10

Repetition and Rehearsal Mechanisms in Models for Short-Term Memory

Robert A. Bjork[1]
University of Michigan

Introduction

Performance in short-term memory situations is, in general, clearly sensitive to variations in conditions of presentation and rehearsal. It is important to consider just what kinds of formal representation of rehearsal and repetition processes are reasonable and possible in models for short-term memory, both to see whether any of the possible mechanisms are supported or ruled out by existing data and, hopefully, to also make more apparent what lines of research are likely to discriminate among the possible mechanisms. This paper is an attempt to outline on the basis of intuition, logic, and current models the likely general mechanisms of repetition and rehearsal in short-term memory, to see any implications of the various mechanisms, and, through a mixture of some

[1] The preparation of this chapter was supported by the Advanced Research Projects Agency, Department of Defense, monitored by the Air Force Office of Scientific Research under Contract No. AF 49(638)-1736 with the Human Performance Center, Department of Psychology, University of Michigan.

relevant data and considerable speculation, to evaluate the relative adequacies of the various mechanisms.

It does not take more than a few minutes staring out the window to decide that the possible number of questions one might ask about repetition and rehearsal exceeds any arbitrary *n*. Furthermore, considering that items of various kinds might be presented in various ways through various sense modes at various rates and tested in various ways, that the subject may apply various rehearsal strategies to various parts of the item at various times in the retention process, and that each and every one of these variations might alter or change the performance characteristics of the repetition or rehearsal processes involved, it is reasonable to consider only the general classes of possible mechanisms and their implications, rather than to attempt the admirable but imprudent task of considering the specific members of those classes.

The infinite domain of this paper generated by the possible variations alluded to above will be drastically reduced to the point of manageability by several restrictions. These restrictions are intended to isolate, on the one hand, the experimental behavior most amenable to the study of repetition and rehearsal effects, and on the other hand, the most fundamental questions about the representation of repetition and rehearsal in this experimental situation. To other workers in the field, of course, the restrictions will seem all but completely arbitrary.

The experimental situations of primary interest are the Brown–Peterson paradigm (Brown, 1954; Peterson and Peterson, 1959) and its main variations. In this experimental situation, a single subspan item (e.g., a word trigram or consonant trigram) is presented and, after some number of seconds filled with additional presentations of the item, opportunities for rehearsal and/or an interpolated interfering activity of some kind, the item is cued for recall. This general paradigm has, in my judgment, some advantages over other procedures with respect to the questions of interest to this paper. It does not seem as complicated as those procedures that require the subject to remember items at or beyond the memory span and that are thereby heavily influenced by response output interference and other factors, nor does it seem as complicated as the various continuous procedures in which the presentation-test trial structure for an item is interlaced with the trial structure of other items, thereby producing significant and largely unfathomable item interactions. The fact that at any one point in time a subject in the Brown–Peterson paradigm is responsible for exactly one subspan item has the following desirable (simplifying) properties in assessing repetition and rehearsal effects: (a) if the item is presented again, it is reasonable to assume that the subject will attend to it rather than to some other item for which he is also responsible, (b) if there is an opportunity to rehearse, the subject can rehearse successfully whatever he remembers of the item (since the item is easily subspan) with negligible memory loss occurring during the rehearsal process itself, and (c) at the time of recall, the subject can similarly report whatever he remembers of the item with negligible loss in memory occurring in the course of the recall itself.

Organization

The remainder of this paper will be organized around two questions on the representation of repetition and two questions on the representation of rehearsal in models for short-term memory.

Repetition

(1) Should each repetition of an item be conceived of as (a) increasing the strength of a single memory trace, (b) increasing the number of traces of the item, or (c) something else? (2) What kind of model structure is required to predict the significant and complex effects of the spacing of repetitions on performance?

Rehearsal

(1) Does rote rehearsal strengthen memory traces by some process or another, or does rote rehearsal simply maintain items in storage? (2) Does a successful rehearsal of an item operate in essentially the same manner as a presentation of the item?

The Representation of Repetition

What Happens?

Though experience is psychological research shapes one to fear making assertions much stronger than obvious tautologies, it does appear safe to say that the presentation of an item builds up something in memory. The nature of the "buildup," as well as the nature of the "something," is a far less obvious matter. It might be (a) that the strength of a unitary trace of the item is increased (e.g., traditional interference theory; Wickelgren and Norman, 1966), or (b) that the number of traces or retrieval routes (Bernbach, Chapter 4) or conditional elements (stimulus sampling theory) are increased, or (c) that there is some probability that the item goes from a less permanent to a more permanent state in memory (the various multistate Markov models). Although it is clear that these three general mechanisms do not exhaust the possibilities, it is uncertain whether there are any reasonable specific mechanisms that do not fall into one of the three classes. There are, of course, possible representations of repetition which combine the properties of two or all three of the classes (Wickelgren, for example, in Chapter 3 assumes that four different memory traces with very different decay characteristics result from a single presentation).

The problem of discriminating among the three mechanisms mentioned above is formidable. There is not much in the way of relevant data; indeed, it is not clear just what constitutes relevant data. Although the specific representatives of one mechanism-class generally have different implications than the specific representatives of another class, it would be difficult if not impossible to construct a specific model of one class, which in its basic properties, could not be exactly duplicated by a model of another class. For example, with respect to predictions of simple retention curves as a function of the number of repetitions, models of different classes are not necessarily identifiable (for a discussion of the identifiability problem, see Greeno, 1967).

