JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR 20, 161–189 (1981)

An Encoding and Retrieval Model of Single-Trial Free Recall

JANET METCALFE AND BENNET B. MURDOCK

University of Toronto

In this paper, the mathematical operations of convolution and correlation will be proposed as mechanisms by which associative encoding and retrieval may occur, and a process model of free recall which conforms to the constraints imposed by these operations will be presented. In the first section of the paper, the operations will be metaphorically described and defined; computer simulations which demonstrate some of the properties will be presented; and restrictions imposed upon a system which uses these operations will be outlined. In the second section, a process model of single-trial free recall that is compatible with the proposed mechanisms is presented. In the third section, a simulation of the model is applied to a variety of experimental situations and the results of the model are compared to observed results. In the final section, some of the strengths and limitations of the model are discussed and contrasted with those of other models of free recall.

THE MECHANISM: CONVOLUTION AND CORRELATION

The approach proposed in this paper to associative encoding and retrieval is not new. As early as 1749, David Hartley presented a theory of memory in which the mechanism for associative encoding consisted of the overlapping of the "vibrations" produced by contiguously presented events. Convolution is very much in the spirit of this early idea, although there was no analog, in Hartley's theory, to correlation—the retrieval mechanism. More recently, a number of theorists (Anderson, 1970; Barret, 1970; Cavanaugh, 1976; Pribram, Nuwer, & Baron, 1974; Van Heerden, 1963) have proposed holographic and neurological models of memory which use the encoding and retrieval operations of convolution and correlation. These operations have also been used in theories of perception (see Murdock, 1979, for a review), and are well known in communication theory.

Convolution and correlation allow the possibility that memory traces may be both distributed and cumulative, that is, they may share the same neural substrate. Memory items are characterized as multidimensional vectors, where the dimensions constituting each item are considered to be the features that make up that item. Items may be spatially or temporally distributed. No particular dimension or feature is critical for identification or memory; but rather the entire pattern determines the identity of a given item. Even when a number of associations formed by convolution are stored in a single cumulative memory vector, so there is no separate representation of each stored trace, retrieval of a single item is then possible using correlation. These characteristics of the proposed mechanism are desirable since there is considerable physiological evidence to support the idea that traces are cumulative, that is, they interact at the synaptic level (Eccles, 1972), and that they are distributed rather than localized (Lashley, 1950; John, 1976).

The present characterization of memory

We would like to thank Eric Eich for the many thoughtful criticisms and insights he shared with us. Peter Liepa and William Batchelder provided mathematical assistance. John Anderson and Peter Polson contributed penetrating reviews of an earlier draft of the manuscript.

This research was facilitated by a National Research Council of Canada (NRC) postgraduate scholarship to the first author and by NRC Grant APA 146 to the second author.

Please send requests for reprints to Janet Metcalfe who is now at the School of Social Sciences, University of California at Irvine, Irvine, Calif. 92717.

0022-5371/81/020161-29$02.00/0

traces is also psychologically plausible. Even when the exact identity of an item cannot be determined, subjects can often provide considerable information about the item (Brown & McNeill, 1966; Hart, 1965). The work of Wickens (1972) suggests that a large number of features may be important for memory. We make the assumption that a large number of features are in fact encoded although we remain neutral with respect to the identity of the features and shall simply consider them to be abstract—probably encompassing both semantic and sensory domains. Although memory traces are often conceptualized as nodes existing in discrete locations in space (Roediger, 1980), there are precedents for representing items as lists of features (Bower, 1967; Flexser & Tulving, 1978; Smith, Shoben, & Rips, 1974; Underwood, 1969; Wickens, 1972).

While the distinction has often been drawn between association by similarity and association by contiguity, only the latter will be considered here. Although it may be possible to extend the theory to similarity effects, these are beyond the scope of the present paper. The discussion of convolution and correlation that follows is directed only at pairwise associations rather than at more complex schemata that may be possible using these operations. The process model which will be outlined shortly uses only pairwise associations.

Metaphors for Convolution and Correlation

Since the operations of convolution and correlation may be unfamiliar to the reader, we shall offer several metaphors to illustrate them. One metaphor is that of simple arithmetic. Suppose two numbers are added to associate them—the operation of addition is similar to convolution. The sum itself is analogous to the associative memory trace. Note that the sum itself does not specify either of the particular two numbers that were added, since the same sum could be obtained in a variety of ways. In order to retrieve one of the original numbers, a "cue" must be provided which is subtracted from the sum. If the cue is precisely one of the original two numbers that were added, the result of subtracting that cue from the sum will be precisely the other number. The addition–subtraction metaphor conveys, correctly, the idea that convolution and correlation are inverse operations. The example just given conveys, incorrectly, the idea that only a single pair of items could be stored in one memory trace. In order to extend the addition–subtraction metaphor to be more consistent with the encoding and retrieval mechanisms proposed here, allow each of the pairs of numbers to add to a value of zero. When this is the case, many pairs can be stored in the same trace and retrieval by subtraction of the cue number will yield the complement. Since the associations may be stored in the same memory trace it is not necessary to assume that a given cue "knows" to which trace it should refer or to postulate any search process.

A second metaphor that conveys some of the flavor of convolution and correlation involves light combination and filtering. Imagine that item A is a blue-green light and that item B is a red light. Shining these two lights together is a metaphor for convolution. The resultant color will be white. In order to retrieve in a manner analogous to correlation, the cue item—the blue-green light, for instance—is used as a filter. The result of passing white light through a blue-green filter is red light. The converse is also true:passing white light through a red filter yields blue-green light. If a second pair of lights also combine to form white light (blue and yellow, for instance), this second pair can be superimposed upon the combination of the first pair and the result will still be white light. Now, however, if a yellow filter is used, blue light will result, even though the red and blue-green lights are still present in the cumulative trace. If the white light (which includes blue and yellow) is passed through a blue-green filter, the result will still be red light, just as before.

A third metaphor that applies to convolution and correlation is that of holography. Since this metaphor has been elaborated in detail elsewhere (see Pribram et al., 1974), we shall not present it here.

Definitions

Convolution and correlation are mathematical operations which can be used to portray the encoding and decoding of to-be-remembered items, given certain assumptions about the nature of the items. The operations can be defined for functions (or waveforms) or for vectors, but in this paper we shall only consider vectors.

We shall denote convolution as * and correlation as #. The convolution of two vectors **F** and **G** is defined as

$$(\mathbf{F} * \mathbf{G})_m = \sum_{i=1}^{n} \mathbf{F}_i \mathbf{G}_{m-i}, \qquad (1)$$

Correlation is defined as

$$(\mathbf{F} \# \mathbf{G})_m = \sum_{i=1}^{n} \mathbf{F}_i \mathbf{G}_{m+i}. \qquad (2)$$

The combination of convolution and correlation which allows the reconstruction or recall of an item is

$$\mathbf{I}_x \# (\mathbf{I}_x * \mathbf{I}_y) = \mathbf{I}_y', \qquad (3)$$

where $\mathbf{I}_x$ and $\mathbf{I}_y$ are two different items and $\mathbf{I}_y'$ is an approximation to $\mathbf{I}_y$. For a more comprehensive discussion of the mathematics and mechanics of convolution and correlation, the reader is referred to Borsellino and Poggio (1973), Bracewell (1965), Lathi (1968), or Murdock (1979).

It is possible to store more than a single convolution in the same associative memory vector (**A**). Suppose, for instance, that item 1 ($\mathbf{I}_1$) and item 2 ($\mathbf{I}_2$) are convolved to form the associative vector **A** at time 1 (i.e., $\mathbf{A}_1$). Another pair of items ($\mathbf{I}_3$ and $\mathbf{I}_4$) may also be convolved and added to **A** to form the associative vector at time 2 (i.e., $\mathbf{A}_2$). Schematically this would look as follows:

$$\begin{aligned} \mathbf{I}_1 * \mathbf{I}_2 &= \mathbf{A}_1 \\ \mathbf{A}_1 + (\mathbf{I}_3 * \mathbf{I}_4) &= \mathbf{A}_2 \\ \mathbf{A}_2 + (\mathbf{I}_5 * \mathbf{I}_6) &= \mathbf{A}_3 \\ &\vdots \\ \mathbf{A}_{n-1} + (\mathbf{I}_j * \mathbf{I}_k) &= \mathbf{A}_n . \end{aligned} \qquad (4)$$

$\mathbf{A}_n$ is the sum of $(\mathbf{I}_1 * \mathbf{I}_2) + (\mathbf{I}_3 * \mathbf{I}_4) + (\mathbf{I}_5 * \mathbf{I}_6) + \ldots + (\mathbf{I}_j * \mathbf{I}_k)$. We assume that the distribution of the elements of each of the item (I) vectors is symmetric around zero. If this assumption is not met, then when successive convolutions are entered into the associative vector the mean value of the associative vector will be correlated with the number of entries. We have also made the assumption (which may in fact be a parameter of a more complete model which includes, for instance, attentional differences) that the power of all the item vectors, as measured by the dot product of an item with itself, is constant. By virtue of the fact that the values on each dimension were chosen randomly in the simulations that follow, the dot product of any vector with any other vector has an expected value of zero. Given these assumptions, if $\mathbf{I}_2$ is correlated with $\mathbf{A}_n$ above, $\mathbf{I}_1'$ will be the result; if $\mathbf{I}_3$ is correlated with $\mathbf{A}_n$, $\mathbf{I}_4'$ will be the result. Since we did not know how increasing the number of entries in the associative memory vector would affect the goodness of recall, we ran a number of simulations, which will be reported shortly, to determine the extent of interference.

Memory Levels

In the present model we conceive of the memory system as consisting of at least two levels, and probably more. The two necessary levels will be designated the I level and the A level, where the I level consists of items, and the A level consists of the sum of the associations. In order to go from the I level to the A level, convolution occurs, and the makeup of the A level is modified. In order to go from the A level to the I level, correlation must occur, and the A level is not changed. There may be other levels involved, however. For instance, in order to analyze sensory input up to the I level, or to convert an item at the I level into an overt response, certain processes are no doubt necessary. Similarly, there may be levels higher than the A level. However, for the sake of simplicity we shall deal only with the I and A levels.

The A level acts as a memory store and is composed of a *single* vector containing the sum of successive convolutions. The I level acts as a pattern recognizer and consists of a lexicon of item vectors which refer to responses. It should be noted that recognition at the I level does not consist of determining whether or not a generated item was a member of the list. Rather, it is an identification akin to perceptual identification, designating the generated vector as such and such an item. This identification process is necessary because the vector generated by means of the correlation of an item with the A level may be somewhat noisy; the correlation of I_1, say with A_n yields I_2'—an approximation to I_2.

Two vectors from the I level may be convolved and entered into the A level to change the vector that constitutes that level. A single vector from the I level may be correlated with the A level to generate a vector that is recognized at the I level as being most like item "X" which may then be given as a response or may itself be correlated with the A level to generate another vector.

Simulations

It is important to demonstrate that if two items are combined by convolution, then when one of the items is correlated with the combination, an approximation to the *other* item will result. It has also been claimed above that more than a single pair of items can be convolved and entered into the same associative memory vector and that the correlation of an item (the cue) with the associative memory vector will still generate an approximation to the item with which the cue item was initially convolved. To substantiate these claims, we ran a number of computer simulations of convolution and correlation. It should be stressed that the program was in no way constrained to produce the results that follow—the operations used to encode and retrieve gave rise to the results. Beyond demonstrating that convolution and correlation can be used as associative encoding and retrieval mechanisms, we wished to determine how (or indeed if) the number of features in the item vectors, and the number of entries in the associative memory vector, influence the clarity of the vector generated by correlation.

