

## Distinctiveness and serial position effects in recognition

IAN NEATH

Purdue University, West Lafayette, Indiana

Digitized photographs of snowflakes were presented for a recognition test after retention intervals of varying durations. While overall accuracy and discrimination remained constant, as the retention interval increased, primacy increased from chance to reliably better than chance while recency decreased to chance levels. A variation of Murdock's (1960) distinctiveness model accounted for the changing primacy and recency effects observed in both between- and within-subjects designs. The generality of the model was examined in two different paradigms: lexical access during sentence processing, and free recall in the continual distractor paradigm. In both cases, the model made accurate qualitative predictions for both latency and accuracy measures.

Recently, there has been a flurry of articles concerned with the idea that recollection of an item, in a variety of different tasks and paradigms, depends critically on its distinctiveness—that is, on the degree to which a given item stands out among the other items in the set (e.g., Crowder & Neath, 1991; Gardiner & Hampton, 1988; Glenberg & Swanson, 1986; Johnson, 1991; McDaniel & Einstein, 1986; Neath & Crowder, 1990; Schmidt, 1991). The idea, of course, is not new: first formulated verbally by Koffka (1935), and then quantitatively by Murdock (1960), the central idea can be traced back to Aristotle (Burnham, 1888). In this paper, I demonstrate that a variation of Murdock's model of distinctiveness can account for serial position effects observed in recognition memory and that this model can be applied to other paradigms as well.

Murdock (1960) defined distinctiveness as the extent to which a given stimulus "stands out" (p. 17) from other stimuli and noted that "the concept of distinctiveness refers to the relationship between a given stimulus and one or more comparison stimuli, and if there are no comparison stimuli the concept of distinctiveness is simply not applicable" (p. 21). By means of an analogy with visual perception, one can demonstrate that just as the insertion of space between items in the visual field should make them more distinctive *spatially*, the addition of an interval between items in a list will make them more distinctive *temporally* (Glenberg & Swanson, 1986; Neath & Crowder, 1990). An item can also be made more dis-

tinct by the manipulation of physical properties (Calkins, 1894; von Restorff, 1933), or, presumably, of any other dimension of interest. This view of distinctiveness can predict the general bowed shape of the serial position function (Murdock, 1960; Neath, in press): the most recent to-be-remembered items in a uniformly spaced list are more distinct, temporally, than middle items, and the first few items also are more distinct, temporally, than middle items. Thus, both primacy and recency effects, the enhanced recall of early and late items as opposed to middle items, can be readily predicted and explained by this view.

The recency effect is one of the most ubiquitous findings in research on memory (see Neath, in press). Perhaps because of this, negative recency effects—the finding that the items at the end of a list do not always enjoy an advantage over earlier items—have figured prominently in theoretical reasoning. They also appear, at first glance, to contradict what a theory based on distinctiveness would predict. How can the final item, the most recently presented item, be less distinct than older items?

Craik (1970) reported a pronounced negative recency effect when subjects were given a surprise final free recall test of all the lists in the session. He observed that although end-of-the-list items were recalled well on the individual, immediate trials, these items were recalled most poorly during the final test. Madigan and McCabe (1971) also demonstrated this effect with paired-associate probe testing, showing that an initial (or immediate) test was not required for the negative recency effect to be observed in the subsequent (or delayed) test.

Cornell and Bergstrom (1983) studied serial position effects in infants' face recognition. Although they presented the same items repeatedly, they generally observed evidence of primacy and recency effects. Of most interest is that they did not observe a recency effect for the last face with a 5-min retention interval. In contrast, Potter and Levy (1969) found better recognition of the final item and worse recognition of the initial item with a 0-sec retention interval.

Some of these results were presented at the 33rd annual meeting of the Psychonomic Society, St. Louis, November 1992. I wish to thank Vincent Biedron, Michelle Hrabovsky, and Victoria Smejkal for running the subjects in Experiment 1, and Kimberly Croley for running the subjects in Experiment 2. In addition, I thank Robert G. Crowder and Aimée M. Surprenant for helpful and insightful discussions. Correspondence concerning this article may be addressed to I. Neath, 1364 Psychological Sciences Building, Purdue University, West Lafayette, IN 47907-1364 (e-mail: neath@brazil.psych.purdue.edu).

Wright, Santiago, Sands, Kendrick, and Cook (1985) demonstrated the emergence of negative recency in a probe recognition paradigm. They presented a series of kaleidoscopic pictures to subjects; each series of four pictures was followed by a recognition judgment after a retention interval of varying duration. Although the overall level of performance remained fairly constant with different retention intervals, the size of the primacy and recency effects changed. With short retention intervals, there was no primacy effect and a substantial recency effect; but as the retention interval increased, the magnitude of the primacy effect increased while that of the recency effect decreased. Moreover, the same pattern was observed in both pigeons and monkeys.

One can draw an analogy between an immediate test of free recall and its pronounced recency effect to the short retention interval conditions of Wright et al.'s (1985) study, and between a delayed test of free recall and the longer retention interval conditions (see Bjork & Bjork, 1992, for a discussion of this trend in other areas).<sup>1</sup> The traditional modal model, both Wright et al. (1985) and Madigan and McCabe (1971) argued, cannot explain this pattern of results without substantial modification. Whereas the decrease in recency can be explained by assuming that there are no longer any items left in short-term memory, the increase in primacy is more problematic. Unless one adds another mechanism or process, it is difficult for the model to explain why an item can be successfully recalled from long-term memory after a 5-sec, 10-sec, or 1-min pause, but not after a 1-sec pause. A consolidation argument might work, but such a theory has not withstood empirical tests very successfully (Crowder, 1982).

The purpose of this paper is, first, to replicate Wright et al.'s (1985) results and, second, to present a distinctiveness model that predicts this pattern. The stimuli used by Wright et al. and those used here are intended to reduce verbal mediation to a minimum. For Wright et al., this was necessary so that cross-species comparisons could be facilitated; for the distinctiveness theory at this stage of development, it is necessary because verbal mediation could lead to re-presentations of a given stimulus at positions in the list other than the original. Because distinctiveness is based on an item's temporal (rather than simply serial or ordinal) position in the list, it is necessary to use stimuli that have been shown to be unaffected by verbal mediation even when explicit verbal associations are required by the task (see, e.g., Goldstein & Chance, 1970). It should also be noted that while primacy and recency effects can be observed when subjects rehearse, they can also be observed when there is little or no evidence of verbal rehearsal (Greene, 1986; Neath, in press).

