses in the *same* condition are considerably slower than the fastest *different* responses. This result immediately causes problems for those models that propose a fast identity matcher for positive responses; this point will be discussed later. Second, the data show dramatic effects of interchanges on both reaction time and accuracy. When adjacent letters are switched, accuracy is poor and reaction time is long; when nonadjacent letters are switched, accuracy is better and reaction time is shorter. Third, performance is better when two letters are replaced than when two adjacent letters are switched. As will be discussed later, these results rule out any model that assumes independent comparisons of single letters and require the development of a model that accounts for interactions between adjacent letters.

Fitting the Model to the Data

In fitting the model to the data from Experiment 1, the overall model is decoupled into the memory and retrieval components. Fits of the retrieval model are made to the reaction time and accuracy data, yielding d'_a measures of the discriminability of each negative condition from the positive condition. These 25 values of d'_a are the data to which the memory model is fitted.

Table 2
Accuracy and Mean Reaction Time for Experiment 1 Averaged Over Subjects

Item	Proportion correct				Reaction time for correct (msec)			
				Single replace				
1	.956				511			
2	.879				582			
3	.889				580			
4	.834				654			
5	.881				665			
				Double replace				
	2	3	4	5	2	3	4	5
1	.972	.973	.976	.975	495	491	506	499
2		.955	.958	.967		518	531	520
3			.941	.966			533	530
4				.935				564
				Switch				
	2	3	4	5	2	3	4	5
1	.940	.959	.973	.972	564	527	522	508
2		.753	.855	.919		738	649	595
3			.655	.875			795	667
4				.734				853
				Same				
	.834				709			

Note. Number of observations for *same* is about 14,400 and for each *different* condition about 1150. Standard error in reaction time for *same* is 3 msec; representative values for the *different* conditions are 500 ± 6 msec, 650 ± 10 msec, 800 ± 15 msec. In the single replace condition, the item number indicates the item replaced. In the double replace and switch conditions, the row and column numbers indicate the pair of items replaced or switched.

Retrieval Model

In this section the retrieval model will be presented and fitted to the data from Experiment 1. The model assumes a comparison process that is a steady accumulation of evidence towards either a match or a nonmatch decision. The process is more easily understood in the discrete rather than the continuous random-walk formulation, so the discrete case is discussed first. In the discrete case a count is kept of the number of feature matches and nonmatches between the study and test strings; if a feature match occurs, the counter is increased by one, and if a feature nonmatch occurs, the counter is decreased by one. The counter begins at some base number of counts z. If the counter reaches a counts, a match results, and if the counter reaches 0 counts, a nonmatch results. In the continuous case evidence is continuously accumulated over time. The rate of accumulation is called the drift rate; the better the match between the study and test strings, the faster the drift toward the match boundary; the poorer the match between the study and test strings, the faster the drift toward the nonmatch boundary.

The comparison process has several parameters. First, the distances from the starting point to the match ($a-z$, Figure 1) and nonmatch ($z-0$) boundaries determine how much evidence is needed for a match or nonmatch. These parameters can be varied to model speed–accuracy trade-offs. Second, drift (derived from the relatedness of the study and test strings being compared) has a variance (s^2) associated with it. Variability in drift represents the noise associated with the comparison process. In the discrete feature-matching conceptualization, this variance arises from randomness in the order of feature matches and nonmatches. In the continuous case this variance is assumed to arise because of noise in the assessment of overlap between study and test strings. Third, it is assumed that another source of variance is associated with the average value of relatedness or overlap. Thus the average drift rate varies across trials (for a single experimental condition) and this variance is separate from the variance within a comparison. Variance in the average drift rate across trials could arise from variability in the encoded representation of the study string. For example, on some trials some letters may be better encoded (leading to less spread into adjacent positions), but on other trials these letters may not be encoded as well (leading to more spread). The distributions of relatedness between the study and test strings are assumed to be normal distributions; for matches, the distribution is $n(u, \eta)$ and for nonmatches, $n(v, \eta)$.

It has been demonstrated (Ratcliff, 1978; Reed, 1976) that it is necessary to assume that there are two sources of variance (as described above) in order to deal with results from experiments that examine the growth of accuracy as a function of time (e.g., response signal or deadline procedures; see Experiment 2 below). If there is no variability across trials ($\eta^2 = 0$), then accuracy will not approach asymptote as a function of time (see Experiment 2); rather it will grow as a function of the square root of time. The variability within a trial ($s^2 > 0$) is necessary because otherwise accuracy would rise to asymptote as soon as the comparison began (as soon as any information was available).

In experiments in which the length of the study and test strings is varied, it is possible that subjects adjust boundary positions according to list lengths (e.g., short lists are easier so boundaries can be set close to the starting point because little evidence is needed for high accuracy). With boundary parameters free to vary, such data do not provide strong tests of models.

In Experiment 1 all the experimental manipulations took place at List Length 5 rather than at different list lengths. Thus, subjects could not anticipate whether the trial was a *same*, single replace, double replace, or switch trial before the test string was presented. Thus, boundary positions could not vary and so, to fit the data from all 25 negative experimental conditions (scaled against the positive condition), the only parameter that could vary was nonmatch relatedness (v). Changes in this one parameter have to account for changes in reaction time, accuracy, and the shapes of reaction time distributions across the 25 conditions.

From the fits of the model to the accuracy and reaction time data, a value of v can be obtained for each negative condition, yielding a corresponding asymptotic d'_a value $(u - v)/\eta$. This value measures the asymptotic discriminability between the match distribution and the particular nonmatch distribution under consideration and hence removes bias dependent on the particular speed–accuracy criterion set by the subject that would be introduced into d' measures derived from raw accuracy scores. The memory model is then fitted to this set of 25 d'_a values.