A lexicon of 100 items was constructed. Each item in the lexicon was represented as a vector of features, where the value of each feature was randomly selected from a uniform distribution with a range of -1 to $+1$. Thus, on the average, the mean value of each vector was zero. The vectors were normalized so that they were of equal power (i.e., the expectation of the dot product of any vector with itself equaled the expectation of the dot product of any other vector with itself). The number of features in the vectors was manipulated.

A varying number of randomly selected pairs of item vectors from the lexicon were convolved and the results of the convolutions were added into an associative memory vector (corresponding to the A level above). The subroutine used for convolution and correlation is outlined in Murdock (1979).[1] The associative memory vector consisted of the sum of the convolutions from 1 to 21 pairs of items. After con-

[1] When two vectors are convolved, the resultant vector is of larger dimension than the two separate vectors. For instance, convolving two three-element vectors results in a five-element vector. The program used truncates the resultant vector to the central N elements (three in this case). A similar result occurs with correlation, and the program also truncates the generated vector. Most of the information is contained in the central elements and the effect of this truncation is small.

structing the associative vector, one of the item vectors that had initially been convolved (i.e., a *cue*) was correlated with it. This operation resulted in a generated vector (G). The question of interest was: How good an approximation was G to the item that had initially been convolved with the cue? To illustrate, consider a case in which five entries are made into the associative memory vector, as follows:

$$\mathbf{A} = (\mathbf{I}_a * \mathbf{I}_b) + (\mathbf{I}_c * \mathbf{I}_d) + (\mathbf{I}_e * \mathbf{I}_f) + (\mathbf{I}_g * \mathbf{I}_h) + (\mathbf{I}_i * \mathbf{I}_j), \quad (5)$$

where each I represents a different item vector randomly chosen from the lexicon. One of the encoded vectors, say $\mathbf{I}_f$, was correlated with A to result in a generated vector:

$$\mathbf{I}_f \# \mathbf{A} = \mathbf{G}. \quad (6)$$

Since $\mathbf{I}_f$ was originally convolved with $\mathbf{I}_e$, the question was: How well does G correspond to $\mathbf{I}_e$?

In order to determine the similarity of G to $\mathbf{I}_e$, G was matched to every item in the lexicon by taking its dot product with every lexical item. The item in the lexicon that demonstrated the highest dot product (i.e., that was the best match with G) was considered to be what G was identified as. If the lexical item $\mathbf{I}_e$ had the highest dot product then a correct recall was said to have occurred. If anything else matched best then the response was considered incorrect. Since the lexicon consisted of 100 items, this was not a trivial test.

Figure 1 illustrates the results of these simulations. Each point in Figure 1 is based on 100 independent runs of the entire simulation outlined above. The top panel illustrates that the correct identification of the generated vector decreases as the number of entries in the associative vector increases. The bottom panel replots the same results to show that the correct identification of G increases as the number of features in each of the item vectors increases, and that the increase is linear. By interpolating and extrapolating the functions in the bottom panel, the relation between the number of features and the number of pairs in the associative memory vector, for a given level of performance, can be ascertained. The relation may be expressed as

$$\Pr\{\text{Recall}\} = k\frac{F}{E}, \quad (7)$$

where Pr {Recall} is the probability of a correct identification of the generated vector, F is the number of features, E is the number of convolutions entered into the associative memory vector, and the slope k depends upon the lexicon size.

In a second set of simulations the lexicon size was varied. As can be seen in Figure 2, the frequency of correct responses decreased somewhat with increasing lexicon size, but the magnitude of the decrease was small when compared with the effect of increasing the number of entries in the associative memory vector. Presumably a very large lexicon, and a corresponding large number of features would be necessary to realistically simulate human performance.

A third simulation was run in which a particular cue item was convolved with more than one other item. The control associative vector was set up in the same manner as in the above simulations:

$$\mathbf{A}_c = (\mathbf{I}_1 * \mathbf{I}_2) + (\mathbf{I}_3 * \mathbf{I}_4) + (\mathbf{I}_5 * \mathbf{I}_6) + \ldots + (\mathbf{I}_{n-1} * \mathbf{I}_n).$$

The "experimental" associative vector was constructed so that one of the items was involved in multiple convolutions:

$$\mathbf{A}_e = (\mathbf{I}_1 * \mathbf{I}_2) + (\mathbf{I}_1 * \mathbf{I}_4) + (\mathbf{I}_1 * \mathbf{I}_6) + \ldots + (\mathbf{I}_1 * \mathbf{I}_n).$$

Each of the vectors consisted of either 15 or 75 features and the lexicon contained 100 item vectors. When $\mathbf{I}_1$ was correlated with either $\mathbf{A}_c$ or $\mathbf{A}_e$, it was found that the probability of recall of one of the items with which $\mathbf{I}_1$ had been convolved in the experimental condition was higher than the probability of recall of the single item with which $\mathbf{I}_1$ had been convolved in the control condi-

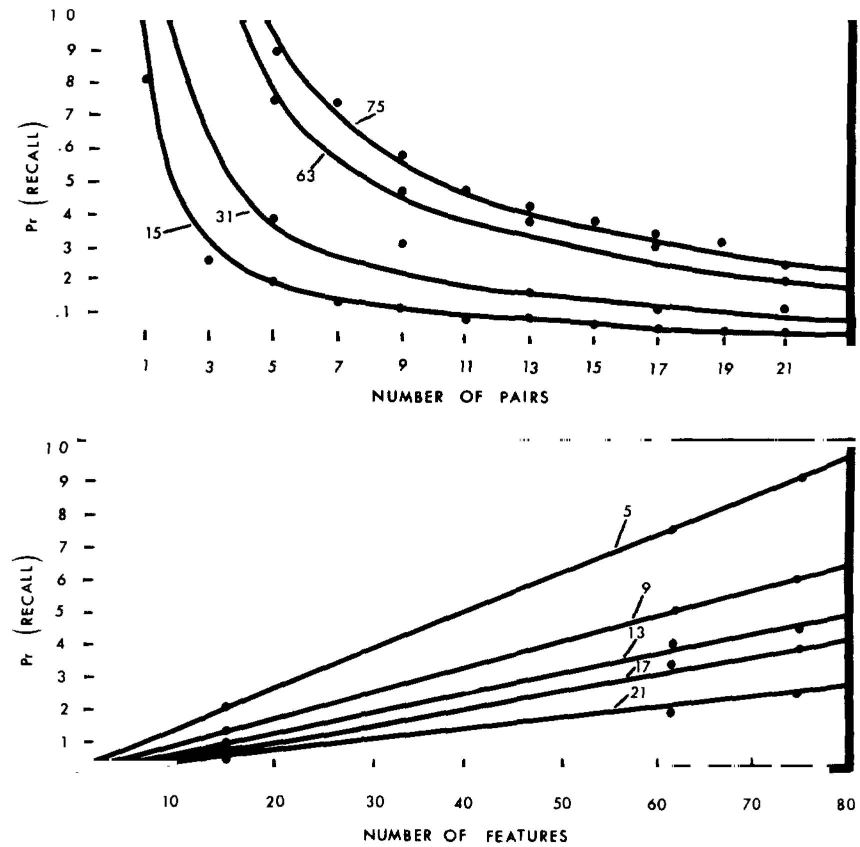

FIG. 1. The relation between the number of features in each vector, the number of entries in the associative memory vector, and the probability of recall when convolution and correlation are used as the associative encoding operations. In the top panel the parameter is the number of features. In the bottom panel the parameter is the number of entries (or pairs) in the associative memory vector.

tion. However, the probability of recall of the particular target item—I_2— was higher in the control condition. On the other hand, when the item involved in multiple convolutions was the target item (i.e., when I_2 was used as the cue to generate I_1 in the above two cases) it made no difference that the item had been entered into the associative memory vector a number of times. The probability of recall was the same when the target item was entered once or many times. This somewhat surprising finding may not be found when both of the convolved items are the same and are entered into the associative memory vector multiple times.

Repercussions of Convolution and Correlation

There are certain repercussions of assuming that convolution and correlation are the processes used to associate and retrieve items, as indicated below.

First, recall based upon the mechanism of correlation must be cued. That is, nothing can be retrieved from the A-level unless a cue or item vector is correlated with the vector constituting the A level.

Second, when an item vector is correlated with the A level, it will generate an item with which it was originally convolved. This is a slight simplification of the model, since if an insufficient number of features has been encoded to permit perfect identification of the vector which is generated, the generated vector may be incorrectly identified as an item other than the one with which the cue-item was convolved. Even in this case, however, the generated item is most likely to be iden-

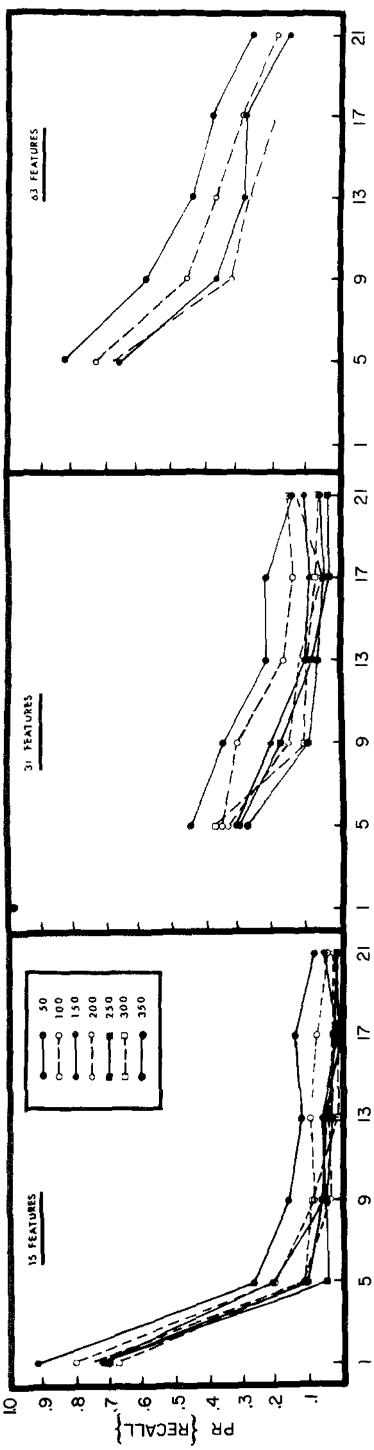

FIG. 2. The decrease in level of recall attributable to an increase in the size of the lexicon when convolution and correlation are used as associative encoding and retrieval operations. The parameter in each panel is the size of the lexicon which was incremented by 50 "words" for each of the functions. The panels show this effect for vector sizes 15, 31, and 63, and a varying number of entries or pairs in the memory vector.

tified as the item that was initially convolved with the cue item.

Third, association occurs by contiguity. It is not necessary to assume that items are similar or preexperimentally related in order for convolution to occur.

Fourth, the clarity of the item generated using correlation is a function of both the number of entries in the cumulative associative vector and the number of features that are encoded. The relation among number of entries, number of features, and the clarity of the generated vector is given by Eq. 7, above.

Fifth, at least two levels are required: a level at which the unit is the item, and a memory level, which has been called the A level.

Sixth, if correlation is the only retrieval mechanism, an "item" must be present in the mind before the first list item is presented so that the first list item may be convolved with something. To illustrate this necessity, consider the case in which a list consists of only one item. In this case, convolution cannot occur unless there is a "start" item which exists prior to list presentation, since two vectors are necessary for convolution. If convolution does not occur, the item cannot be retrieved via correlation.

Finally, the association formed by convolution is symmetric since the operation itself is commutative: $(\mathrm{I}_x * \mathrm{I}_y) = (\mathrm{I}_y * \mathrm{I}_x)$. What have traditionally been called backward and forward associations simultaneously result from the single operation of convolution.

The Process Model

In this section we shall outline a process model that allows for rehearsal and recall in a single-trial free-recall situation. The process is somewhat constrained by the assumption that convolution and correlation are the associative encoding and retrieval mechanisms.