## EXPERIMENT 1

Experiment 1 is modeled on the study conducted by Wright et al. (1985). They observed that as the retention interval increased, recency decreased and primacy increased in a recognition task. The main differences between that experiment and Experiment 1 are that here

(1) no feedback was given to subjects about their accuracy; (2) a between-subjects design was used such that for a given subject, the duration of the retention interval was constant; (3) confidence ratings were collected; and (4) each list contained only novel stimuli never before seen by the subjects. Also, only humans were tested.

## Method

**Subjects.** One hundred and twenty Purdue University undergraduates volunteered to participate in exchange for credit in introductory psychology courses.

**Design and Materials.** Upon arrival in the laboratory, the subjects were arbitrarily assigned to one of six groups. The sole difference between the six groups was the duration of the retention interval. All stimuli were presented on, and responses recorded by, an Apple Macintosh LC computer. The stimuli were 593 photographs of snowflakes from Bentley and Humphreys (1962). The snowflakes were digitized in such a way that they were displayed in 16 levels of gray against a black background and occupied approximately 7 cm<sup>2</sup> at the center of the CRT. On each trial, four novel snowflakes, never before seen by the subject, were displayed one at a time for 1 sec each. There was a 1-sec interpresentation interval (IPI), but the duration of the retention interval (RI) was 0, 1, 2, 5, 7.5, or 10 sec, depending on the group. After the RI, a fifth flake was displayed, with subjects responding either "old" if it was one of the four flakes seen in the immediately preceding list, or "new" if it was a flake that had not been seen before. The subjects were informed that each list comprised novel snowflakes and that the fifth flake was either a novel flake or one of the four. The subjects made each response by clicking once with the mouse on an appropriately labeled button. In addition, the subjects indicated their confidence that they had answered the recognition question correctly on a scale from 1 (*a guess*) to 5 (*absolutely positive*). They had 8 sec in which to respond. Three out of every seven trials was a *new* trial; of the remaining four, each serial position was tested once. Order of the trials was random. The subjects received a total of 84 lists, preceded by two practice trials.<sup>2</sup> No feedback was given to each subject about performance level.

**Procedure.** The subjects were told that we were interested in how accurately they could remember which snowflakes they had seen before. It was emphasized that each four-item list contained novel flakes, and that the test snowflake was either novel or from the immediately preceding list. During the instructions, the subjects practiced responding with the mouse, and no subject reported any difficulty with this method of responding. Two practice trials were conducted with the experimenter watching. The subjects were tested individually, were given a rest period approximately half-way through the experiment, and were allowed to adjust the monitor to a comfortable level of brightness.

## Results and Discussion

Figure 1 shows the proportion of correct responses for Experiment 1, as a function of RI and serial position. Chance performance was 50%. First, the main data of Wright et al. (1985) were replicated: with short RIs, there is little primacy and large recency. As the RI increases, however, primacy increases and recency decreases. Overall level of performance was comparable in each of the six conditions, with the proportion of correct responses, for RIs of 0-10 sec, being 0.63, 0.66, 0.65, 0.67, 0.63, and 0.62, respectively. The corresponding false alarm rates were 0.24, 0.22, 0.21, 0.18, 0.17, and 0.17.

A 6 (retention interval)  $\times$  4 (serial position) analysis of variance (ANOVA) on the proportion of items correctly recognized as old (hits) revealed a reliable main effect

of serial position [ $F(3,342) = 39.04$ ,  $MS_e = 0.03$ ,  $p < .01$ ] but not of retention interval [ $F(5,114) < 1$ ]. The interaction was reliable [ $F(15,342) = 3.78$ ,  $MS_e = 0.03$ ,  $p < .01$ ]. Thus, as is apparent in Figure 1, there was no overall loss of accuracy as the retention interval increased, but there was a change in which items would be remembered accurately. Performance on items presented at the first position was no better than chance when the retention interval was 0 sec; however, with a 10-sec RI, it was substantially better than chance [ $t(19) = 4.36$ ,  $p < .01$ ].

Unlike in Wright et al.'s (1985) study, the subjects in the present experiment were not given feedback. They were, however, asked to rate the confidence they had in their answer: a rating of 1 would indicate that they thought they were guessing, and a rating of 5 would indicate that they were absolutely certain they had made the correct recognition judgment. Figure 2 shows the confidence data for hits as a function of serial position and RI for Experiment 1.

The confidence ratings generally bear little resemblance to the accuracy data. Regardless of the retention inter-

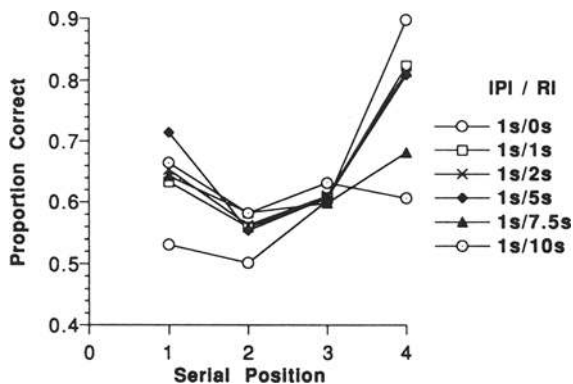


Figure 1. Proportion of snowflakes correctly recognized, as a function of the duration of the retention interval and serial position in Experiment 1.

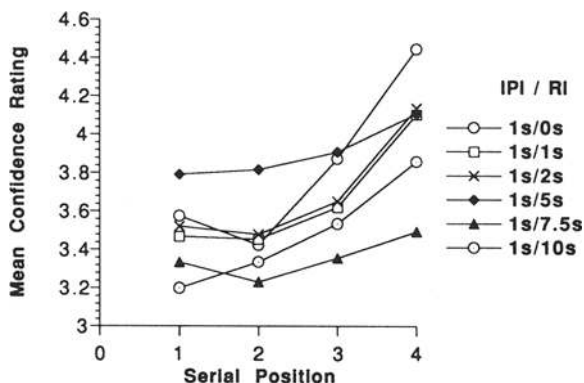


Figure 2. Mean confidence ratings for hits, as a function of the duration of the retention interval and serial position in Experiment 1. 1 = guess, 5 = positive.