Fits of the retrieval model. The retrieval model was fitted to the data of individual subjects rather than data averaged across subjects for two reasons: First, subjects showed quite different levels of performance; thus, fitting the individual data is a useful test of the model. Second, in the average data shown in Table 2, there are speed–accuracy trade-offs due to averaging across different subjects who show different patterns of results. The results of these fits and the data are shown in Figures 3 and 4. (Details of the fitting procedure are presented in the Appendix.) Figure 3 shows fits of the diffusion model to mean reaction times and error rates for the four subjects for the 25 *different* conditions. The dots represent each condition and the solid lines show the theoretical fits. (Note that it would be possible to improve the overall fits for Subject 3 by allowing the fit to the *same* condition to become poorer, falling outside two standard errors, but this was not done.) Figure 4 shows fits of the model to group cumulative distribution functions. There are several points to note about the data and the fits of the model to the data.

Individual differences. The individual

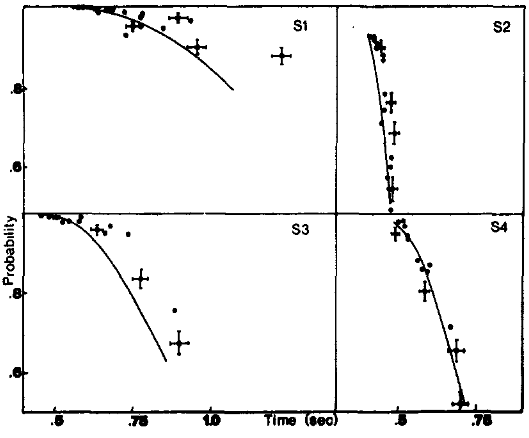

Figure 3. Theoretical predictions and empirical results for probability correct versus mean reaction time for the four subjects. (The error bars [for selected points] represent one standard error. Fits to reaction time and accuracy for *same* responses lie within two standard errors of the data. Parameter values for the diffusion model are as follows: Subject 1, t_{ER} = 295 msec, a = .30, z = .12, u = .38; Subject 2, t_{ER} = 295 msec, a = .08, z = .035, u = .14; Subject 3, t_{ER} = 330 msec, a = .19, z = .08, u = .25; Subject 4, t_{ER} = 335 msec, a = .13, z = .06, u = .23. The v values can be calculated from the d' values in Table 4 using $d' = [(u - v)]/\eta$ where η = .18.)

subjects show vastly different patterns of performance. Subjects 1 and 3 have high accuracy and slow reaction times for the difficult conditions, whereas Subject 2 shows wide variations in accuracy (for *different* judgments, from .93 correct to .49 correct) but relatively fast reaction times that do not vary by more than about 70 msec from the easiest to hardest conditions. The results for Subject 4 fall between these subjects. The fast and inaccurate subject is modeled by having the boundaries of the diffusion process close to the starting point of the process. The slow and accurate subjects are modeled by having the diffusion process boundaries relatively far from the starting point.

The behavior of these types of subjects is reminiscent of the two types of subjects described by Cooper and Podgorny (1976). For their task, which required visual comparison of angular shapes, subjects fell into two different groups: Type 1 subjects showed no reaction time differences as a function of stimulus-test similarity, and Type 2 subjects showed large reaction time differences (the more similar the two patterns, the slower the *different* response). For both types of subjects, they found large differences in accuracy as a function of similarity (the more similar the study and test strings, the more inaccurate the *different* response). By analogy to the results presented here, the differences between the subjects in the task of Cooper and Podgorny (1976) could be accounted for by differences in the settings of the boundary positions in a random-walk comparison process.

Adequacy of retrieval model fits. For Subjects 2 and 4, the fits are rather good, with both reaction time and accuracy falling within two standard deviations of the data for all conditions. For Subjects 1 and 3 (the slow subjects), the fits miss significantly, especially in the conditions in which reaction time was long. An inspection of Figure 4

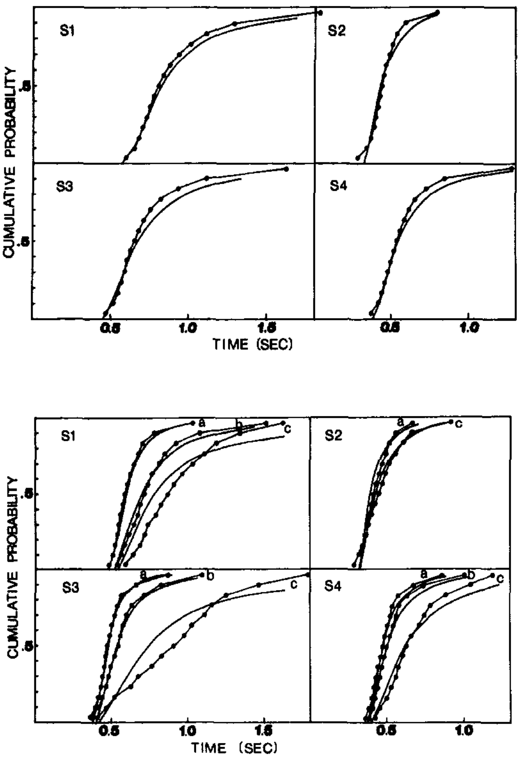

Figure 4. Theoretical (continuous lines) and empirical (dots) cumulative probability functions for *same* responses (top panel) and selected *different* responses. (For the *different* responses *a* represents double replacement of Letters 2 and 4, *b* represents switches of Letters 2 and 5, and *c* represents switches of Letters 3 and 4. Most other negative conditions are similar to condition *a* but these are not shown.)

shows the way in which the model misses in predicting reaction time in these conditions: The experimental reaction time distributions have a greater shift in the relative positions of the negative distributions than the theoretical distributions. However, in general, the fits to the distributions are very good considering that there is only one parameter changing across the 25 *different* conditions. In only 7 out of 104 cases are there serious mispredictions by the model. These mispredictions take the form of the *c* conditions in Figure 4 where the mode of the experimental distribution is slower than the mode of the theoretical distribution and the tail of the experimental distribution is shorter.