As outlined above, there must be two items for convolution to occur. Hence we make the assumption that something exists in the mind at the I level before the first list item is presented. Let us call this something *context* (I_c) and characterize it as the subject's internal state—level of awareness, mood, and so on. (Note that context, as used in the present paper, does not refer to verbal context, and does not bias the meaning of items at the I level. It is also not associated with every item in the list.) Context is assumed to be functionally equivalent to list items in every way except that it is not a word, and so would not be overtly articulated. Additionally it may be present at the I level in the absence of list items. When the first list item is presented, it is encoded up to the I level and associated with context. The result of the association is entered at the A level.

In order to rehearse, the list item that has just been presented, and which is assumed to still be present at the I level, is used as a retrieval cue to generate something from the A level. Since the item generated by the correlation of a cue with the A level will resemble the item with which the cue was originally convolved, using the first list item as a cue will cause context to be generated. The generated item is identified at the I level then used as a retrieval cue to generate from the A level an approximation to the item with which it was initially convolved. The generated item is identified and subsequently used as a cue to generate another item. Thus rehearsal alternates between context and the first list item until the second list item is presented.

When the second list item is presented it is encoded up to the I level and associated with whatever is currently being rehearsed. Thus the second list item may be associated with either the first list item or with context. Suppose, for the purpose of illustration, that the second list item is associated with the first item. The association is entered into the A level and then the second item is used as a retrieval cue to generate an item from the A level. The item which is generated by I_2 is identified at the I level as being I_1. When I_1, in turn, cues the A level, the vector that is generated is actually a combination of I_2 and I_c. This vector is identified as one or the other of I_2 or I_c (we shall shortly outline the process), and the identified item is then used as a cue to generate another item with which it was initially associated.

In summary, when an item is presented, it is encoded to the I level and associated with the item that is currently at the I level. The association is stored at the A level. Then the just-presented item is used as a cue to generate an item with which *it* was associated. That item is identified and used as a cue to generate an item with which it was associated. That item is identified and used as a cue to generate another item. This process continues until either the next item is presented and the whole process repeats or the signal is given to recall.

At time of recall, the last rehearsed item is output as the first recalled item. It is then used as a cue to generate from the A level another item that is identified, recalled and then itself used as a cue. Successive recall and then cuing with the recalled item continues until a certain amount of time has passed without the recall of any previously unrecalled items. Then context is reinstated as a cue and recall begins afresh. The item that context generates from the A level is identified and then used as a cue to generate another item. Recall continues in the same manner as before, and in essentially the same way that rehearsal had occurred, until a certain amount of time passes for the second time without the recall of any as yet unrecalled items.

Simulations of the Model

It can be extremely expensive to simulate convolution and correlation on the computer. Hence, our strategy in applying the model to the free-recall data was to use the information about convolution and correlation to construct a program that behaved as if these mechanisms were actually used. To be sure that we had not misapplied the findings about convolution and correlation, however, we also ran a minisimulation of the process model outlined in the previous section actually using convolution and correlation.

Convolution–correlation simulations of the free recall process. In this simulation, items consisted of vectors of 143 features. When the first "list" item was presented, it was convolved with another item which was designated as "context." Context, like the other items, consisted of 143 features randomly drawn from a uniform distribution as in the simulations already described. The resultant association was added into a memory vector corresponding to the A level. Then the just-presented item was correlated with the associative vector, resulting in a generated vector.

In order to identify the generated vector, a feature was selected at random from the generated vector and matched by multiplying, to the corresponding feature of each item in the lexicon. Features continued to be randomly selected and matched against the lexicon until the sum of the products for one of the lexical items reached a criterion. The lexical item that first reached the criterion was considered to be the item that was rehearsed. Each feature match was assumed to require one unit of time. This process may be viewed as a random walk. The identified item (in this case it would be context) was then correlated with the associative vector—generating another vector which was identified in the same way and which was then itself correlated with the associative vector. Rehearsal thus proceeded according to the process outlined in the previous section. There was a fixed amount of time between the presentation of any two items. When the next item was presented, it was convolved with the last item that had been identified (i.e., the item that was being rehearsed) and the result of the convolution was added into the associative memory vector. Rehearsal then proceeded as before.

Since the variance of the associative memory vector and hence the generated vector increases with the entry of each successive convolution, when the identification process above is used there will be an increasing number of fast incorrect responses as the number of entries into the associative memory vector increases (unless the variance is somehow controlled). There are a number of ways to avoid this problem. In the simulations we normalized the generated vector. One could also normalize the associative memory vector, or could allow the criterion to increase with the number of entries. Coefficients (x and y) could be assigned to $\mathbf{A}_n$ and ($\mathbf{I}_p * \mathbf{I}_q$) when they are added to form $\mathbf{A}_{n+1}$. The coefficients that yield an equal weighting for each entry into the associative memory vector are

$$x = n\left(\frac{1}{n^2 + 1}\right)^{1/2} \qquad (8)$$

and

$$y = \left(\frac{1}{n^2 + 1}\right)^{1/2}, \qquad (9)$$

where n is the number of entries in the associative memory vector. If this method is used, the form of the entry of convolutions into the associative memory vector becomes

$$A_{n+1} = x\,\mathbf{A}_n + y(\mathbf{I}_p * \mathbf{I}_q). \qquad (10)$$

The choice between these methods might be theoretically interesting but is beyond the scope of the present paper.

The results of the simulation were that no errors were made (i.e., no intrusions occurred in rehearsal) until nine pairs had been added into the associative memory vector. After nine entries, intrusions became increasingly more frequent, as would be expected from Eq. 7 for the number of features employed in the simulation. The pattern recognizer introduced variability into the identification of the generated vector. In particular, when an item was convolved with more than one other item, the vector that it generated would sometimes be identified as one item, sometimes another. As more entries were made in the associative memory vector, it required more time to correctly identify a given pattern. It appeared that the number of items recalled per presentation interval was inversely proportional to the number of entries in the associative memory vector, again as would be expected from Eq. 7.

Simulations to generate the data. In the following simulations, convolution and correlation were not used because of the expense. Instead, the computer was programmed to "associate" at the time of presentation and to generate the item with which a cue had initially been associated. This was accomplished by constructing a matrix in which the first entry referenced the second. Thus if A was the item that was being rehearsed when B was presented, the matrix entries (A,B) and (B,A) would be marked. Even though convolution would result in the ability of A to generate B and B to generate A from a single operation, it was necessary to program both associations in the abbreviated program. It was also necessary to include a separate parameter (which will be detailed shortly) to mimic the buildup of noise that occurs with each successive convolution when the actual operations of convolution and correlation are used. Thus, certain properties that fall out when convolution and correlation are used had to be explicitly programmed into the abbreviated simulation that did not use convolution and correlation. The present schematization of the model gives a valid representation of what would happen with convolution and correlation only if it is assumed that a sufficiently large number of features are encoded to allow correct identification to occur consistently. The present simulations are thus a special case of the more general model outlined above.

If an item was associated with more than one other item, the program randomly selected one of the associates. This random selection corresponds to the fact that the random walk pattern recognizer causes variability in the recognition of ambiguous patterns or of vectors consisting of the combination of more than one item. The time taken for the identification process in the last section was simply assigned a parameter value in the present simulation. This parameter was fixed at 50 milliseconds/entry at the A level. It should be noted that we are not particularly committed to this or any other parameter value, but it was necessary to assign values. Our strategy was to assign values in what seemed to us a reasonable manner and to leave them fixed as constants for all the simulations unless a particular experimental manipulation suggested that a certain parameter would change. In this specific instance, 50 milliseconds was chosen because the identification process envisaged bears a resemblance to recognition processes that have been proposed elsewhere (Anderson, 1973; Ratcliff, 1978). Thus we allowed the slope of the identification function to be of the order of magnitude of the reaction time slope found in the Sternberg (1969) paradigm.

We assigned a parameter to the time taken for encoding up to the I level and association formation. The encoding parameter was a random variable with a mean of 1.5 seconds and a standard deviation of .66. If the time required for encoding of a particular item exceeded the time available (i.e., the presentation rate) the item in question was not encoded and the next item was associated with context. Phenomenologically this corresponds to the fact that when subjects are presented with a list at a fast rate, they occasionally allow a word to pass unattended as they prepare to concentrate on the next word. Since it requires some time to say a word aloud, we assigned a value of 0.5 second for vocalization of any recalled or rehearsed item.

The probability of using context as a retrieval cue at time of recall was set at 1. The stop rule, or the amount of time that would be spent retrieving with no new successful recall, was set at 3 seconds. This does not include any time to vocalize or to check whether an item has previously been recalled (although a subroutine to check for previous recall was necessary in the simulations). Thus the 3-second limit reflects "pure" retrieval time.

The only other variables in the model were the presentation rate, list length, and delay before recall. These were not considered parameters of the model and were assigned numerical values in accordance with the experiments modeled. The program using the process and the parameters cited above was run 200 times (i.e., through 200 lists) to generate each of the results described below. The parameters of the model were fixed as noted above unless explicitly stated otherwise in the text.

Simulation Results

Rehearsal

Comparison of the model to rehearsal data is necessarily limited to those experiments in which the overt rehearsal procedure was used. In this procedure, the subject is presented with a list of words and is asked to say aloud those words about which he or she is thinking during list presentation. The rehearsal protocol is recorded and may be analyzed in various ways.

Rehearsal frequency of each serial position. One of the most familiar results stemming from the overt rehearsal procedure is the finding that items presented early in the list are rehearsed more frequently than are those presented in later positions. Figure 3 illustrates this finding. The top solid curve is from Rundus and Atkinson (1970), who presented subjects with 20-item lists of nouns at a 5-second rate, and asked them to repeat aloud items from the current list as they were being studied. The lower solid curve is from Murdock and Metcalfe (1978), whose materials were common words presented at a 5-second rate. The broken line is the result generated by the model when the presentation time was set at 5 seconds. The model qualitatively corresponds quite closely to the data, although it slightly overestimates the frequency of rehearsal of the first item.

Rehearsal frequency at each presentation interval. In the model, context is encoded as the "item" which precedes the

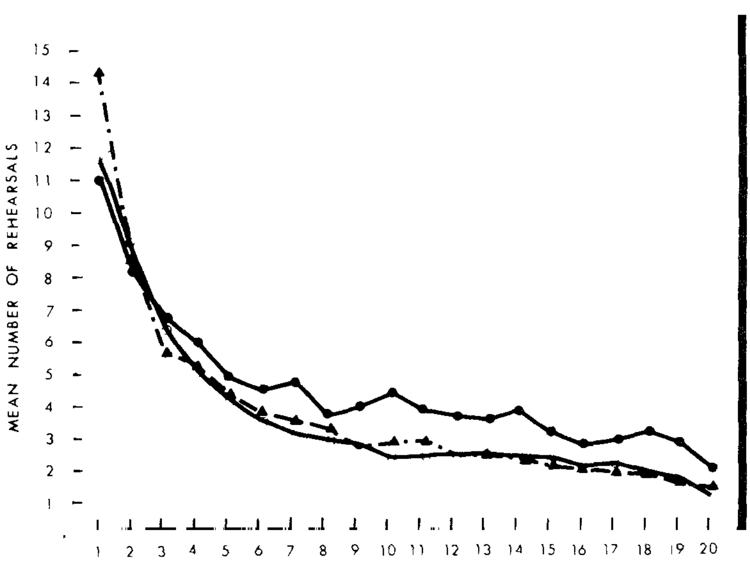

FIG. 3. Mean number of rehearsals for each successive item in a 20-word list presented at a 5-sec rate. The top curve (●——●) is from Rundus and Atkinson (1970), the bottom curve is from Murdock and Metcalfe (1978) (○——○), and the broken line (▲– – –▲) is from the model.

presentation of the first list item, and is thereafter rehearsed as if it were a list item except that it is not overtly articulated. If this role of context is correct, one might expect to observe an irregularity in the pattern of overt rehearsals which allows the detection of the rehearsal of context. An abnormally low frequency of overt rehearsals should be found over the first few presentation intervals where context is being rehearsed frequently but covertly. In the model, the number of rehearsals between the presentation of any two items is inversely proportional to the number of entries at the A level. The fine broken line in Figure 4 shows this relation. The heavy broken line shows the result of the model when context (which is included in the fine line) is not counted as a rehearsal, as it could not be in the experiment since it is nonverbal. The top solid line is from Rundus and Atkinson (1970) and the bottom solid line is from Murdock and Metcalfe (1978). Both sets of data show the influence of the covert rehearsal of context at the beginning of the list. Also, the figure illustrates that there is about a one-word decrease in the frequency of rehearsal from presentation interval 3 to the end of the list, whereas the model shows about a two-word decrease. There are two possible explanations for the discrepancy between the model and the data. The first is that the parameter value for the increase in identification time was set too high; a lower value yields a shallower slope. The second possible explanation is that there is no provision for repetition in the model, that is, an item may be rehearsed only by being cued by another item. In the present version of the model, an item at the I level is never repeated by rote. If repetition occurred in the model, and more particularly if it occurred more frequently towards the end of the list, a shallower slope would be obtained.