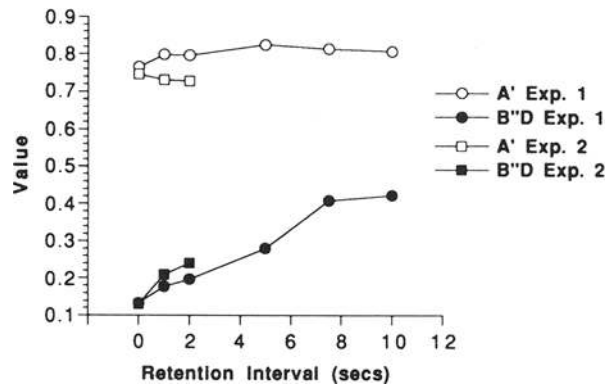


Figure 3. Discrimination ( $A'$ ) and Bias ( $B'D$ ), as a function of the duration of the retention interval in Experiments 1 and 2.

val, the subjects generally had higher confidence for the last item in the list than the first, even though this did not always reflect their true accuracy. The overall confidence ratings for hits, in the conditions 0-sec to 10-sec, respectively, were 3.83, 3.66, 3.70, 3.91, 3.36, and 3.48. The corresponding ratings for misses were 3.77, 3.63, 3.56, 3.38, 3.06, and 3.46. An analysis of these data revealed a reliable effect of retention interval [ $F(5,114) = 2.25$ ,  $MS_e = 0.73$ ,  $p = .05$ ] and a marginal difference between ratings for hits as opposed to ratings for misses [means of 3.65 and 3.48,  $F(1,114) = 3.31$ ,  $MS_e = 0.68$ ,  $p < .10$ ]. The interaction was not reliable ( $F < 1$ ). Of interest are the elevated ratings of subjects in the 5-sec RI condition, in comparison with the ratings of subjects in the other groups; recognition performance in this condition also varied the most from what was expected by the model.

The overall correlation between the confidence ratings and proportion of hits as a function of serial position, although not large, was significant [ $r(118) = 0.25$ ,  $p < .05$ ]. However, further analyses showed that as the retention interval increased, the accuracy of the confidence ratings became worse: For the retention intervals from 0 sec to 10 sec, the correlations were 0.44, 0.33, 0.35, 0.12, 0.13, and 0.00, respectively. This corresponds to the fact that the subjects were more likely to give a higher confidence rating for the item in the final serial position as opposed to earlier positions, even though the proportion of hits did not necessarily follow this pattern.

Recognition data can also be analyzed according to discriminability and bias. Figure 3 represents the results of an  $A'$  and  $B'D$  analysis for Experiment 1 (as well as for Experiment 2).  $A'$  is a nonparametric analog of  $d'$  (Pollock & Norman, 1964) and has been shown to be highly correlated with  $d'$  (Snodgrass, Volvovitz, & Walfish, 1972). It ranges from 0 to 1, with 0.5 reflecting chance, and has the advantages of allowing analysis of subjects who may have hit or false alarm rates of 0 and of not requiring all the assumptions of  $d'$  concerning variance.  $B'D$  ranges from  $-1$  to  $+1$ , with 0 indicating no bias; a posi-

tive number indicates a conservative bias (see Donaldson, 1992).

Of principal interest is the fact that  $A'$  remains constant across retention intervals and is well above chance levels; discriminability is just as good with a 0-sec RI as with a 10-sec RI [ $F(5,114) = 1.41$ ,  $MS_e = 0.01$ ,  $p > .20$ ]. The changes in the serial position function, then, cannot be due to changes in discriminability. There is, however, a trend toward more conservative bias as the RI increases [ $F(5,114) = 2.46$ ,  $MS_e = 0.12$ ,  $p < .05$ ]. This is reflected in the slightly increasing proportion of correct rejections: for retention interval durations of 0-10, the mean proportions of correct rejections were 0.76, 0.79, 0.79, 0.82, 0.83, and 0.83, respectively. Although a change in bias is reflected in the correct rejection data, it is not sufficient to explain the change in the shape of the serial position function.

## EXPERIMENT 2

Experiment 1 replicated the main findings of Wright et al. (1985): as the retention interval increases, recency decreases but primacy increases while the overall level of performance remains constant. Wright et al. used a within-subjects design, but the retention intervals were blocked and the stimuli were partially repeated. Experiment 2 was similar to Experiment 1, except that each subject experienced each retention interval, with the order of the conditions occurring randomly.

### Method

**Subjects.** Thirty-eight Purdue University undergraduates participated in exchange for credit in introductory psychology courses.

**Design, Materials, and Procedure.** The materials, design, and procedure were identical to those of Experiment 1, except for the following: There were three retention intervals (0, 1, or 2 sec), and each subject experienced all three. There was a total of 105 trials, 35 in each condition. Of these 35 trials, 15 were *new* trials, and the remaining 20 were equally divided among the four serial positions. The order of the trials and conditions was random and different for each subject.

### Results and Discussion

Figure 4 shows the proportion of correct responses (hits) for Experiment 2 as a function of RI and serial position. First, the findings of Experiment 1 were replicated, but now in a within-subjects design: with short RIs, there is little primacy and large recency. As the RI increases, however, primacy increases and recency decreases. A 3 (retention interval)  $\times$  4 (serial position) ANOVA revealed a reliable main effect of position [ $F(3,111) = 31.71$ ,  $MS_e = 0.06$ ,  $p < .01$ ] but not of retention interval [ $F(2,74) = 1.39$ ,  $MS_e = 0.07$ ,  $p > .20$ ]. The interaction was just reliable [ $F(6,222) = 2.10$ ,  $MS_e = 0.06$ ,  $p = .05$ ]. The changing serial position function, then, is not due to developing specific strategies to cope with one retention interval, nor is it due to experiencing several different intervals. With a 0-sec retention interval, subjects did not perform better than chance at the first serial position ( $t < 1$ ), whereas with a 2-sec retention interval,

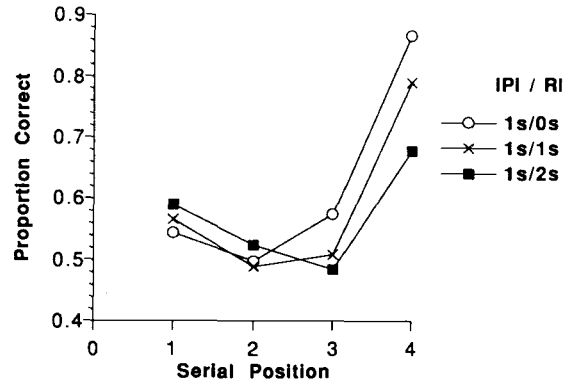


Figure 4. Proportion of snowflakes correctly recognized, as a function of the duration of the retention interval and serial position in Experiment 2.

subjects performed reliably better than chance [ $t(37) = 2.18$ ,  $p < .05$ ]. The increase in primacy is very small, but this is not surprising, given only a 3-sec difference in the duration of the retention intervals.