Error reaction times. Table 3 shows error reaction time data and fits of the diffusion model to those data. The results show that error reaction times are not well predicted by the model. In general, error reaction times are predicted to be slower than reaction times for correct responses, but the data show that for some subjects errors are faster. Ratcliff (1978) demonstrated that in the diffusion model, predictions of error reaction times are heavily dependent on the shapes of the extreme tails of the relatedness distributions. In contrast, predictions about reaction times for correct responses, shapes of reaction time distributions, and error rates are dependent on the mean and spread of the distributions of relatedness (meaning approximately interquartile spread rather than the variance that depends on the tail of the distribution). It is possible to adjust the distribution shape to keep approximately constant the reaction time for correct responses, reaction time distribution shape, and error rate while varying error reaction time from very fast to very slow (by as much as a second or two; Ratcliff, 1978, p. 87). Thus, if a relatedness distribution other than the normal distribution were chosen, it would be possible to vary error reaction time without affecting the reaction time for correct responses or error rate.

Table 3
Theoretical and Empirical Error Reaction Times (msec)

Response	Subject 1	2	3	4
Same				
Empirical	1634	414	536	543
Theoretical	2850	524	1376	901
Different				
Empirical	1460	425	681	575
Theoretical	2130	479	1134	826

The question of the relative speed of correct and error responses is still a major puzzle in most experimental paradigms, and attempts to account for the data continue. Laming (1979) has provided results on the random-walk model that suggest that the reason that error responses are faster than correct responses in choice reaction time is that the subject begins sampling before the stimulus is presented (thus bringing variability into the starting point of the random walk). Variability in the starting point of the diffusion process along with distributions over relatedness would be sufficient to account for any pattern of error reaction time. To get fast errors the variability in the starting point would be large, and for slow errors variability would be small (slow errors coming from the tails of the relatedness distributions). Thus, Laming's suggestion may apply to the data presented here. In any case, it is clear that the unmodified model does not do a good job of accounting for error reaction times and that much more work is required to decide what factors are responsible for producing slow errors in some conditions and fast errors in other conditions.

Reaction time differences between the same and different conditions. By adding the very difficult switch *different* conditions to the experiment, the usual result that *same* responses are faster than *different* responses has been reversed. This is not owing to the use of long strings, because I have replicated this using three-letter strings (in an unpublished experiment), nor is this result due to the high frequency of *different* trials (2:1 over *same* trials) in the experiment. Ratcliff and Hacker (in press) have shown similar results using equal-frequency conditions. In the model the relative speeds of *same* and *different* responses depend on the position of the criterion on the relatedness scale, and the positions of the response boundaries in the diffusion process (note that these criteria settings also account theoretically for effects of the unequal probability of *same* and *dif-*

ferent trials). Comparing the case in which *different* responses are easy (when there are no switch conditions in the experiment) to the case in which some *different* responses are very difficult, there is a greater range of *different* distributions on the relatedness axis in the latter case (see Figure 5 and Krueger, 1979) than in the former case. Thus, the criterion separating *same* and *different* relatedness values must be placed relatively near the *same* distribution; this leads to slow *same* responses, slow difficult *different* responses, and fast easy *different* responses. Why *same* responses are fast when there are only easy *different* responses in the experiment is discussed by Krueger (1978, 1979).

Given the generally good fit of the retrieval model to the data, we can proceed to connect the retrieval model to the memory model. The retrieval model provides parameter estimates for the mean relatedness values for the *same* condition (u) and the 25 *different* conditions (v). From these values and η it is possible to calculate 25 d' values for the *different* conditions scaled against the *same* condition using $d'_a = (u - v)/\eta$. These values are shown in Table 4 (from these values and the values for u and η in Figure 3, the v values can be computed). These values of d'_a are the basic data to which the memory model is fitted.

Fits of the Memory Model

The difference in amount of overlap between study and test strings for the *same* condition and the particular negative condition studied is transformed to a d' value

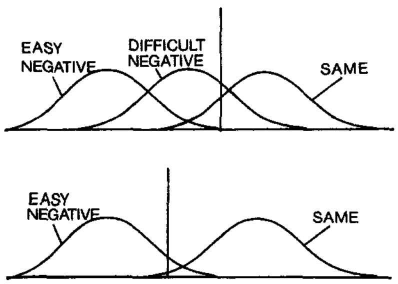

Figure 5. Distributions of total overlap (relatedness) showing criterion placement for conditions in which difficult negatives are included (top panel) and in which there are only easy negatives (bottom panel).

by a power function transformation ($d'_m = a\ (o_s - o_d)^b$ where o_s = *same* overlap and o_d = *different* overlap). These values are then fitted to the d'_a values (shown in Table 4) of the retrieval model by least squares. Values of the five variance parameters and the transformation parameters a and b are presented in Table 5. Table 6 shows the fits of the model. In general the fits of the model are reasonable; there do not seem to be any systematic deviations of the computed d'_m from the d'_a values across subjects.