Nominal and functional serial positions. The overt-rehearsal procedure has been used to analyze data in terms of functional serial position as well as nominal serial position (Brodie, 1975; Brodie & Murdock, 1977; Melton & Glenberg, cited in Bjork & Whitten, 1974). The term nominal serial position refers simply to input or presentation position, and is identical to what is normally meant by serial position. The functional arrangement uses the rehearsal data generated by subjects to order the items. The last item rehearsed (which may or may

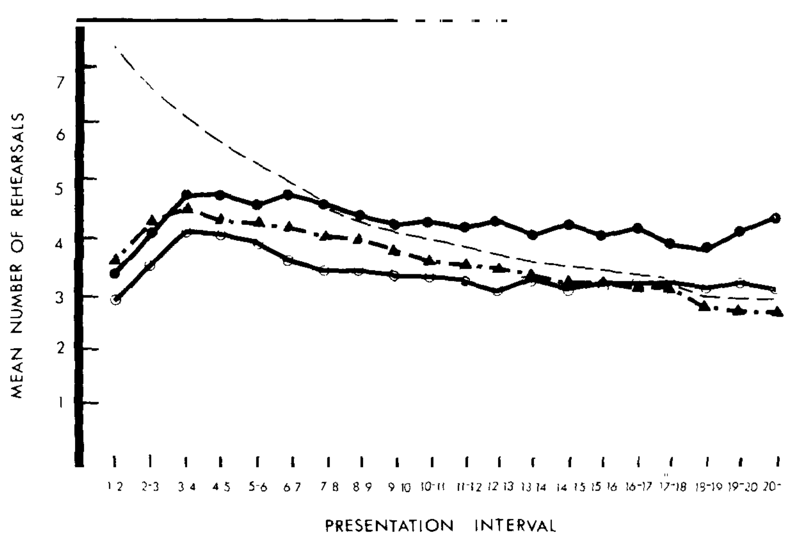

FIG. 4. Mean number of rehearsals in each presentation interval—Rundus and Atkinson (●——●), Murdock & Metcalfe (○——○), model (▲———▲). The presentation interval is the (5 sec) period of time elapsing between the presentation of one item and the next.

not be the last item presented) is the last functional item; the second last item rehearsed is the second last functional item; the third last item rehearsed is the third last functional item, and so on. The first item to stop being rehearsed is the first functional item.

In Figure 5, the mean functional position of each nominal item has been plotted for two presentation rates—1.25 and 5 seconds. The broken lines represent the results of the model, the solid lines are estimated from Figure 6 in Brodie and Murdock (1977). As can be seen from the left panel, at a fast presentation rate both the data and the model show that the nominal and functional serial positions are virtually identical, as would be expected if little or no rehearsal occurred. The right panel shows the obtained and predicted results at a slow rate. Under this condition the early nominal items drift to later functional positions and late items regress to earlier functional positions. Of more interest is the finding that very early items move to later functional positions than do early–middle items. In the model the structure of associations generated by means of rehearsal gives rise to this result. In particular, the associative structure which occurs by means of the rehearsal process is "thicker" for early items than for later items, that is, early items are associated to more other items than are later ones. The result is that there are more retrieval routes available for earlier items, or more cues that will be effective in retrieving the early items, and so they have a tendency to be retrieved more frequently throughout rehearsal.

Although the model does not perfectly reflect the rehearsal of subjects it does qualitatively show the pattern that is exhibited. The decreasing number of rehearsals for each serial position is shown by the model. The increase in the number of rehearsals in the first few presentation intervals is predicted, as is the subsequent decrease in number of rehearsals by presentation interval. The high correspondence between nominal and functional serial positions at a fast rate as well as the breakdown in this correspondence at a slow rate are also shown by the model. While the particular form of the breakdown of the correspondence between nominal and functional serial positions is not exactly the same in the data and the model, the general pattern—namely, that are very early items drift to later functional serial positions than do early–middle items—is predicted by the model.

Recall

List length. One of the most basic findings in single-trial free recall is the finding of serial position effects. At various list lengths, the recency effect remains relatively constant while the probability of recall of earlier items decreases as list length increases. In order to simulate the effects of list length, the parameters of the model were held constant and only the list length was varied. Figure 6 (bottom) shows the list length predictions of the model based on 200 simulated lists for each curve. Figure 6 (top) shows the result of varying list length in Murdock's (1962) experiment. In the experiment, lists were presented at two different rates—2 seconds/item for the 10-, 15-, and 20-word lists, and 1 second/item for the 20-, 30-, and 40-word lists. The interitem interval was set at 1.25 seconds in the model. As can be seen from the figure, the model, like the data, shows a recency effect of comparable magnitude at all list lengths while the level of recall of prerecency items decreases with increasing list length. In the model, recency occurs because the last item at the I level prior to recall (and hence the first item recalled) is most likely to be the last nominal item. This item provides an entry point into the end of the list. Primacy occurs because context is used as a retrieval cue and context was nearly always associated with the first item, although it may also have been associated with other items. Context provides an entry point into the beginning of the list. The primacy effect decreases with an increase in list length because with longer lists there is

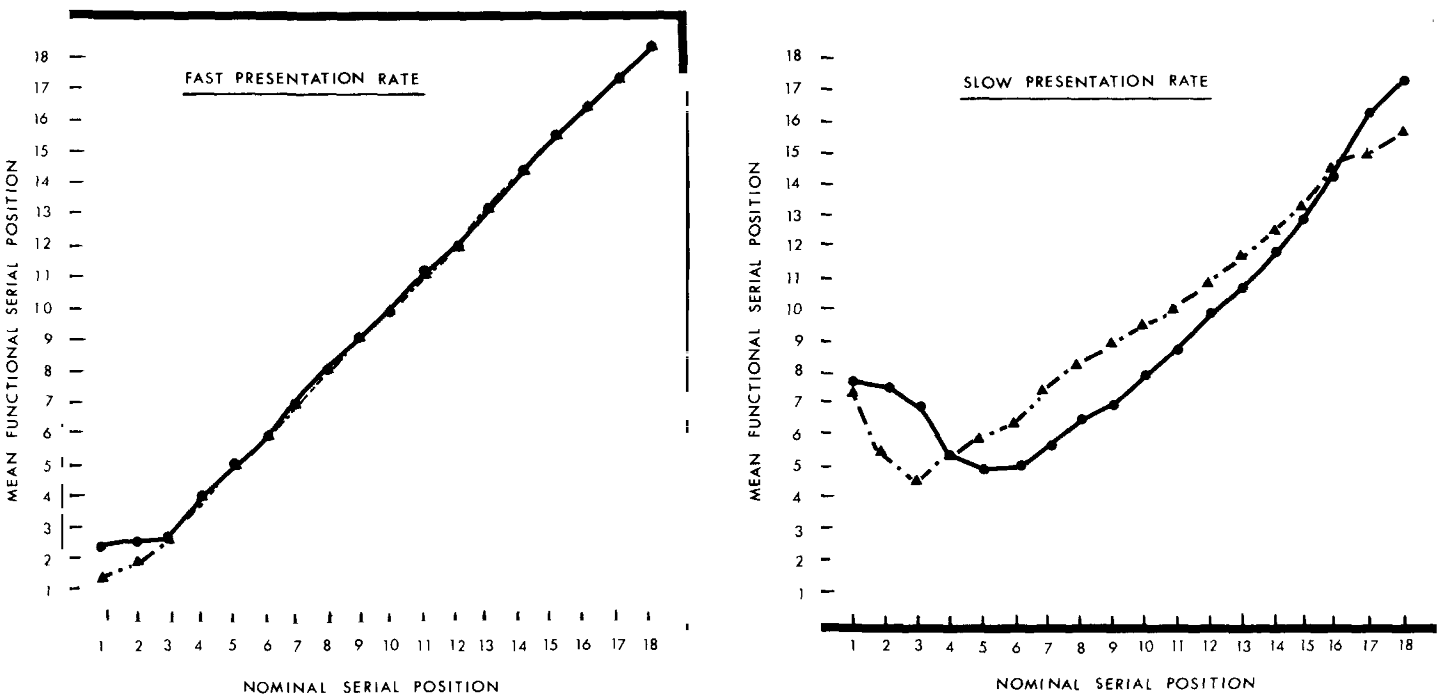

FIG. 5. Relation between input, or nominal serial position, and functional serial position, or when an item is last rehearsed. The mean functional positions for each nominal position from Brodie and Murdock (1977) (●——●) and the model (▲– – –▲) are shown for a 1.25 and 5 sec presentation rate in the left and right panels, respectively.