Figure 5 shows the confidence ratings of the subjects in Experiment 2. Unlike the data from Experiment 1, these ratings bear a more marked resemblance to accuracy: the subjects were most confident at Serial Position 5 for the 0-sec RI condition and least confident for the 2-sec RI condition; they were least confident at Serial Position 1 for the 0-sec RI condition. The overall confidence ratings for hits were 3.76, 3.72, and 3.62 for the 0-sec, 1-sec, and 2-sec RI conditions, respectively; the corresponding ratings for misses were 3.43, 3.46, and 3.50. There was no difference in overall ratings as a function of the duration of the RI ( $F < 1$ ), but there was a reliable difference between the mean rating for a hit (3.70) and the mean rating for a miss [3.46;  $F(1,37) = 5.01$ ,  $MS_e = 0.63$ ,  $p < .05$ ]. The interaction was not reliable ( $F < 1$ ). The correlation between the confidence ratings and proportion of hits as a function of serial position, although not large, was significant [ $r(36) = 0.43$ ,  $p <$

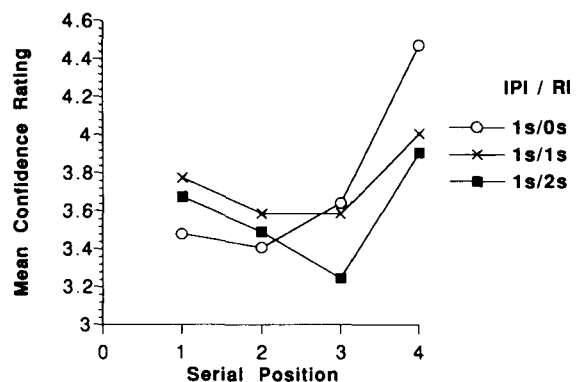


Figure 5. Mean confidence ratings for hits, as a function of the duration of the retention interval and serial position in Experiment 2. 1 = *guess*, 5 = *positive*.

.01]. As in Experiment 1, the correlation between confidence and performance decreased across retention intervals (for RIs of 0, 1, and 2 sec,  $r = 0.50, 0.47,$  and  $0.33,$  respectively).

Figure 3 shows the results of an  $A'$  and  $B''_D$  analysis; as in Experiment 1, there was no change in discriminability as a function of retention interval [ $F(2,74) < 1$ ]. Although the figure suggests that the subjects became more conservative in their bias as the duration of the retention interval increased, this was not supported statistically [ $F(2,74) < 1$ ]. The average proportions of correct rejections for the three conditions were 0.76, 0.77, and 0.77 for the 0-sec, 1-sec, and 2-sec conditions, respectively. The average proportions of hits were 0.62, 0.59, and 0.57 for the same conditions.

Experiments 1 and 2 both replicated the main findings of Wright et al. (1985), despite the variations in design, procedure, and stimuli: with a short retention interval, performance on the first item is no different from chance levels but there is substantial recency; as the retention interval increases, however, the pattern reverses, with primacy increasing to above-chance levels and recency decreasing. The fact that the same pattern is observed in both between- and within-subjects designs rules out specific strategies arising from contingencies peculiar to the design.

### THE MODEL

Within the field of human cognition in general, and with human memory in particular, there is an increasing awareness that verbal models are no longer sufficient to accurately describe our current knowledge and understanding of mental processes (see, e.g., Hintzman, 1991). The advantages of formal models are well known: (1) formal models can be easily and unambiguously described; (2) they usually state more explicitly the assumptions about and relationships between the concepts involved; and (3) their predictions can be precisely stated.

At the same time, many researchers warn against developing formal models that are too complex (e.g., Broadbent, 1987). The goal in this section of the paper, therefore, is to present a relatively simple model of serial position effects in recognition memory that preserves the advantages of precise statement and that makes quantitative predictions for the home domain and (precise) qualitative predictions elsewhere. For example, intuition suggests that the final item of a list should be more distinctive than earlier items; the formal definition of distinctiveness presented below demonstrates that this intuition can sometimes be incorrect.

The model, based on Murdock (1960), compares memory with visual perception: the most recent to-be-remembered items in the list, temporally, are the most distinct, just as in visual perception the closest items, spatially, are the most distinct. Although that which follows is based on distinctiveness along a temporal dimension, in theory this model could be applied to any stimuli lying along any ordered dimension (e.g., physical or semantic dimensions;

see Neath, in press, for a more complete discussion). To model other forms of distinctiveness, it would be necessary to create a ratio scale along the dimension of interest and to calculate the distinctiveness of each item according to that scale. This appears to be more straightforward for physical dimensions, but it should, in principle, be applicable to dimensions such as semantic distinctiveness. Once such a scale is constructed, the application of the model is the same as it is for temporal distinctiveness.

One difference between Murdock's formulation and the present model is that the durations of the IPI and the RI, rather than the serial (ordinal) position alone, play a critical role in determining the distinctiveness of an item.<sup>3</sup> The temporal position of an item is calculated by giving each item an initial value of 1. The final item in the list is multiplied by the duration of the RI; all other items are multiplied by the duration of the IPI and then added to the sum of the value of the following item. For example, if the RI is 2 sec and each IPI is 3 sec, the serial/temporal values ( $s$ ) of a four-item list, beginning with the first item, would be 11, 8, 5, and 2. The final item in the list is a distance of 2 from the recaller, and the first item is a distance of 11 from the recaller. These  $s$  values then undergo a log transformation, a process based on Helson's (1964) adaptation level theory, yielding the preliminary discriminability of that position ( $d$ ).