Power transformation. Before continuing, it is necessary to examine the power transformation relating overlap and discriminability. One assumes that subjects are using amount of overlap as the dimension on which to make their judgment about similarity or difference. However, there is no need to assume that overlap directly measures discriminability. Rather, the best that one can hope for is that discriminability is a monotonic transformation of overlap. Monotonic transformations (e.g., a power function) have often been assumed for the relationship between the perception of a stimulus and the physical stimulus. For letter matching, not only should the transformation be monotonic but it would also be most useful if it magnified small differences in order to aid the most difficult discriminations, as the power transformation does (with exponent less than 1).

Figure 6 shows that the power function describes the relation between the amount of overlap and the d'_a values from the retrieval model for the four subjects (using the values for the variance parameters in Table 6). As can be seen, all the curves show some nonlinearity (each dot represents one negative condition), and the power function (the solid line) captures the relationship between overlap and d'_a.

The overlap model provides a good ordinal representation of the data. Rank order correlations between the data and the overlap model are .956, .959, .986, and .944 for Subjects 1 to 4, respectively. Figure 7 shows a plot of d'_a from the retrieval model versus d'_m from the overlap memory model. This is essentially a plot of the data from Tables 4 and 5 and shows that there is no systematic deviation from linearity as there is in Figure 6.

Table 4
Asymptotic (d'_a) Values for Individual Subjects for Fits of the Retrieval Model to the Data From Experiment 1

	Subject															
Item	1				2				3				4			
	Single replace															
1	4.75				2.28				4.72				3.22			
2	3.94				1.50				3.61				2.39			
3	4.75				1.33				3.61				2.39			
4	3.50				1.06				3.06				2.28			
5	3.22				1.50				3.06				2.56			
	Double replace															
	2	3	4	5	2	3	4	5	2	3	4	5	2	3	4	5
1	5.44	5.72	5.17	5.33	2.44	2.44	2.44	2.44	5.00	5.00	4.72	4.72	3.39	3.39	3.39	3.50
2		5.44	5.06	5.06		2.17	2.00	2.17		4.44	4.17	4.17		2.94	3.22	3.22
3			5.17	5.17			1.89	2.17			3.61	3.89			3.17	3.39
4				4.33				1.78				3.61				2.94
	Switch															
	2	3	4	5	2	3	4	5	2	3	4	5	2	3	4	5
1	3.86	4.56	4.56	5.17	2.06	2.28	2.28	2.28	3.61	4.17	4.44	4.44	3.17	3.22	3.50	3.39
2		3.78	4.06	4.06		.78	.94	1.61		2.33	2.94	3.33		1.67	2.50	2.83
3			3.50	4.06			.67	1.22			1.67	2.78			1.28	2.39
4				2.94				.89				1.94				1.83

Note. In the single replace condition, the item number indicates the item replaced. In the double replace and switch conditions, the row and column numbers indicate the pair of items replaced or switched.

Table 5
Variance and Transformation Parameters for the Memory Model for Experiment 1

	Parameter						
Subject	s_1	s_2	s_3	s_4	s_5	a	b
1	1.24	1.83	1.39	4.58	2.92	6.11	.21
2	.59	1.43	1.63	2.05	1.42	2.70	.50
3	1.05	2.17	3.23	6.08	3.93	5.81	.26
4	.65	1.62	2.11	2.01	1.53	3.78	.28

Note. Parameters a and b are from the power function transformation $d'_m = a\ (o_s - o_d)^b$.

The power function was the third transformation considered for relating d'_m values to amounts of overlap. The first was $d'_m = a\ (o_s - o_d)$, which is the simplest possibility. This produced very poor fits to the data. The second transformation considered was the linear transformation, $d'_m = a\ (o_s - o_d) + b$, which was designed to improve the bad fits obtained from the first transformation on the conditions in which d' was small. Fits were better, but there were systematic deviations in the adjacent switch conditions consistently across subjects. Furthermore, the linear transformation was difficult to interpret in that the parameter b produced a discontinuity in the d'_m scale (d'_m could not be below b). The fits of the linear transformation suggested that the scale of overlap needed stretching at low d'_m values (i.e., when the amount of *different* overlap was very close to the amount of overlap for the *same* condition), and contracting at high d'_m values. The power function does just that.

Adequacy of memory model fits. There are seven parameters in the model, five variance parameters and two transformation parameters. These parameters are not in a one-to-one correspondence with the data points. For example, a change in one variance parameter affects 9 out of the 25 d'_m values, and changes in the transformation parameters affect the levels of all 25 d'_m values. Thus attempting to improve the fit to one data point by adjusting just one parameter may perturb fits to at least eight other conditions. This indicates that the structure of the model captures relationships between the different experimental conditions. From this and the fact that the two versions of the linear transformation did not produce good fits to the data, it can be concluded that the data provide reasonably tight constraints on the model.

The parameter values across subjects show interesting patterns. First, the values of the variance parameters are in the same range across all subjects (except for two high values, 6.08 and 4.58). Much of the difference between subjects arises in the transformation parameters. This suggests that the limit on performance is in the transduction process between the memory representation and the d' values (i.e., drift in the diffusion comparison process).

Second, the values of the overlap parameters in some cases are fairly large. With a standard deviation of 2, 20% of the distribution is in the correct position, and 16%, 12%, and 7% extend to the successively adjacent positions. Thus, it is necessary to assume that there is considerable positional uncertainty.

The Time Course of Processing

There are two complementary methods for investigating the temporal properties of information processing. The first is exemplified by Experiment 1, in which reaction time and accuracy measures were obtained. This kind of processing can be termed *information controlled* (Ratcliff, 1978), because the subject is free to respond when he or she has obtained enough information for a decision. The second method involves determining the growth of accuracy as a function of time by requiring the subject to respond at experimenter-determined time deadlines, that is, to respond immediately when a signal is given. This kind of processing, termed *time controlled,* is investigated in Experiment 2.