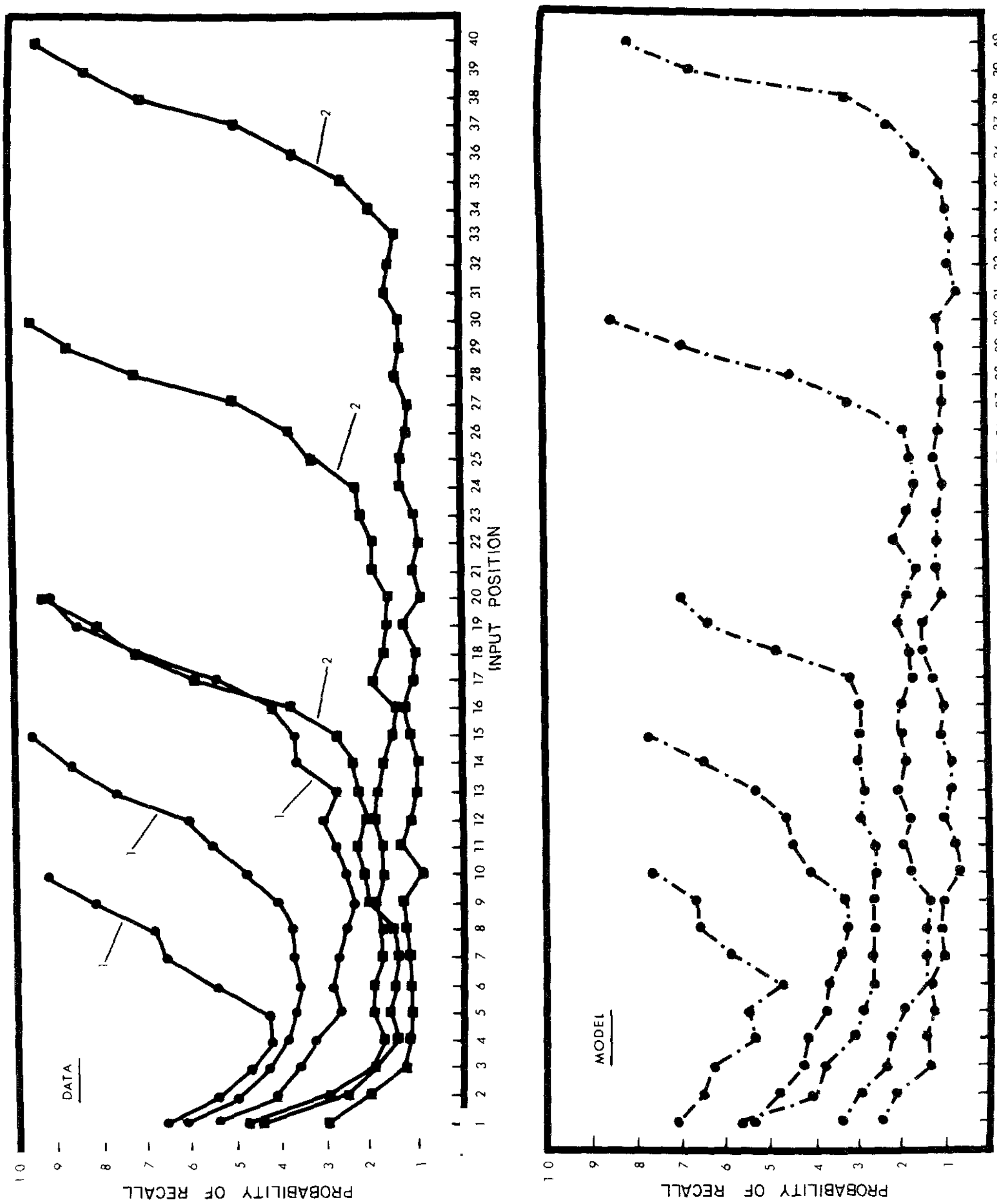

FIG. 6. Recall as a function of list length from Murdock (1962) in the top panel, with a presentation rate of either one sec/word or two/sec word. The simulations were set at a rate of 1.25 sec/word (bottom panel).

a greater chance that context will also be associated with something other than the first list item. The overall probability of recall is lower with long lists than with short ones because there are more entries at the A level with the longer lists. When there are more entries at the A level there is more interference (as in Eq. 7), and the generated items are noisier.

Presentation rate. The effects of presentation rate have often been cited as evidence for two-store models of memory such as those of Waugh and Norman (1965) or Atkinson and Shiffrin (1968). The finding that rate of presentation does not affect the recency portion of the nominal serial position curve but does change the primacy and asymptotic parts of the curve has been taken as evidence for two distinct underlying stores. Recently, however, Brodie and Murdock (1977) have demonstrated that this effect does not hold up when the serial position curves are plotted functionally. They have claimed that "these presentation rate results contradict predictions from encoding, storage and retrieval models" (p. 199). Since Brodie and Murdock have made a strong claim about the theoretical implications of the effects of presentation rate (although Kintsch & Polson, 1979, have developed a storage model which seems to contradict the claim for at least one class of models implicated), it would seem a worthwhile endeavor to examine the results generated by the present encoding–retrieval model with respect to presentation rate.

The top two panels of Figure 8 reproduce Brodie and Murdock's findings; the bottom panels are the results generated by the present model. It can be seen that for both the data and the simulation the nominal and functional serial position recall curves are much more similar at a fast than at a slow presentation rate. This is, of course, what would be expected since at a fast rate the nominal-to-functional correspondence is high, as in Figure 5. The nominal serial position curves show that the main advantage of the slow rate over the fast rate is in the prerecency positions. The functional serial position curve generated by the model at a slow rate was more bowed than was that shown by the data. By increasing the interitem interval in the model the amount of bowing in the functional serial position curve can be made to appear more similar to the functional data. However, this manipulation causes a decrease in the level of recall of the last nominal item, which is already underpredicted by the model. It seems likely that this underprediction results because subjects change strategies (to a rote repetition strategy) toward the end of the list, whereas the model does not. (See also Watkins & Watkins, 1974.) The main difference between presentation rates in the functional serial position curves was in the midlist items, in both the data and the model.

There is a difference in overall level of recall between the fast and slow rates because, first, at a fast rate there is a slightly smaller chance that an item will be encoded. However, separate simulations were conducted in which the probability of encoding was held constant and the difference in level of recall was nevertheless obtained. The second reason for the difference is that the structure of associations formed at a slow rate when rehearsals are allowed to intervene is more interconnected. At a slow rate there are a great number of items which allow recall of several other items and hence the continued recall of previously unrecalled items is easier than at a fast rate.

The effects of delay. In the model, the recency effect is quite fragile, as appears to be true in the data as well. The left panel of Figure 8 depicts the results of the Glanzer and Cunitz (1966) experiment in which recall was delayed for either 0, 10, or 30 seconds. During the delay, the subjects performed a minimal task of counting aloud. The 15-item lists were presented at a 3-second rate. In the simulation of this result, the interitem interval was set at 3 seconds, rehearsal was allowed to continue for either 0, 10, or 30 seconds after the end of the list,

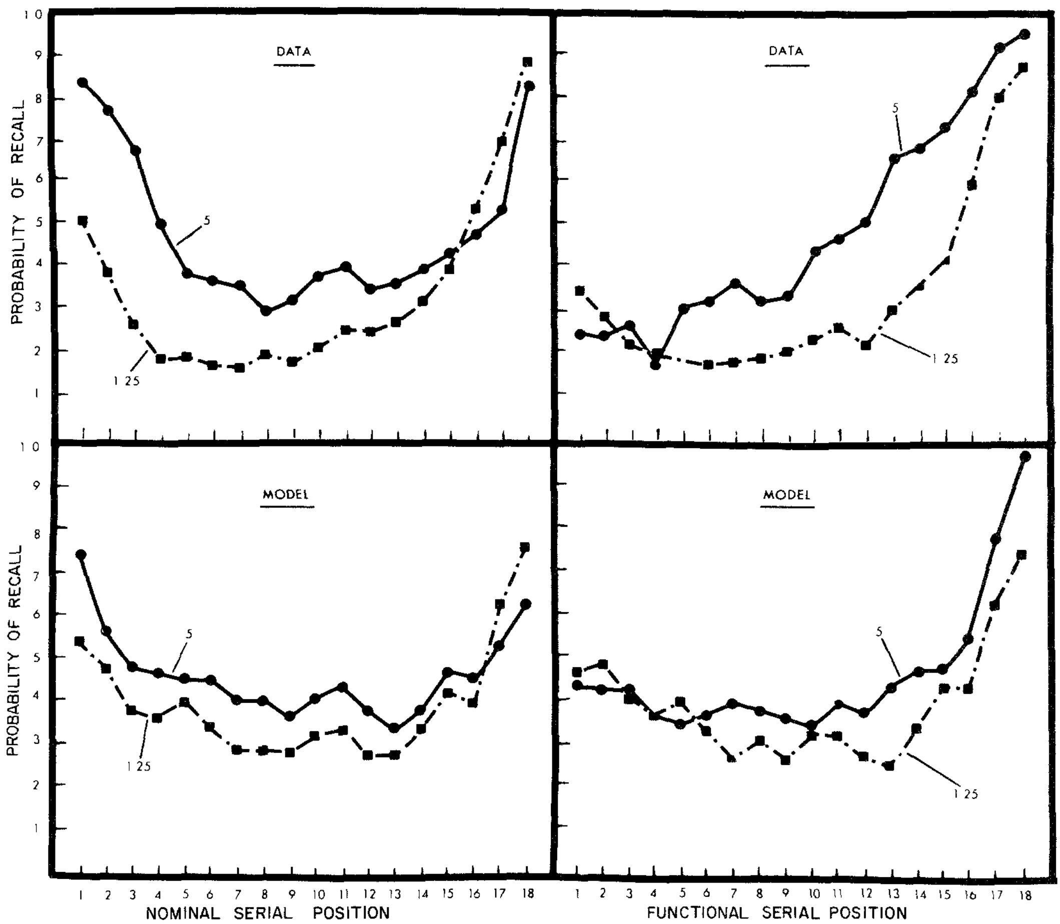

FIG. 7. Effect of presentation rate on nominal and functional serial position recall curves from Brodie Murdock (1977; top panels) and the model (bottom panels). The two rates were 1.25 sec/word and 5 sec/word.

and then recall occurred as in the simulations reported above. The right panel shows the results of the model, where each of the curves are based on 200 simulated lists, as before. Both the model and the data show a larger deficit in the recall of recency items at longer delays. It is also notable that in this experiment the level of recall of the last few items was lower than is often found in immediate recall tasks, and more consistent with the predictions of the model. This may have occurred because subjects were less inclined to switch rehearsal strategies at the end of the list because of the possibility that they would have to perform the distracting task, and such a strategy change would be inefficient. No attempt was made to adjust the overall level of performance of the model to that of the data. Raymond's results (cited in Glanzer, 1972) for high- and low-scoring subjects show that the effects of delay are much the same regardless of the overall level of recall.

Notice that the decrease in recency is attributable, in the model, to continued *free* rehearsal. The last presented item is displaced from the I level by other items that are rehearsed and thus the optimum entry point into the end of the list is lost. In an experiment such as that of Bjork and Whitten (1974), in which the subject is instructed to use *constrained* rehearsal of only the last presented item or pair, the optimum retrieval cue into the end of the list

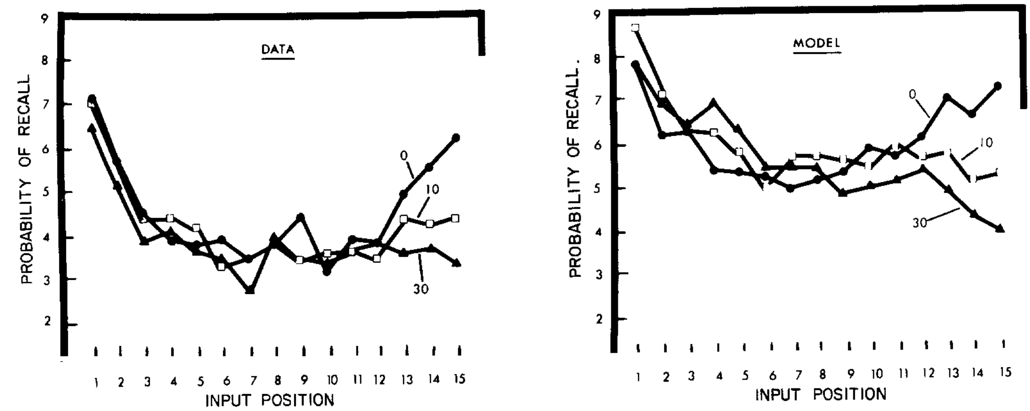

FIG. 8. Effect of a 0, 10 or 30 sec delay from Glanzer and Cunitz (1966; left panel) and the model (right panel).

might still exist at the I level. The present explanation of the effects of delay suggests that it may be possible to reinstate the lost recency effect caused by a delay interval by means of appropriate cuing.

Modality effect. In the simulations above, the parameters of the model have remained unchanged as outlined earlier, and only experimental variables such as presentation rate and list length were altered to correspond to the experimental situations. However, there are several experimental manipulations which could reasonably be expected to alter the parameter values in the model. One experimental manipulation that might be expected to alter the encode time is the modality of presentation. If, as has been suggested by Sperling (1967) and others, a translation operation occurs in the early encoding of visually but not auditorily presented words, then it might be expected that it would take more time to encode visually presented materials than it would for auditorily presented materials. Thus in this simulation, the encoding parameter was varied to mimic the effect that presentation modality might have on the difficulty of encoding items up to the I level. The results of the model are presented in Figure 9, panel B. As can be seen from the figure, the most noticeable effect of this parameter change occurs in the recency section of the serial position curve, as in the data from Murdock and Walker (1969) which are shown in panel A. A less obvious result, but one that other simulations showed to be a prediction of the model, is that the asymptote for the slower (visual) encode time is actually higher than for the faster time. This result occurs because when an item is not encoded, the next item is associated with context. Thus, when encoding is more time consuming, more midlist items are associated with context, and they benefit from this at time of recall.