A value of each item's distinctiveness ( $\delta$ ) is calculated according to the following formulae, where  $n$  is the number of items in the list and  $k$  is the serial position of the current item. Equation 1 applies when  $k$  is either 1 or  $n$ , whereas Equation 2 applies to the remaining items. The difference between the two equations is that the middle items are adjusted by using a constant ( $c$ ) to avoid symmetries: without this correction, for example, Items 2 and 3 of a four-item list would have the same predicted value.

$$\delta_k = \sum_{j=1}^n \sqrt{(d_k - d_j)^2} \quad (1)$$

$$\delta_k = \frac{c \sum_{j=1}^n \sqrt{(d_k - d_j)^2} + 2 \left[ \sum_{j=1}^n \sqrt{(d_{k-1} - d_j)^2} + \sum_{j=1}^n \sqrt{(d_{k+1} - d_j)^2} \right]}{c + 2} \quad (2)$$

Finally, the distinctiveness values from Equations 1 and 2 are normalized, giving the predicted performance level. Normalizing is appropriate here, because we are concerned with the distinctiveness of an item relative to all other items in the set (see Murdock, 1960).

#### Fit of the Model

Figures 6 and 7 display the predictions of the model and the normalized data from Experiments 1 and 2, respectively. Predictions for a given set of IPI and RI durations are the same, regardless of the experimental design. All of the parameters are determined by the duration of the IPI (always 1 sec in these experiments) and the RI (varying from 0 to 10 sec, depending on the condition). The only

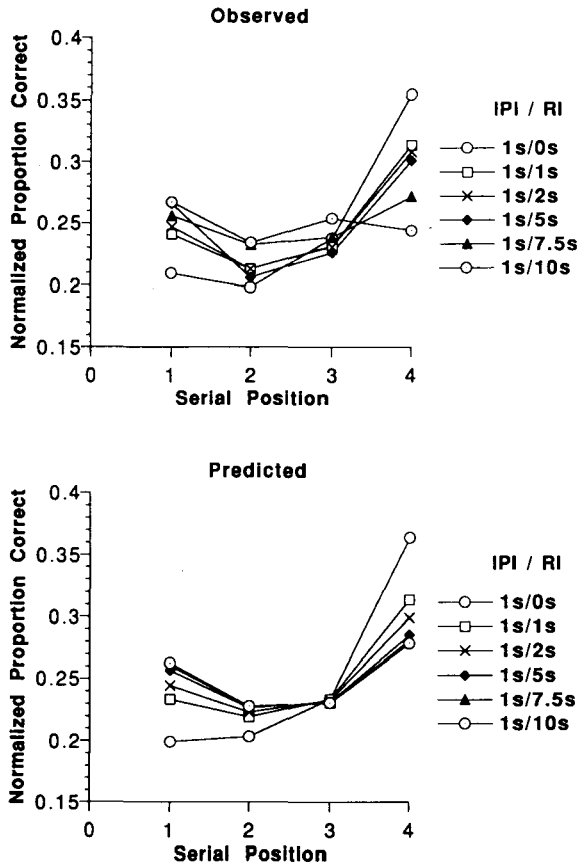


Figure 6. Normalized proportion of snowflakes correctly recognized from Experiment 1, as a function of condition and serial position (top panel) and the predictions of the distinctiveness model (bottom panel).

exceptions are that for the 0-sec condition a value of 0.1 was used—because of the log transformation—and that the constant  $c$  was arbitrarily set to 9. Note that  $c$  was the same value for all serial positions, for all conditions, and for both experiments: once set,  $c$  does not change.

For the data from Experiment 1, when normalized, the model accounted for 90.09% of the variance (RMSD = 0.012); for the data from Experiment 2, when normalized, the model accounted for 91.93% of the variance (RMSD = 0.013). The parameters for both fits either were identical (viz.,  $c$ ) or depended solely on the duration of the IPI and the RI. The model is clearly capturing the important aspects of the data: as the retention interval increases, recency declines and primacy increases. The rank ordering of the conditions for both the first and the last items is predicted accurately. Note that the definition of distinctiveness offered here can predict less distinctiveness for the final item in a series than for earlier items.

**Generality of the Model**

If the model is capturing something fundamental to memory, it should make accurate predictions in paradigms other than probe recognition. One effect reminiscent of

the changing serial position function occurs in the study of lexical access in sentence processing.

Lexical access can be inferred from the time to respond to simple verification questions that follow presentation of a sentence. Gernsbacher, Hargreaves, and Beeman (1989) showed that for sentences such as “Tina gathered the kindling as Lisa set up the tent,” there is greater lexical access (i.e., there are faster response times) to the word “Lisa” than to “Tina” with a 0-msec RI, but that with a 2,000-msec RI, the reverse is true. Figure 8 shows the data reported by Gernsbacher et al. (1989) and the predictions of the model.

The predictions of the model were made by assuming that the first-mentioned participant appeared in the first serial position and the second-mentioned participant appeared in the sixth serial position. The parameter  $c$  was set to 3 (for all conditions), and the values of the IPI and RI were determined from the actual experimental conditions. Gernsbacher et al. (1989) held the IPI constant, and the on-screen duration of each word was approximately constant. If one assumes that reaction time is inversely related to distinctiveness, the model can predict, qualitatively, Gernsbacher et al.’s data. Similar assumptions con-