Table 6
d'_m *Values for Fits of the Memory Model to* d'_a *Values Derived From the Retrieval Model (Table 4)*

	Subject															
Item	1				2				3				4			
Single replace																
1	4.78				2.09				4.49				3.21			
2	4.41				1.41				3.76				2.54			
3	4.67				1.32				3.40				2.36			
4	3.64				1.18				2.89				2.39			
5	4.00				1.41				3.23				2.58			
Double replace																
	2	3	4	5	2	3	4	5	2	3	4	5	2	3	4	5
1	5.34	5.47	5.04	5.16	2.52	2.47	2.40	2.52	4.98	4.84	4.68	4.78	3.56	3.49	3.50	3.57
2		5.27	4.74	4.90		1.93	1.84	2.00		4.29	4.07	4.21		2.98	3.00	3.11
3			4.94	5.08			1.78	1.94			3.79	3.96			2.89	3.01
4				4.45				1.84				3.67				3.02
Switch																
	2	3	4	5	2	3	4	5	2	3	4	5	2	3	4	5
1	3.85	5.03	4.77	5.02	1.83	2.29	2.30	2.52	3.53	4.34	4.49	4.61	2.90	3.30	3.41	3.56
2		3.67	3.77	4.40		.84	1.32	1.88		2.23	2.90	3.49		1.72	2.41	2.95
3			3.40	4.36			.67	1.46			1.61	2.44			1.55	2.44
4				2.29				.76				1.42				1.79

Note. In the single replace condition, the item number indicates the item replaced. In the double replace and switch conditions, the row and column numbers indicate the pair of items replaced or switched.

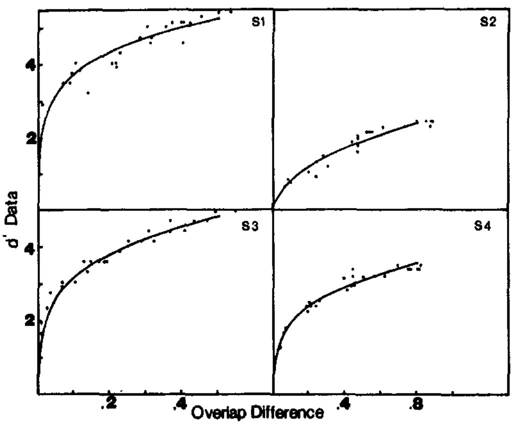

Figure 6. d' derived from the retrieval model and the experimental data plotted against the difference in overlap between the *same* condition and the particular *different* condition. (Each dot represents one *different* condition and the solid line represents the power function.)

The diffusion retrieval model makes a strong prediction about the form of the function describing the growth of accuracy as a function of time. The expression for d' as a function of time is derived as follows: First, the boundaries on the decision process are removed so that there is no criterion for amount of evidence accumulation. Second, it is assumed that if the diffusion process is on the match side of the starting point of the process when the signal to respond is given, then the subject initiates a *same* response. If the process is on the nonmatch side, the subject initiates a *different* response.

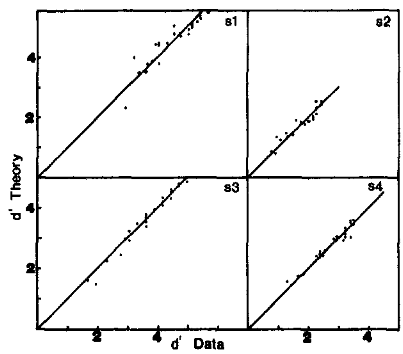

Figure 7. d' data refers to the asymptotic d'_a values derived from the data using the retrieval model and d' theory refers to the fits derived from the memory (overlap) model. (These data are plotted from Tables 4 and 6. Each dot represents one negative condition and the solid line is the line of perfect prediction.)

The expression that describes d' as a function of time is given by:

$$d'(t) = \frac{d'_a}{\sqrt{1 + s^2/(\eta^2[t - t_{\mathrm{ER}}])}}, \quad (2)$$

where d'_a is the asymptotic d' value for the particular negative condition, s^2 is the variance in the comparison process, η^2 is the variance in drift, t_{ER} is the encoding and response parameter, and t is elapsed time from the presentation of the test string (see Ratcliff, 1978). The form of this function is the same for each *different* condition and the only free parameter between conditions is the asymptotic d'_a value for each curve. Thus, fits of the model to the response-signal curves are another reasonably constrained test of the retrieval model. Furthermore, the values for the asymptotic d's can be used in fitting the memory model in the same way as in Experiment 1.

An important point to note is that according to Equation 2, at any given time t, the ratio of $d'(t)$ for the *different* conditions will be identical to the ratio of d'_a for those conditions. This shows that the memory model can be tested by performing experiments in which subjects respond at a single constant deadline (not at one of several deadlines as was done in Experiment 2). The parameter that changes to account for the difference between the $d'(t)$ value and the d'_a value is a in the power transformation.

Experiment 2

Method

Experiment 2 was designed to provide further tests of both the memory model and the retrieval model by providing data complementary to those of Experiment 1, that is, data showing the growth of accuracy as a function of time.

The matching conditions used in Experiment 2 included the *same*, single replace *different*, and switch *different*. Double replace trials were dropped to reduce the total number of sessions required per subject. Details of the procedure followed Experiment 1 except that subjects were required to respond at one of five response-signal lags.

Subjects. Two undergraduate volunteers served as subjects and were paid at a rate of $3 per 50-min. session. The subjects served in 18 50-min sessions preceded by two practice sessions.

Stimuli. Construction of the stimuli was the same as for Experiment 1. Subjects received 18 blocks of 45 trials (15 *same*, 10 single replace, and 20 switch trials).