It is of interest to point out that a result which looks somewhat similar to the present result may be obtained by distraction (see Fig. 8). In the model, however, these apparently similar results are attributed to different causes. One finding which suggests that the present interpretation of the modality effect is correct concerns output orders for auditory and visual presentation (Nilsson, Wright, & Murdock, 1979). Recall, given auditory presentation, appears to begin several words from the end of the list and proceeds mainly in a forward order, whereas with visual presentation the last presented word is usually the first word recalled. In the model, a high level of recall of the last few words combined with a word other than the last presented word being recalled first could only occur if the last word were not only encoded and associated with another word, but if there were sufficient time left over for rehearsal to occur. That recall in the visual case started from the last presented word suggests that there was probably an insufficient time for rehearsal and (as we have assumed in the present simulation) possibly also an insufficient time for encoding since the level of recall of the last word was low in spite of the fact that it was usually recalled first. Other results—such as the finding that in a mixed auditory and visual list, the visual deficit is not restricted to the recency portion of the serial position curve (Murdock & Walker, 1969), and that the effect obtains only at fast presentation rates (Penney, 1975)—provide converging evidence that the locus of the modality effect in free recall may be prior to association formation.

State dependence. The subject's internal state, which has been called context in the model, was assumed to be what was present at the I level before any list items were presented and hence what was associated with the first list item. At time of recall, the subject's state, which was assumed to be perceptually present whenever he chose to monitor it, serves as a retrieval cue—one that accesses the early items in the list and accounts for the primacy effect. This context cue, in all the simulations above, was assumed to be available with a probability of 1. However, if the internal state of the subject were to change from time of study to time of test, it is reasonable to suppose that the effectiveness of this cue would be lessened. Thus, the model predicts that recall performance will be poorer when a subject is in a different state at the time of test. Eich (1977) has reviewed the literature on state dependence in free recall in which subjects were either in the same state at time of study and test or in a different state from the time of study to the time of test. In 11 of the 12 experiments, performance was poorer in the different state than in the same state conditions. The main effect of same versus different state without regard for serial position or output order effects is trivial to model since the context parameter refers directly to state changes. Since no metric has been developed to assess the degree of state change and the nature of the serial position data is unknown for experiments that have explicitly manipulated the change on the subject's internal state from study to test, we have not attempted to simulate the results of the 2 × 2 design.

However, serial position data have been presented by Darley, Tinklenberg, Roth, Hollister, and Atkinson (1973) on the main effects of marihuana on immediate free recall. Subjects in this experiment studied a list of 20 unrelated words presented at a rate of 5 seconds per item, either after having ingested marihuana or in their normal nondrugged state. Other researchers have shown that there are marked changes in subjects' perception of time, subjective high ratings, and in physiological indicators following the administration of marihuana (Kopell, Roth, & Tinklenberg, 1978). In terms of the model, it seems reasonable to assume that context was different at time of study than at time of test in the Darley et al. (1973) experiment. The results of this study are presented in Figure 10, left panel. To simulate this experiment, the interitem presentation interval was set at 5 and parameters were as outlined previously except that in the "changed state" condition, the probability of effectively using context as a cue was set at .6 (rather than at 1). The results of the model are presented in the right panel of Figure 10.

It is of interest to point out that a similar pattern of results was obtained when the presentation rate was altered (see Fig. 7), but that these apparently similar patterns are attributable to different causes. Converging evidence that the context cue, as specified in the model, is primarily responsible for the phenomenon of state dependence comes from several sources. Eich (Note 1) has reported that there is a difference in the output order of subjects who are tested in the same state or a different state from that in which they had studied the list. Specifically, those subjects who were tested in the same state tended to begin recall from the beginning of the list, as would be expected from the model in a much delayed recall test (often several days in the

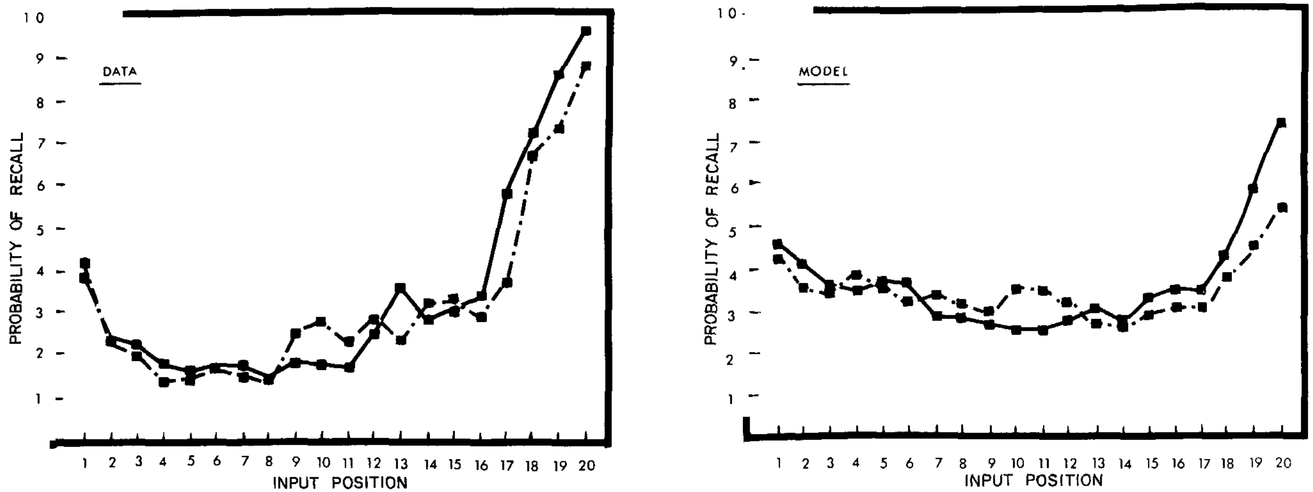

FIG. 9. Effect of auditory (■——■) and visual (■———■) presentation modality (left panel) from Murdock and Walker (1969). The right panel shows the effect of manipulating the encoding time in the model.

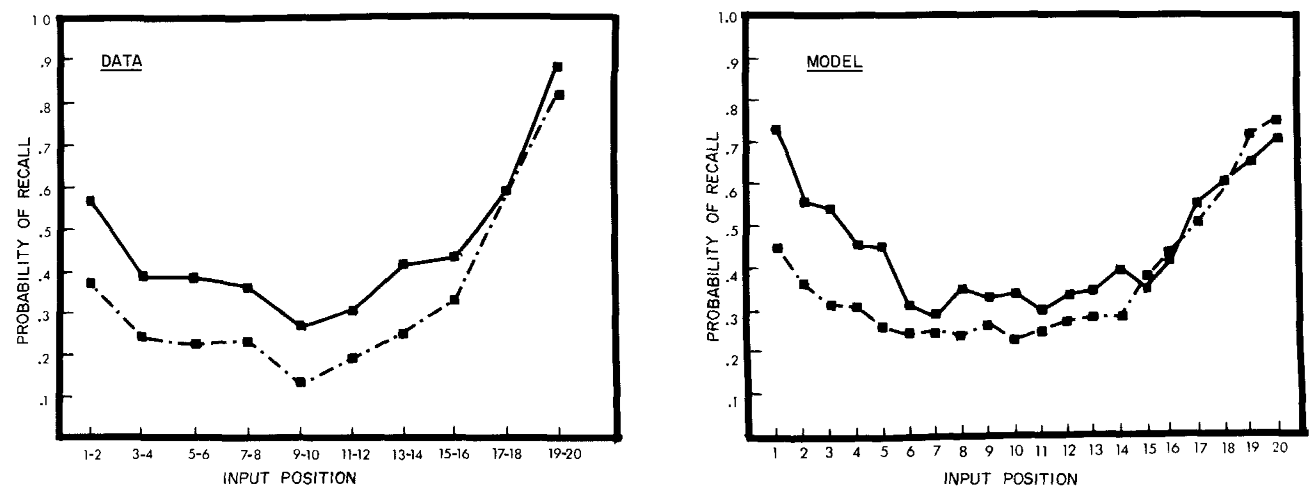

FIG. 10. Effect of marijuana on recall (left panel) from Darley et al. (1973) and the result of altering the context cue in the model (right panel). Marihuana (■———■), control (■——■).

experiments) since the context cue gives access to the early items. Those subjects who were in the different state condition tended not only to recall fewer words but also to recall in a haphazard manner.

Although the output order results provide some evidence that context functions in the manner specified in the model, it would further strengthen the argument if the function of the context cue could be mimicked. Although not intending to test the model, Eich (Note 1) has in fact performed a relevant experiment. He presented subjects with lists of categorized words consisting of 12 categories. At the time of the test, in which the subjects were in a changed state, he provided the subject with either the first, second, or sixth category name as a cue. If the assumptions in the model are correct—that context provides access to the beginning of the list, and that the associative structure generated by rehearsal is thicker at the beginning of the list—then presentation of the first category name, which also can serve as a cue into the beginning of the list, should minimize state dependence. In fact, this is what was found. Presentation of the other category names had a smaller effect.

So far, experiments which study the effect of changed states on free recall offer good support for the role of context as delineated in the present model.

Output Order

In the model, there are two cues that are available, either of which may initiate recall—the last rehearsed item which is assumed to be available at the I level, and the subject's internal context. In the simulations that have been presented, recall began with the last item at the I level and continued until a certain amount of time had passed with no new item being generated. At this point the context cue was used to reinitiate recall. The alternative strategy, of using context as the first retrieval cue is also a possibility, albeit not one implemented in the simulation model above. Bjork and Whitten (1974) and Hogan (1975) have presented data on output order. Both found that early and late items tend to be emitted as responses earlier in recall than do middle items. We have reanalyzed the output order data from the uncontrolled rehearsal conditions of Murdock and Metcalfe (1978). In that experiment, 20-item lists of words were presented at a 5-second rate, and immediate recall was required. Each list was classified into one of three categories—either recall began with one of the first six words, with one of the center eight words, or with one of the final six words. It was found that 66.7% of the lists began with recall of one of the final six words, 25.3% began with recall of one of the first six words, and only 8.0% began with recall of one of the middle eight words. Since there were quite a few observations in which recall began from either the beginning or end of the list we were able to analyze these two cases in more detail. The modal serial positions for each output position for recall which began in the first and last part of the list are presented in Figure 11. Since there were an unequal number of entries in each cell, the modal output serial positions are presented only if the cell had more than 12 observations. The figure shows that when recall began with one of the last six items, it was most likely to begin with the last item. The next recalled item was most likely to be the second last item. Recall then proceeded in a backward order for several items. By output position 5 recall shifted to the first item presented (as also occurs in the model because the context cue is evoked to reinitiate recall). Recall then proceeded in a forward order. It should be noted that while we have presented only the item which demonstrated the highest frequency of recall at each output position, after about output position 3, the frequency distribution becomes bimodal such that there is a tendency for a peak not only at serial positions 18 and 19, but also at serial position 1. This occurs in the model as well as the data. In those cases in which recall began with one of the first six items, it was most likely to begin with item 1 and then to proceed in a forward order. Although this method of analysis, like that of Bjork and Whitten (1974) and Hogan (1975), is not immune to the criticism that a mixture of recall orders could give rise to the results, it does show, in a rough manner, what those orders are. The systematicity of the recall order suggests that the particular cues specified in the model very likely are the cues controlling recall. It also supports the idea that associations are formed between contiguous items, as in the present model. The dotted lines on the figure represent the modal serial positions for each output position generated by the simulation when recall is initiated by the last item present at the I level. This output order corresponds reasonably well to that which was used two-thirds of the time in the data. As an exercise we also ran the simulation model with context as the cue initiating recall. In this simulation, the probability of recall of the recency items was quite low and the output order was in a forward direction.