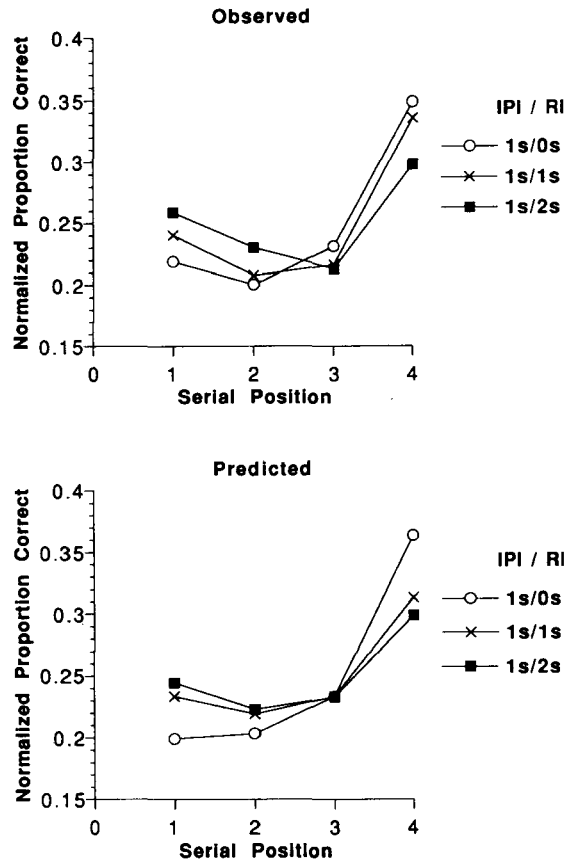


Figure 7. Normalized proportion of snowflakes correctly recognized from Experiment 2, as a function of condition and serial position (top panel) and the predictions of the distinctiveness model (bottom panel).

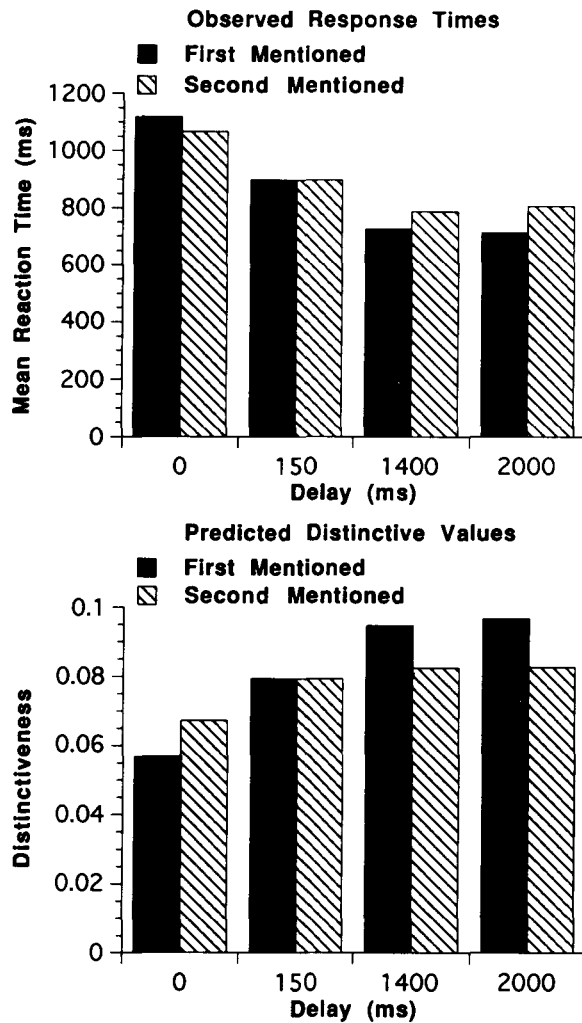


Figure 8. The predicted distinctive ( $d$ ) values (bottom panel), according to the model, for the data presented by Gernsbacher et al. (1989) (top panel). From "Building and Accessing Clausal Representations: The Advantage of First Mention Versus the Advantage of Clause Recency" by M. A. Gernsbacher, D. J. Hargreaves, and M. Beeman, 1989, *Journal of Memory and Language*, 28, p. 742. Copyright 1989 by Academic Press. Adapted by permission.

cerning the relationship between accuracy and latency measures have been made before (e.g., as regards strength theory; see Murdock, 1985). The suggestion is that the more distinctive an item, the faster the subject will be able to determine its identity and to make an appropriate recognition judgment. This explanation suggests that lexical access effects are due, in part, to the serial position—or more accurately, to the temporal position—of the presented items.

There has been much research relating various reading measures with various memory measures (e.g., Baddeley, 1992). This model suggests that some lexical access and comprehension effects may be due, in part, to the distinctiveness of the items in memory. The exact nature of

the function that relates latency and accuracy, although probably exponential (see, e.g., Murdock, 1974), still needs to be determined. A strong prediction of the model, then, is that latency measures in probe recognition should be inversely related to the distinctiveness values.

#### Extensions of the Model

One potential weakness of the model is the fact that the predictions are normalized. Normalization is not always necessary; the reason for normalizing is to transform the distinctive values into a particular scale so that comparisons across conditions will be meaningful. An alternate method is to divide the distinctive values ( $\delta$ ) by the sum of the log values ( $d$ ), which restricts the values to a range from 0 to 1. This method can be seen as taking into account the total amount of distinctiveness; for example, one would want the model to predict little or no memory, regardless of the durations of the IPI, if the RI were substantially longer (e.g., 50 years). This method of scaling takes this reflection into consideration and enables the model to predict the relative number of items recalled in different paradigms.

Neath and Crowder (1990, Experiment 3) demonstrated that altering the duration of the IPIs within a list had consistent effects on the shape of the serial position curve. Studying free recall in the continual distractor paradigm, they showed that the relative level of recall of a given item was successfully predicted using only the time separating item  $n$  from item  $n-1$  and the time until item  $n$  was recalled. This so-called extended ratio rule states that the larger this ratio, the better the overall level of recall at that position, regardless of the actual time scale. In addition, the median ratio of the list as a whole should predict the overall level of performance. Neath and Crowder used three decreasing schedules, in which the duration of the IPI decreased from 8 to 4 to 2 to 0 sec (Decreasing 1), from 8 to 4.5 to 2.5 to 1.5 sec (Decreasing 2), or from 4 to 2.5 to 1.5 to 1 sec (Decreasing 3). The increasing condition was the reverse of Decreasing 1, and the constant condition had a uniform 2-sec IPI. The RI was always constant at 2 sec. The extended ratio rule predicted that overall performance should be, from best to worst, Decreasing 1, Decreasing 2, Decreasing 3, constant, and increasing, because the median ratios were 1.0, 0.75, 0.53, 0.42, and 0.27, respectively. The numerical qualitative predictions from the present model are 0.378, 0.356, 0.307, 0.267, 0.256 for the same conditions, thus making the same rank-order predictions. Note this ordering, despite the fact that the total amounts of distraction for these conditions were 56, 66, 36, 32, and 56 sec, respectively.