Procedure. Each trial began with a fixation point that remained on for 1 sec. The study string then replaced the fixation point and subjects studied it for 1 sec. The study string then disappeared and was replaced after 50 msec by a five-asterisk mask that remained on for 200 msec. Subjects were presented with one of five response-signal lags: 100, 200, 400, 900, or 2000 msec. For the 100-msec lag, after presentation of the mask a 50-msec blank period preceded presentation of the test string. The test string was removed 100 msec after it appeared, and the signal to respond was then presented (a row of + signs one line below the test string). For the other response-signal lags, the test stimulus remained on for 200 msec and was then removed before the signal to respond was presented. Subjects were instructed to respond within 200 to 300 msec of the presentation of the response signal. Following the response of the subject, feedback as to the response latency was presented for 250 msec. Next, the fixation point was again presented and the next trial began. The five response-signal lags were randomly assigned to the *same* and *different* conditions so that on average about seven responses per session were made for each *different* condition at each lag. Subjects were allowed a self-paced rest after each 45 trials.

Results

The main results are shown in Figure 8. Reaction time to the response signal is included in the computation of the response time in Figure 8.

The fit of the function $d'(t)$ allows estimates to be obtained for the asymptotic d'_a values for each negative condition, the encoding and response parameter t_{ER}, and the ratio s^2/η^2. Fits of the diffusion model are shown along with the experimental results in Figure 8. The parameter estimates for the d'_a values for the negative conditions are shown in Table 7.

The memory model is fitted to these d'_a values as in Experiment 1. The parameter values are shown in Table 7 and the fits are shown in Table 8. Again we see large individual differences between subjects. Subject 1 is quite accurate in the switch conditions; thus, estimates of the item distribution standard deviations are quite small. In contrast, Subject 2 is quite inaccurate in switch conditions; thus, the estimates of the standard deviations are large.

General Discussion

Evaluation of Alternative Models

Most of the models that would be considered competitors to the model presented

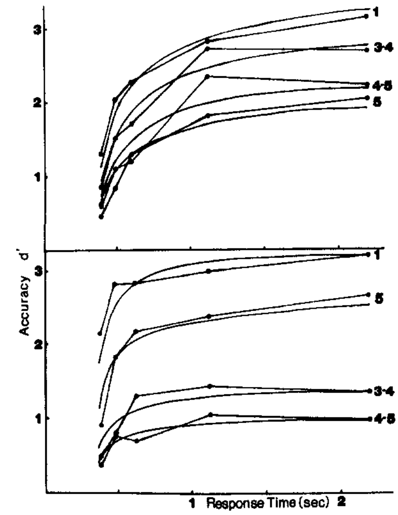

Figure 8. d' as a function of response time for Experiment 2. (The dots represent the experimental data and the continuous lines represent the fits of Equation 2. Only 4 of the 15 curves were selected for display; 1 and 5 refer to Single Replace Conditions 1 and 5, respectively, and 3-4 and 4-5 refer to Switch conditions between Items 3 and 4 and 4 and 5, respectively. Parameters for the fits are as follows: Subject 1, t_{ER} = 347 msec, s^2/η^2 = 370 msec; Subject 2, t_{ER} = 320 msec, s^2/η^2 = 167 msec. Values for the asymptotic d' values can be found in Table 8. Standard errors for representative d' values are as follows: .5 ± .1, 1.0 ± .12, 2.0 ± .14, 3.0 ± .2.)

above would assume letter-by-letter comparisons between the study and test strings. One such model is that of Shiffrin and Cook (1978), designed to account for ordered recall of items from short-term memory. The model assumes that items are stored so as to represent item information that allows identification of a presented item and to include order links between adjacent items. The retrieval process in ordered recall is assumed to be a serial process. First, the order

Table 7
Variance and Transformation Parameter Values for the Overlap Memory Model

Subject	Parameters						
	s_1	s_2	s_3	s_4	s_5	a	b
1	.52	.75	.71	.82	1.11	4.14	.53
2	1.03	1.97	2.82	5.94	5.06	4.60	.24

Note. Parameters a and b are from the power function transformation $d'_m = a\,(o_s - o_d)^b$.

Table 8
Asymptotic d′ *Values Computed From the Response Signal Curves (Experimental) and From the Overlap Memory Model (Theoretical) for Experiment 2*

Item	Experimental				Theoretical			
				Subject 1				
Single replace								
1		3.52				3.33		
2		2.75				2.85		
3		3.15				2.92		
4		2.53				2.73		
5		2.07				2.37		
Switch	2	3	4	5	2	3	4	5
1	3.42	3.97	4.50	4.37	3.59	4.47	4.38	4.16
2		3.21	3.68	4.07		3.04	3.89	3.75
3			2.99	3.47			2.90	3.61
4				2.37				2.30
				Subject 2				
Single replace								
1		3.54				3.64		
2		3.16				3.15		
3		3.14				2.90		
4		2.62				2.44		
5		2.60				2.53		
Switch	2	3	4	5	2	3	4	5
1	3.31	3.35	3.47	3.63	2.98	3.56	3.65	3.69
2		1.75	2.56	2.97		2.06	2.57	2.94
3			1.42	2.18			1.54	2.12
4				1.02				1.13

Note. In the single replace condition, the item number indicates the item replaced, and in the switch condition, the row and column numbers represent the items switched.

of items is determined by intact order links. Then the subject guesses at the remaining unordered items using item information. To apply this model to the perceptual matching task, the obvious strategy is to compare serially items in the study and test strings. For the experiments presented above, loss of item and order information should not be a problem according to this model. There should be less than 3% chance that an order link is lost between study and test (Figure 9 in Shiffrin & Cook, 1978). As it stands this model is unable to account for the large difference in performance between the double replace condition and the condition in which adjacent items are interchanged. The reason is that as soon as a nonmatching item is encountered in the serial comparison process, a negative response will be made. However, it should be noted that Shiffrin and Cook's model was not designed to account for reaction time data, and perhaps in light of these data, the retrieval assumptions could be modified and some mechanism added that would allow an item adjacent to the item being processed to have an effect on the comparison process.