Interresponse Times

In the model as recall proceeds, the number of items which must be checked to see if the just recalled item has already been recalled increases, as does the number of items that must be retrieved before a not yet recalled item is found. Thus an increase in interresponse time from beginning to end of recall as has been found by Roediger (1974), and by Roediger, Stellon, and Tulving (1977) is expected.

However, in the simulations of the model, recall stopped if it took longer than 3 seconds to retrieve an item. It appears that this stop rule (which, in fact was chosen rather arbitrarily) is wrong. One of the results of the particular rule used in the simulation to end recall is that while the

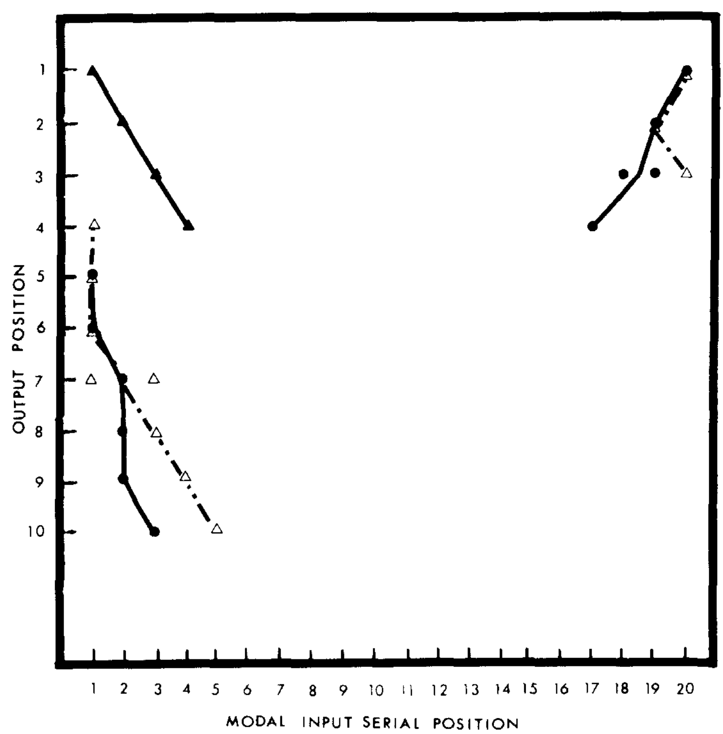

FIG. 11. Output order in Murdock and Metcalfe (1978) when recall began with one of the last six items (●——●), one of the first six items (▲——▲), and the predicted output order (△– – –△).

time spent recalling each successive item does increase, it does not increase exponentially as do the interresponse times found by Murdock and Okada (1970). The stop rule in the model imposes a ceiling on the amount of time spent for any recalled item. The mean interresponse times for each output position from Murdock and Okada are presented in Table 1 (column 1) along with the interresponse times from the model (column 2). The stop rule is the particular feature of the model that gives rise to this aberrant result since when the stop rule is modified the interresponse time results are more consistent, at least qualitatively, with the data. It seems plausible to us that there is, indeed, a stop rule in free recall, particularly when the subject is free to say that he or she has recalled as much as possible and is ready for the next list. It appears, however, that the stop rule used in the simulations does not correspond to that used by subjects.

The model also predicts that there should be a rather long pause when subjects revert to the context vector and begin recalling from the start of the list. Referring to the Murdock and Okada (1970) study, Murdock (1972) says: "After subjects have recalled the last chunk, there seems to be a discontinuity in the retrieval process." He examined cases in which subjects recalled items 17, 18, 19, 20, and then some nonterminal item. The mean interresponse times were 426, 639, 752, and *2830* milliseconds. We have reanalyzed the Murdock and Okada data, looking at the modal serial position for each output position. The order of recall revealed by this analysis is: 18, 19, 20, 20, 1, 1, 1. There is a clear shift at output position 5. Furthermore, this shift is not to just any nonterminal item but to item number 1, much as was found in the Murdock and Metcalfe (1978) experiment (see the previous section), and as would be expected if subjects switched to the context cue. The interresponse times for these modal outputs were 583, 677, 1022, *5051*, 5272, 8550. The increase in interresponse time that goes with the shift to item 1 is not attributable solely to the fact that there is a general increase in interresponse time throughout recall. In those cases in which item 20 was recalled in output position five, the interresponse time was only 1303 milliseconds. The increase in interresponse time with the shift to item 1 is more than twice as large as the *mean* increase in interresponse time between any two output positions. These interresponse time data offer support for the idea that the retrieval of the first item differs from the retrieval of other items. The difference postulated by this model is that the context vector is first covertly reconstructed before retrieval of the first item occurs.

TABLE 1
INTERRESPONSE TIME (IN SEC) AS A FUNCTION OF OUTPUT POSITION

Output position	Murdock & Okada (1970)	Model
1–2	0.7	1.6
2–3	1.1	2.8
3–4	2.3	2.6
4–5	3.6	2.7
5–6	5.3	2.5
6–7	6.3	2.2
7–8	7.6	2.4

Part–List Cuing

The storage of traces in this model is clearly what Slamecka (1968) would call "dependent." "Storage dependence means that traces are associated with or in contact with each other, so that the fate of one affects the status of another" (Slamecka, 1968, p. 505). The method of retrieval postulated in the present model depends critically on the supposition that items are associatively stored.

Slamecka argued that if dependent-trace storage were the case, then at time of recall the presentation of some of the items in the list should improve recall of the remainder of the list because some items would be accessed via their association with the experimenter-presented items. If storage were independent, the presentation of some cue words from the list would be expected to have no effect on the recall of the remaining items. In order to test for dependent or independent storage, Slamecka (1968), in a series of experiments, presented various proportions (ranging from .17 to .97) of the words in the free-recall lists to subjects as cue words. The control groups received no cues. The variable of interest was the proportion recalled based only on those words that were not cue words in the experimental conditions. The counterintuitive result which emerged from this series of experiments was that recall in the cued conditions was not only not better, but in most cases was actually worse than in the control conditions.

On the surface, this finding appears to pose difficulties for a model in which the idea of dependent storage is central. However, a small simulation of conditions similar to those in Slamecka's experiments revealed that these expected difficulties did not in fact materialize. The simulation for the control condition used a 30-word list presented at a 2-second rate, as in Slamecka's experiment. It differed from the experiment insofar as the list was presented only once (since the model is not yet designed to deal with multiple presentations). This difference resulted in a lower overall recall level than was found in the experiment, in which the 30-word list was presented twice at a 2-second rate. To simulate the experimental condition, a randomly selected half of the list was inserted at the beginning of recall, and recall proceeded, in both conditions, as usual.

In Experiment 1 (Slamecka, 1968) when half of the words in the list were presented as cues, the ratio of cued recall to uncued recall was .76 for rare words, .84 for common words, and .86 for high associates of the word *butterfly*. In Experiment 2 when the proportion of cues was .5 and the materials were rare words, the ratio of cued to uncued recall was .76; and in Experiment 4 in which the words were category exemplars, the ratio was .70. The ratio of cued to uncued from 200 runs per condition of the simulation was .77—well within the range found by Slamecka.

Why does a negative cuing effect obtain? According to the model, the subject recalls by correlating the item at the I level with the A level, which is unchanged by the correlation. During recall only one item exists at the I level and hence only one item may be used as a cue. A single item is generated from the A level which then becomes the item at the I level and is used as a cue to generate another item. The item at the I level may be recalled if it has not already been recalled. To determine if a word has already been recalled, the model checks through the previously recalled words when each item is generated. In the simulation above, the last presented item (not the last cue item) was used as the cue to initiate recall. This is an optimum strategy since if the last cue item were used to initiate recall, performance would be even worse in the cued condition. (See also the section on delay interval.) In the cued conditions the stop rule is encountered sooner than in the uncued conditions because occasionally words that were presented as cue words are recalled. Unlike the Rundus (1973) model of part–list cuing, the "strength" of associations to the cue items was not increased in the present model. It is interesting that Slamecka (1969) suggested a retrieval procedure fairly similar to the one used in the present model but did not apply it to his own earlier (1968) study.

It would be reassuring to have some independent evidence for the recall process delineated. Hogan's (1975) data offer some support. He instructed his subjects to vocalize everything that they are thinking about during recall. As in the overt rehearsal procedure, there is the problem in this technique that subjects may not say everything they are thinking. Nevertheless, Hogan found that nearly 40% of the responses that were emitted were repetitions.

The second finding of interest from Hogan's study concerns the retrieval strategy. He tested under three conditions: the sixth word in the list was presented to subjects at the start of recall; the thirteenth word was presented; or no word from the list was presented. If the presentation of intralist cues altered the retrieval strategies employed, the expectation would be that the output order in these three conditions would differ radically. Specifically, the items near in serial position to the presented item should be output first. Although a very small effect of the sort expected was obtained, "in fact, the final serial position remained the most likely initiation point in each of the three conditions" (p. 202). These results suggest that rather than using the presented items as cues to help access other items, subjects recall in the order that they normally would, regardless of the presentation of part of the list.

To summarize, although Slamecka presented a strong case against dependent storage, it is not definitive. The experiment does offer support for the idea that retrieval in free recall proceeds in a particular way—by use of the item just recalled as a cue to generate the "next" item, rather than by use of the nominal cues that the experimenter has provided.

Discussion

In this final section we shall briefly contrast the present model with some other ideas and models of free recall. Before doing so, however, some limitations of the present model should be noted. First, the representational assumptions apply only to unrelated items. At the present time we do not know how or if changing the similarity of the vectors representing items would alter the results. Obviously this is a problem of considerable interest. Second, at the present time the model extends only as far as single-list presentation. Extending the model to multiple lists poses a number of problems such as how previous lists are stored such that they are noninterfering with the current list, how list discrimination is accomplished, and what the relation is between multitrial and final free recall. Third, the model at present does not delineate how episodic recognition might occur. The above problems may be tractable within the framework proposed for single-trial free recall with some relatively simple modifications. However, it is doubtful that the model will ever be able to parse a sentence or to do arithmetic—both things that humans quite clearly can do.

We shall not attempt to compare the present model quantitatively to other models, since the parameter values were, at best, rough estimates. However, certain conceptual differences and similarities are worth pointing out. To this end we shall briefly look at generate and edit models as summarized by Watkins and Gardiner (1979), two-store models, and at semantic search models as exemplified by FRAN (Anderson, 1972).

The present model is a form of generate and recognize model insofar as referencing the A level with a retrieval cue gives rise to a generated vector which is identified at the I level. Thus, in the present model, as in other generate and recognize models (Anderson & Bower, 1972; Kintsch, 1970), there are two stages in the recall process—the generate stage and the recognize stage. The present model differs from other generate and recognize models concerning the nature of these two stages. In the modal generate and recognize model, response candidates are generated from semantic memory. The candidates are then edited in such a way that only those candidates which were recognized *as belonging to the list in question* are expressed as responses. It has been pointed out that this form of generate and recognize model cannot account for the finding of recognition failure of recallable words (Tulving & Thomson, 1973). In the present model, generation does not occur from semantic memory, but rather from the A level which consists of a cumulative memory vector formed by the combination of words that were in the list; when something is generated, the chance that that something will be correctly identified is quite high (in fact it is perfect if a sufficient number of features were encoded and are used in the identification process). The process of generation does not give rise to a number of response candidates, but to one vector which was initially associated with the cue, and therefore was from the list. Since the generated vector is not a perfect replica of the initially encoded item, but may be noisy, an identification stage is required. The identification stage does not consist of recognizing that the generated item was a member of the list, rather it consists of recognizing what the generated item *is*. The recognition stage is thus similar to the identification stage involved in the initial perception of items that are externally presented (which may also be more or less noisy). It is not necessary in the present model to recognize that an item was a member of a particular list in order for recall to occur, and hence the present model is not embarassed by the finding of recognition failure of recallable words.