A replication, with different stimuli and different time scales, basically mirrored the preceding results (Neath & Crowder, 1993). The retention interval was always 1 sec. The four conditions, or schedules, were: (1) constant, with 1 sec between each letter pair, yielding a median ratio of 0.42; (2) increasing, with successively longer durations between the letter pairs, going from 1 to 2 to 4 to 8 sec,

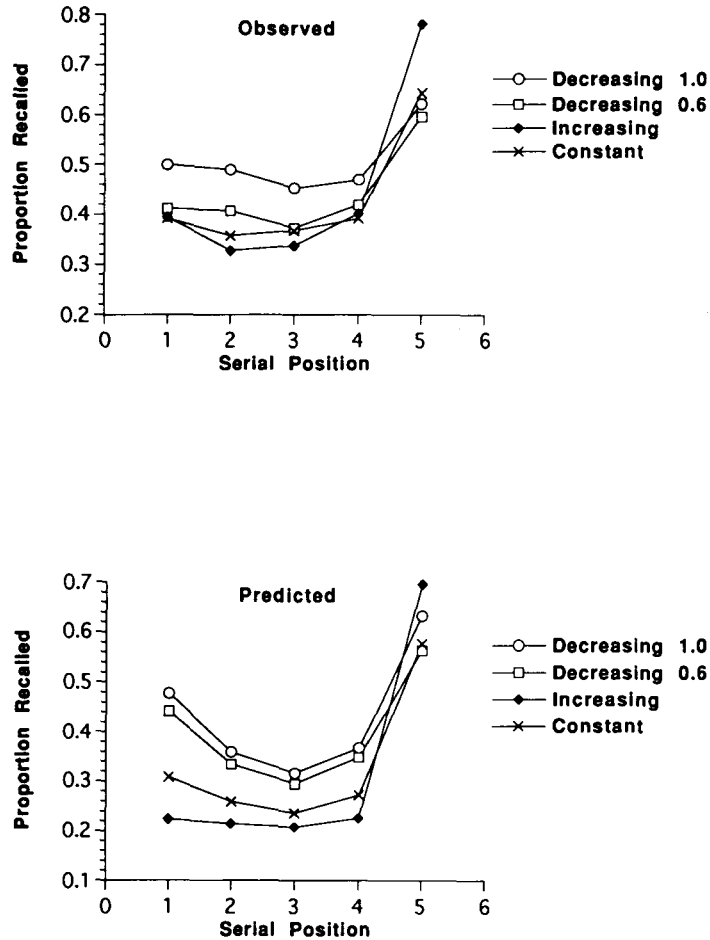


Figure 9. The proportion of letters freely recalled, as a function of serial position and condition from Neath and Crowder (1993)—demonstrating that performance depends on the median ratio rather than absolute temporal duration (top panel). The bottom panel shows the predictions of the distinctiveness model. The data in these panels were not normalized.

yielding a median ratio of 0.32; (3) decreasing 1.0, which was the reverse of the increasing condition, yielding a median ratio of 1.0; and (4) decreasing 0.6, where the duration of the IPIs systematically decreased from 2.5, to 1.5, to 1, to 0.5 sec, yielding a median ratio of 0.61. Including the retention interval, the total duration of distractor activity for each schedule was 16 sec in decreasing 1.0, 6.5 sec in decreasing 0.6, 16 sec in the increasing, and 5 sec in the constant lists.

Figure 9 displays the observed and predicted results. Only two parameters were needed to obtain a fairly accurate fit: the value of the RI was multiplied by 4 and the value of each IPI was multiplied by 17. Of more interest, however, was the observation that even if these parameters are not used, the model produces the correct ordering of the conditions. The present model, then, produces the same general pattern of results as does the extended ratio rule: Recall will be better, the larger the ratio; it is the ratio of times that is important, not the absolute amount of time, and the distinctiveness model captures this important distinction. One advantage of the present

conception over the extended ratio rule is that it allows predictions to be made for the first item in a list.

### Evaluation

The model has three main problems that appear inter-related: it is restricted to fitting normalized data when making quantitative predictions; it becomes increasingly inaccurate for longer RIs; and it uses a log transformation rather than a power function.

Fitting normalized data restricts the possible range of the data and may artificially inflate the performance of the model. On the other hand, the definition of distinctiveness used in the model emphasizes the fact that an item is distinct only in its relationship to the other items in the ensemble, and that the concept of distinctiveness does not apply to items in isolation. Although normalizing may not be unreasonable, then, it may limit the predictions that can be made. An additional side effect is that the model becomes increasingly inaccurate for longer RIs. Although the model *does* predict worse recall at the final position for longer RIs, the differences predicted are



smaller than those actually observed. An alternate scaling method, discussed above, partially overcame this problem, but rendered quantitative fits untenable. The third problem is that both versions rely on a log transformation when most psychophysicists regard power functions as a more appropriate transformation. Future investigations will explore this possibility.

## CONCLUSIONS

In two experiments, the main findings of Wright et al. (1985) were replicated, demonstrating that the shape of the serial function changes in predictable ways with increases in retention interval. A distinctiveness model, based on one proposed by Murdock (1960), accounted quantitatively for the data. Although the model was found to have three weaknesses, it also has distinct strengths. First, the model has minimal free parameters: all but one of the parameters are given in the experimental design, and the same parameters fit data from both between- and within-subjects designs. Second, the model made accurate qualitative predictions in areas outside the domain that it was originally developed for: it predicted the pattern of lexical access reaction times in text processing and the results of the extended ratio rule in free recall in the continual distractor paradigm. Third, because the model requires no peculiarly verbal skills on the part of the rememberer, it is potentially applicable to animals other than humans.

The model provides a simple, precise definition of *distinctiveness*—a potentially vague and ambiguous notion—and demonstrates that such a concept can predict and explain many serial position effects in both recognition and recall. As such, the model can be seen as one way of specifying in more detail a component of the more general theories of distinctiveness.