Other models of the perceptual matching process (Bamber, 1969; Taylor, 1976) are inadequate in the same way as the Shiffrin and Cook model. Any unelaborated serial or parallel matching process must predict that reaction time and accuracy are the same for double replace conditions and conditions in which two items in the study string are switched; this prediction is clearly disconfirmed by the data.

Krueger's model (1978) is very closely related to the retrieval model presented above, but there are some differences. First, Krueger's model has no provision for distributed representations of letters. The model has only one source of variability and that is variability in the comparison process. There is no allowance for variability in the quality of the encoding, so accuracy should not level off with time as in Experiment 2 (unless a deadline is reached). Second, in Krueger's model, letters are represented as feature bundles, and letters that do not match have a surprising number of features in common. In the worst case only 11 out of 100 features are assumed to mismatch physically. This seems to conflict with the idea of a dot-matrix representation as described by Krueger (1978) because one would expect far less overlap between different letters. Furthermore, the way multielement strings are processed seems arbitrary: If one mismatch is registered in one position, then the probability of registering a mismatch in an adjacent position is reduced. In the memory model presented here, this kind of assumption is not necessary. Another difference between the models is that Krueger's model assumes discrete processing, one comparison taking about 200 msec, whereas the diffusion overlap model assumes continuous processing. However, even though the models do differ considerably in detail, they are very much alike in their approach to retrieval processes. Both assume a sequential sampling scheme; both are applied to accuracy, reaction time, and reaction time distributions; and both reduce the importance of fast *same* judgments from a separate process to little more than a criterion effect.

Application to Other Paradigms

One important property of a model is its ability to generalize across experimental paradigms. In this section, I consider how the model presented in this paper generalizes to some related paradigms.

The converse of the perceptual matching task. Taylor (1976) has introduced a variant of the Bamber (1969) task in which subjects have to respond positively if one or more letters match in position, and negatively if no letters match in position. Results show that as the number of matching letters decreases, reaction time increases and accuracy becomes poorer. Krueger (1978) obtained adequate fits of his retrieval model to these data. Therefore, it is likely that fits of the diffusion retrieval model presented here would also be adequate. In the overlap model the parameter that would vary is the relatedness parameter for positive comparisons: If all letters matched, there would be considerable overlap between the study and test strings. As the number of matching letters decreased, the overlap would decrease and reaction times and error rates would increase.

Multiple element comparison task. Angiolillo-Bent and Rips (Note 1) have performed two experiments designed to test various models of item and order information. In their experiments three letters were presented as the study string. The test string was either a permutation of the study string, to which the subject was to respond positively, or a permutation with one letter replaced by a new letter, to which the subject was to respond negatively. They found that performance was best, with low error rate and fast responses, when the study and test strings were identical (e.g., study ABC and test ABC). The next best test condition was ACB, followed by BAC. The remaining three permutations were all equally worse.

The overlap model will predict exactly this pattern of results. The letter *A* has smaller variance in the fits of the model to the data than the other letters because it is in Position 1; in fits of the model to the experiment presented above (and other data), Position 1 consistently has the smallest variance estimates. Thus, it would be expected that ACB would have greater overlap with ABC than would BAC. The other three permutations all have three letters exchanged, so they should have the poorest performance.

The model proposed by Angiolillo-Bent and Rips (Note 1) includes a position-sensitive component such that the number of steps necessary to permute the test string to produce the study string helps account for reaction time differences. This mechanism captures the kind of structure proposed in the overlap model. Angiolillo-Bent and Rips's proposal will not, however, be able to account for the effects found in Experiment 1.

Their permutation scheme will require some other decision mechanism; if the study and test strings had only two adjacent letters switched, only one switch would be needed to convert the test string to the study string, so performance should be good, but in fact it is poor.

Conclusion

The theory developed here integrates a random-walk retrieval model and an overlap model for the representation of order information into a single scheme for the perceptual matching task. The strength of the theory is that it provides integration of several different areas of research: The overlap model is related to the perturbation model of Lee and Estes (1977) that accounts for data from short-term memory recall studies. The retrieval model is a variant of the random-walk model that has been used to model retrieval in a number of experimental paradigms. The theory also provides an account of many aspects of the data, including reaction time, accuracy, and the time course of processing with the retrieval model; and performance on several kinds of *different* trials (including the new switch condition) with the memory model.

An important aspect of the memory model presented here is the assumption that the representation of an item is distributed over position rather than being precisely located in a single position. This conception goes against the view of letter-matching processes, in which it is assumed that items are compared one by one with no effect of nearby items. In contrast, in research in the areas of letter and word perception, the assumption of positional uncertainty is well established. Thus, the model presented here may help relate the areas of letter perception and letter matching.

Reference Note

1. Angiolillo-Bent, & Rips, L. *Order information in multiple element comparison.* Manuscript submitted for publication, 1981.

References

Bamber, D. Reaction times and error rates for "same"-"different" judgments of multidimensional stimuli. *Perception & Psychophysics*, 1969, *6*, 169–174.

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

Cooper, L., & Podgorny, P. Mental transformations and visual comparison processes: Effects of complexity and similarity. *Journal of Experimental Psychology: Human Perception and Performance*, 1976, *2*, 503–514.

Estes, W. K. An associative basis for coding and organization in memory. In A. Melton & E. Martin (Eds.), *Coding processes in human memory.* Washington, D.C.: Winston, 1972.