The present model may also be thought of as a two-store model, although the primary memory component is minimal. If we consider that the I level is equivalent to primary memory then the maximum capacity of primary "memory" in this model is two items, since two items must simultaneously converge on that level for an association to be formed. The maximum *recall* possible from primary "memory" or the I level is one item. We do not consider the item at the I level to be in memory proper but rather to be the item of which the person is conscious—it has already been recalled but that recall has not necessarily been made overt. In order to recall a second item, it must be retrieved from the A level. It seems unlikely that a system using associations as the basis of encoding can be devised with a smaller primary memory capacity. However, postulating a larger capacity seems to us to be arbitrary, particularly since a considerable debate has ensued of late over the capacity and measurement of the capacity of primary memory (see Watkins, 1974), and recent evidence shows that the decay rate from primary memory may be much more rapid than was initially supposed (Muter, 1980). This is not to say that theorists who have assumed that the capacity of primary memory is greater than outlined in the present model are wrong. It may well be that the present model is wrong and that there is some mechanism by which items may be stored and retrieved within primary memory. Without specifying how this might occur, however, it seems more parsimonious to use the minimal assumption.

The present model bears some resemblance to FRAN (Anderson, 1972) since both are associative models of free recall, and so are constrained in some of the same ways. So far as we can determine, FRAN does not predict the rehearsal results in Figures 4 and 5, and perhaps not in Figure 3 either. However, it is difficult to know this for certain since rehearsal data are not explicitly modeled in FRAN. This is also true for the data on state dependence and output order. The recency effects generated by the two models are similar and for similar reasons (if the five item short-term buffer in FRAN is equated to the *single* item used as a cue into the end of the list in the present model). Both models predict the effects of delay on the recency effect, but for different reasons. The LIST MARKER in FRAN, like "context" in the present model, generates the primacy effect. However, the associative mechanism in FRAN does not require a LIST MARKER, whereas in the present model context is logically required—the associative process could not get started without it. Certain interesting predictions, such as the rehearsal curves and state dependence, fall out of the characterization of context as in the present model (and do not, so far as we can determine, fall out of the LIST MARKER in FRAN). The most important difference between the present model and FRAN, however, is the mechanism for association. In the present model, items are associated because a particular operation—convolution—is performed to associate them. As has been demonstrated, having performed this operation enables the inverse operation—correlation—to generate an item when the other item is given. In FRAN, items are considered associated because a tag is affixed to a preexperimental association. Even if it were possible to affix tags to associative pathways, as is assumed in FRAN, the question remains (as has been pointed out by Postman, 1975) as to how the items were initially associated. It is also not obvious to us that any relation will allow retrieval. Certain relations may allow retrieval while others do not. We do not deny the idea that any item may be associated with any other. (Convolution allows us to do just that.) Nor do we deny the idea that people can find similarities between any two items. The vectors representing the items, in the simulations, were allowed to have many features in common. However, it is not clear how having features in common or dictionary definitions in common provides the sort of relation that allows retrieval (especially when the information necessary for identification involves the features that are not shared, and also when the lists consist of unrelated words). The specification of an encoding and retrieval mechanism that *demonstrably* allows retrieval is, in our view, the principle advantage that the present model enjoys over FRAN. Some of the predictions of the two models are the same, but the conceptual underpinnings of the two are quite different.

In conclusion, we have presented a mechanism—convolution and correlation—which, as has been demonstrated, allows items to be associated and once that association has been formed allows one of the items to generate the other. The method of storage of associations is cumulative and hence there is no search process required for the retrieval of items. A process model of single-trial free recall which is compatible with the encoding and retrieval mechanism was proposed. This process model was found to account for a wide range of results.

References

Anderson, J. A. Two models for memory organization using interacting traces. *Mathematical Biosciences*, 1970, **8**, 137–160.

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

Anderson, J. R. FRAN: A simulation model of free recall. In G. H. Bower (Ed.), *The psychology of learning and motivation*. New York: Academic Press, 1972. Vol. 5.

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

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

Barret, T. W. Holography, information theory, and the cerebral cortex. *Mathematical Biosciences*, 1970, **9**, 49–60.

Bjork, R. A., & Whitten, B. W. Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology*, 1974, **6**, 173–189.

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

Bower, G. H. A multicomponent theory of the memory trace. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation*. New York: Academic Press, 1967. Vol. 1.

Bracewell, R. *The Fourier transform and its applications*. New York: McGraw–Hill, 1965.

Brodie, D. A. Free recall measures of short-term store: Are rehearsal and order of recall data necessary? *Memory & Cognition*, 1975, **3**, 653–662.

Brodie, D. A., & Murdock, B. B., Jr. Effect of presentation time on nominal and functional serial-position curves of free recall. *Journal of Verbal Learning and Verbal Behavior*, 1977, **16**, 185–200.

Brown, R., & McNeill, D. The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior*, 1966, **5**, 325–337.

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

Darley, C. F., Tinklenberg, J. R., Roth, W. T., Hollister, L. E., & Atkinson, R. C. Influence of marihuana on storage and retrieval processes in memory. *Memory & Cognition*, 1973, **1**, 196–200.

Eccles, J. C. Possible synaptic mechanisms subserving learning. In A. G. Karzman & J. C. Eccles (Eds.), *Brain and human behavior.* New York: McGraw–Hill, 1972.

EICH, J. E. State-dependent retrieval of information in human episodic memory. In I. M. Birnbaum & E. S. Parker (Eds.), *Alcohol and human memory.* Hillsdale: Erlbaum, 1977.

FLEXSER, A. J., & TULVING, E. Retrieval independence in recognition and recall. *Psychological Review,* 1978, **85**, 153–171.

GLANZER, M. Storage mechanisms in recall. In G. H. Bower & J. T. Spence (Eds.), *The psychology of learning and motivation.* New York: Academic Press, 1972. Vol. 5.

GLANZER, M., & CUNITZ, A. R. Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior,* 1966, **5**, 351–360.

HART, J. T. Memory and the feeling of knowing experience. *Journal of Educational Psychology,* 1965, **56**, 208–216.

HARTLEY, D. *Observations on man: his frame, his duty, and his expectations.* Gainesville, Fla: Scholar's Facsimilies & Reprints, 1966. (Originally published in 1749).

HOGAN, R. M. Inter-item encoding and directed search in free recall. *Memory & Cognition,* 1975, **3**, 197–209.

JOHN. E. R. A model of consciousness. In G. E. Schwartz & D. Shapiro (Eds.), *Consciousness and self-regulation.* New York: Plenum, 1976.

KINTSCH, W. *Learning, memory and conceptual processes,* N.Y.: Wiley, 1970.

KINTSCH, W., & POLSON, P. G. On nominal and functional serial position curves: Implications for short-term memory models? *Psychological Review,* 1979, **86**, 407–413.

KOPELL, B. S., ROTH, W. T., & TINKLENBERG, J. R. Time course effects of marijuana and ethanol on event-related potentials. *Psychopharmacology,* 1978, **56**, 15–20.

LASHLEY, K. S. In search of the engram. *Symposia of the Society for Experimental Biology,* 1950, **4**, 454–482.

LATHI, B. P. *Communication systems.* New York: Wiley, 1968.

MURDOCK, B. B., JR. The serial-position effect of free recall. *Journal of Experimental Psychology,* 1962, **64**, 482–488.

MURDOCK, B. B., JR. Short-term memory. In G. H. Bower & J. T. Spence (Eds.), *The psychology of learning and motivation.* New York: Academic Press, 1972. Vol. 5.

MURDOCK. B. B., JR. Convolution and correlation in perception and memory. In L.-G. Nilsson (Eds.), *Perspectives in memory research: Essays in honor of Uppsala University's 500th anniversary.* Hillsdale: Erlbaum, 1979.

MURDOCK, B. B., JR., & METCALFE, J. Controlled rehearsal in single-trial free recall. *Journal of Verbal Learning and Verbal Behavior,* 1978, **17**, 309–324.

MURDOCK, B. B., JR., & OKADA, R. Interresponse times in single-trial free recall. *Journal of Experimental Psychology,* 1970, **86**, 263–267.

MURDOCK, B. B., JR., & WALKER, K. D. Modality effects in free recall. *Journal of Verbal Learning and Verbal Behavior,* 1969, **8**, 665–676.

MUTER, P. A. How fast can forgetting be? *Memory & Cognition,* 1980, **8**, 174–177.

NILSSON, L.-G., WRIGHT, E., & MURDOCK, B. B., JR., Order of recall, output interference, and the modality effect. *Psychological Research,* 1979, **41**, 63–78.

PENNEY, C. G. Modality effects in short-term verbal memory. *Psychological Bulletin,* 1975, **82**, 68–84.

POSTMAN, L. Verbal learning and memory. *Annual Review of Psychology,* 1975, **26**, 291–335.

Pribram, K., Nuwer, M., & Baron, R. The holographic hypothesis of memory structure in brain function and perception. In R. C. Atkinson, D. Krantz, R. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology.* San Francisco: Freeman, 1974.

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

ROEDIGER, H. L. Inhibiting effects of recall. *Memory & Cognition,* 1974, **2**, 261–269.

ROEDIGER, H. L. Memory metaphors in cognitive psychology. *Memory & Cognition,* 1980, **8**, 231–246.

ROEDIGER, H. L., STELLON, C. C., & TULVING, E. Inhibition from part-list cues and rate of recall. *Journal of Experimental Psychology: Human Learning and Memory,* 1977, **3**, 174–188.

RUNDUS, D. Negative effects of using list items as recall cues. *Journal of Verbal Learning and Verbal Behavior,* 1973, **12**, 43–50.

RUNDUS, D., & ATKINSON, R. C. Rehearsal processes in free recall: A procedure for direct observation. *Journal of Verbal Learning and Verbal Behavior,* 1970, **9**, 99–105.

SLAMECKA, N. J. An examination of trace storage in free recall. *Journal of Experimental Psychology,* 1968, **76**, 504–513.

SLAMECKA, N. J. A temporal interpretation of some recall phenomena. *Psychological Review,* 1969, **76**, 492–503.

SMITH, E. E., SHOBEN, E. S., & RIPS, L. J. Structure and processes in semantic memory: A featural model for semantic decisions. *Psychological Review,* 1974, **81**, 214–241.

SPERLING, G. Successive approximations to a model for short-term memory. *Acta Psychologica,* 1967, **27**, 285–292.

STERNBERG, S. Memory scanning: Mental processes revealed by reaction-time experiments. *American Scientist,* 1969, **57**, 421–457.

TULVING, E., & THOMPSON, D. M. Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 1973, **80**, 352–373.

UNDERWOOD, B. J. Attributes of memory. *Psychological Review*, 1969, **76**, 559–573.

VAN HEERDEN, P. J. Theory of optical information storage in solids. *Applied Optics*, 1963, **2**, 393–399.

WATKINS, M. J. Concept and measurement of primary memory. *Psychological Bulletin*, 1974, **81**, 695–711.

WATKINS, M. J., & WATKINS, O. C. Processing of recency items for recall. *Journal of Experimental Psychology*, 1974, **102**, 488–493.

WATKINS, M. J., & GARDINER, J. M. An appreciation of generate-recognize theory of recall. *Journal of Verbal Learning and Verbal Behavior*, 1979, **18**, 687–704.

WAUGH, N. C., & NORMAN, D. A. Primary memory. *Psychological Review*, 1965, **72**, 89–104.

WICKENS, D. D. Characteristics of word encoding. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory*. Washington, D.C.: Winston, 1972.

REFERENCE NOTE

1. EICH, J. E. *Dissociation of memory: a state of the research-art report*. Paper presented at the Conference on Alcohol and Human Memory, Laguna Beach, California, August 1976.

(Received June 5, 1980.)