## REFERENCES

- BADDELEY, A. D. (1992). Working memory. *Science*, **255**, 556-559.
- BENTLEY, W. A., & HUMPHREYS, W. J. (1962). *Snow crystals*. New York: Dover.
- BJORK, R. A., & BJORK, E. L. (1992). A new theory of disuse and old theory of stimulus fluctuation. In A. F. Healy, S. M. Kosslyn, & R. M. Shiffrin (Eds.), *From learning processes to cognitive processes: Essays in honor of William K. Estes* (Vol. 2, pp. 35-67). Hillsdale, NJ: Erlbaum.
- BROADBENT, D. (1987). Simple models for experimentable situations. In P. E. Morris (Ed.), *Modeling cognition*. Chichester, UK: Wiley.
- BURNHAM, W. H. (1888). Memory, historically and experimentally considered. *American Journal of Psychology*, **2**, 39-90, 255-270, 431-464, and 566-622.
- CALKINS, M. W. (1894). Association. *Psychological Review*, **1**, 476-483.
- CORNELL, E. H., & BERGSTROM, L. I. (1983). Serial position effects in infants' recognition memory. *Memory & Cognition*, **11**, 494-499.
- CRAIK, F. I. M. (1970). The fate of primary memory items in free recall. *Journal of Verbal Learning & Verbal Behavior*, **9**, 143-148.
- CROWDER, R. G. (1976). *Principles of learning and memory*. Hillsdale, NJ: Erlbaum.
- CROWDER, R. G. (1982). General forgetting theory and the locus of amnesia. In L. S. Cermak (Ed.), *Human memory and amnesia* (pp. 33-41). Hillsdale, NJ: Erlbaum.
- CROWDER, R. G., & NEATH, I. (1991). The microscope metaphor in human memory. In W. E. Hockley & S. Lewandowsky (Eds.), *Relating theory and data: Essays on human memory in honor of Bennet B. Murdock, Jr.* (pp. 111-125). Hillsdale, NJ: Erlbaum.
- DONALDSON, W. (1992). Measuring recognition memory. *Journal of Experimental Psychology: General*, **121**, 275-278.
- GARDINER, J. M., & HAMPTON, J. A. (1988). Item-specific processing and the generation effect: Support for a distinctiveness account. *American Journal of Psychology*, **101**, 495-504.
- GERNSBACHER, M. A., HARGREAVES, D. J., & BEEMAN, M. (1989). Building and accessing clausal representations: The advantage of first mention versus the advantage of clause recency. *Journal of Memory & Language*, **28**, 735-755.
- GLENBERG, A. M., & SWANSON, N. C. (1986). A temporal distinctiveness theory of recency and modality effects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 3-24.
- GOLDSTEIN, A. G., & CHANCE, J. E. (1970). Visual recognition memory for complex configurations. *Perception & Psychophysics*, **9**, 237-241.
- GREENE, R. L. (1986). Sources of recency effects in free recall. *Psychological Bulletin*, **99**, 221-228.
- HELSON, H. (1964). *Adaptation-level theory*. New York: Harper & Row.
- HINTZMAN, D. L. (1991). Why are formal models useful in psychology? In W. E. Hockley & S. Lewandowsky (Eds.), *Relating theory and data: Essays on human memory in honor of Bennet B. Murdock, Jr.* (pp. 39-56). Hillsdale, NJ: Erlbaum.
- JOHNSON, G. J. (1991). A distinctiveness model of serial learning. *Psychological Review*, **98**, 204-217.
- KOFFKA, K. (1935). *Principles of Gestalt psychology*. London: Routledge & Kegan Paul.
- MADIGAN, S. A., & MCCABE, L. (1971). Perfect recall and total forgetting: A problem for models of short-term memory. *Journal of Verbal Learning & Verbal Behavior*, **10**, 101-106.
- MCDANIEL, M. A., & EINSTEIN, G. O. (1986). Bizarre imagery as an effective memory aid: The importance of distinctiveness. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 54-65.
- MURDOCK, B. B., JR. (1960). The distinctiveness of stimuli. *Psychological Review*, **67**, 16-31.
- MURDOCK, B. B., JR. (1974). *Human memory: Theory and data*. Potomac, MD: Erlbaum.
- MURDOCK, B. B., JR. (1985). An analysis of the strength-latency relationship. *Memory & Cognition*, **13**, 511-521.
- NEATH, I. (in press). Contextual and distinctive processes and the serial position function. *Journal of Memory and Language*.
- NEATH, I., & CROWDER, R. G. (1990). Schedules of presentation and temporal distinctiveness in human memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 316-327.
- NEATH, I., & CROWDER, R. G. (1993). *Distinctiveness and very short-term recency effects*. Manuscript submitted for publication.
- POLLACK, I., & NORMAN, D. A. (1964). A non-parametric analysis of recognition experiments. *Psychonomic Science*, **1**, 125-126.
- POTTER, M. C., & LEVY, E. I. (1969). Recognition memory for a rapid sequence of pictures. *Journal of Experimental Psychology*, **81**, 10-15.
- SCHMIDT, S. R. (1991). Can we have a distinctive theory of memory? *Memory & Cognition*, **19**, 523-542.
- SNODGRASS, J. G., VOLVOVITZ, R., & WALFISH, E. R. (1972). Recognition memory for words, pictures, and words + pictures. *Psychonomic Science*, **27**, 345-347.
- VON RESTORFF, H. (1933). Analyse von Vorgängen im Spurenfeld: I. Über die Wirkung von Bereichsbildung im Spurenfeld. *Psychologische Forschung*, **18**, 299-342.
- WRIGHT, A. A., SANTIAGO, H. C., SANDS, S. F., KENDRICK, D. F., & COOK, R. G. (1985). Memory processing of serial lists by pigeons, monkeys, and people. *Science*, **229**, 287-289.

## NOTES

1. It is also possible that spontaneous recovery in verbal learning is another manifestation of this phenomenon. For example, in an A-B

A-D paired-associate design, recall of D items is better after transfer than recall of B items, but later the pattern can reverse (see Crowder, 1976).

2. Because of a programming error, groups 0 and 1 sec received 21 more trials (3 *old* trials for each position and 9 *new* trials) than the remaining groups. The presence of these trials did not affect the results in any detectable way.

3. The model currently ignores both the duration of the to-be-remembered item and the fact that it takes time to recall more than one item. This was done intentionally to simplify formulation at this stage of development.

(Manuscript received December 7, 1992;  
revision accepted for publication March 1, 1993.)