Krueger, L. E. A theory of perceptual matching. *Psychological Review*, 1978, *85*, 278–304.

Krueger, L. E. A model of unidimensional perceptual matching. *Journal of Experimental Psychology: Human Perception and Performance*, 1979, *5*, 277–288.

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

Laming, D. Choice reaction performance following an error. *Acta Psychologica*, 1979, *43*, 199–224.

Lee, C., & Estes, W. K. Order and position in primary memory for letter strings. *Journal of Verbal Learning and Verbal Behavior*, 1977, *16*, 395–418.

Lee, C., & Estes, W. K. Item and order information in short-term memory: Evidence for multilevel perturbation processes. *Journal of Experimental Psychology: Human Learning and Memory*, 1981, *7*, 149–169.

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

McCloskey, M., & Glucksberg, S. Decision processes in verifying category membership statements: Implications for models of semantic memory. *Cognitive Psychology*, 1979, *11*, 1–37.

Murdock, B. B. Item and order information in short-term serial memory. *Journal of Experimental Psychology: General*, 1976, *105*, 191–216.

Nickerson, R. On the time it takes to tell things apart. In J. Requin (Ed.), *Attention and performance VII.* Hillsdale, N.J.: Erlbaum, 1978.

Ratcliff, R. A theory of memory retrieval. *Psychological Review*, 1978, *85*, 59–108.

Ratcliff, R. Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin*, 1979, *86*, 446–461.

Ratcliff, R., & Hacker, M. The relative speed of same and different judgments: Another tradeoff. *Perception & Psychophysics*, in press.

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

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

Shiffrin, R., & Cook, J. Short-term forgetting of item and order information. *Journal of Verbal Learning and Verbal Behavior*, 1978, *17*, 189–218.

Stone, M. Models for choice reaction time. *Psychometrika*, 1960, *25*, 251–260.

Taylor, D. Effect of identity in the multiletter matching

task. *Journal of Experimental Psychology: Human Perception and Performance*, 1976, *2*, 417–428.
Thomas, E., & Ross, B. On appropriate procedures for combining probability distributions within the same family. *Journal of Mathematical Psychology*, 1980, *21*, 136–152.

Appendix

Fitting the Retrieval Model to Data

Mathematical Expressions

In order to fit the retrieval model to the data (reaction time, distributions, and accuracy) it is necessary to use expressions that relate the theoretical parameters (a, z, s, η, u, v, and t_{ER}) to the empirical results. These expressions can be found in the appendix to Ratcliff (1978). Equation A8 gives the probability of a nonmatch $\gamma\_$ (ξ) for a particular drift rate ξ. Equation A12 gives the expression for the first passage time distribution function G _ (t, ξ) (the cumulative density function) for a particular drift rate ξ. Equations A24 and A25 give the expressions for the first passage time distribution functions averaged over a distribution of drift rates—n(u, η), for example. These expressions are sufficient to allow the model to be fitted to mean reaction time (by integrating t times the density function), reaction time distributions, and accuracy.

It should be noted that in all these expressions the encoding and response output time parameter t_{ER} has been ignored. In the fits to the model it is estimated along with the other parameters. It is assumed that the encoding and response output processes have variability small enough to be ignored.

Fits of the Retrieval Model

Fitting all 25 sets of data using some kind of minimization routine would have taken far too much computer time, so the fits were performed in a more informal manner. First, group reaction time distributions (Ratcliff, 1979; Thomas & Ross, 1980) were obtained for each subject for each condition across sessions. Ratcliff (1979) has shown that this method introduces no serious bias into the determination of the average reaction time distribution. Even if there are significant practice effects across sessions, the distributions across conditions will maintain their relative shapes (a condition with a long tail will still have a long tail relative to other conditions). In fact, eliminating the first session from analyses had no effect on the distribution parameters except that everything was a few milliseconds faster.

These group distributions (one for each subject in each condition) were then fitted by the convolution model shown by Ratcliff and Murdock (1976) to describe the shape of empirical reaction time distributions. The convolution model (the convolution of an exponential and a normal distribution) uses three parameters to describe distribution shape. These parameters are the exponential parameter τ (the mean of the exponential) and μ and σ (the mean and standard deviation of the normal). Roughly, μ describes the behavior of the mode or fastest responses and τ describes the behavior of the tail of the distribution. Thus, fitting the convolution model to the group distribution for each condition gave 25 sets of three parameters. These parameters were fitted by the retrieval model. Specifically, the equations for the theoretical reaction time distributions and error rates were used to generate predicted empirical reaction time distributions and error rates. Convolution parameters for these predicted distributions were then compared to the convolution parameters obtained from the data. Thus, the convolution model serves as a meeting point for the theoretical and empirical distributions. This procedure does not introduce serious bias into the fitting process (Ratcliff, 1978, 1979). Although the empirical and theoretical distributions are fitted to each other through the convolution model, the fits that are presented in Figures 3 and 4 are direct from retrieval model to data; the convolution model parameters are not shown. This is a critical point. If the fitting procedure described above introduced any bias, then this bias would show up as a poor fit of the diffusion model to the data in Figure 4.

The fits of the model to the data were achieved by adjusting parameters to obtain the best fits to accuracy and reaction time distributions for the 25 *different* conditions and the *same* condition. The parameters s (variability in drift in the diffusion process) and η (variability in relatedness) were kept constant at the values used by Ratcliff (1978), .08 and .18. Adjusting these parameters over a small range (30%) did not affect the fits but did alter other parameter estimates (though not the relative values). The results of these fits and the data are shown in Figures 3 and 4.

Received May 15, 1981
Revision received July 21, 1981 ■