The problem of deriving implications that hold for all models which embody a particular repetition or rehearsal mechanism as opposed to implications that hold for most or some members of the class needs and merits elaboration. Consider Hellyer's (1962) straightforward experimental assessment of the effects of multiple presentations on retention. Hellyer presented consonant trigrams 1, 2, 4, or 8 times and measured retention following 3, 9, 18, or 27 sec of interference. His results are shown in Fig. 1.

If, in an effort to account for Hellyer's findings, one postulates a specific model based on one of the repetition mechanisms listed above, the model will in general make different predictions about the shapes of the curves in Fig. 1 than will some other specific model based on one of the other repetition mechanisms; it is not clear, however, that it is possible to state any one possible property of Hellyer's curves that is

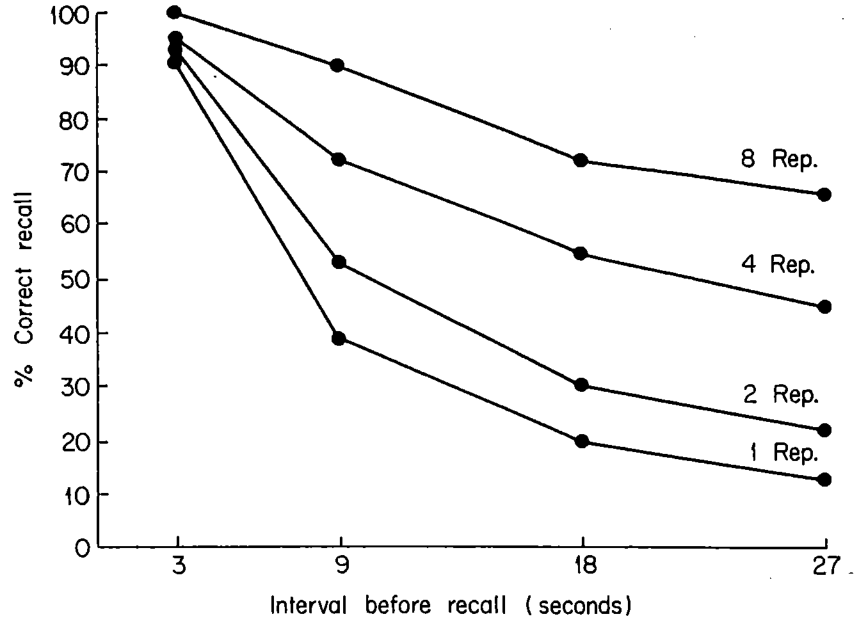

FIG. 1. Percentage of consonant trigrams recalled completely correctly as a function of the number of presentations of the trigram (Hellyer, 1962).

uniquely predicted by all specific models of a given class. To illustrate, consider the following three simple models and their general implications concerning the family of curves in Fig. 1.

Model 1

As a result of presentation, a memory trace is established to a momentary maximum strength sufficient to result in perfect immediate recall. During an interfering activity, this trace strength decreases exponentially with time and the probability that an item is recalled correctly is directly proportional to the momentary trace strength. Finally, each subsequent presentation both reinstates the trace to the maximum momentary strength and increases the resistance of the trace to interference which results in a decrease in the rate of exponential loss of trace strength with time during the retention interval.

This model predicts a family of exponential curves all of which start at unity (when there is no interference) and asymptote at zero (chance performance).

Model 2

The first presentation of an item results in a copy of the item being stored in memory. With each subsequent presentation of the item, there is some probability an additional item is stored in memory—the exact probability is some inverse function of the number of copies already in memory. During each unit of interfering time, any one copy in memory is, independent of other copies, lost from memory with some probability. Finally, if one or more copies of the item are in memory at the time of recall, the item is recalled; if no copies are in memory, the item is not recalled.

This model predicts a family of retention curves that start at unity, go eventually to zero, and become increasingly sigmoid-shaped with increasing presentations.

Model 3

An item is at any point in time in exactly one of three states: not in memory, in short-term memory, or in long-term memory. Items not in memory are not recalled, items in short-term or long-term memory are recalled perfectly. Items in long-term memory are impervious to interference, items in short-term memory are lost in an all-or-none fashion with some probability (greater than zero and less than one) during each unit of interfering time. With each presentation of an item, there are some probabilities p and $1 - p$ that the item, if not in memory, will go to the long-term and short-term states, respectively. If the item is in the short-term state when presented, there is some fixed probability it will go to the long-term state and, failing that, it will remain in the short-term state. An item in the long-term state when it is presented will remain there.

This model predicts a family of retention curves that also start at one and are geometric in shape, but go to nonzero asymptotes, the asymptotes increasing with increasing presentations.

The three models outlined above were chosen so that each of the three types of repetition mechanisms introduced at the start of this section would be represented. In evaluating the models with respect to Hellyer's data, it is clear that the obtained curves do not look like those predicted by models 1 and 3 in that they are not geometric in shape. However, although the increasingly sigmoidal curves do offer general support for model 2 (the multiple copy notion), it is not clear that the curves would go eventually to zero for retention intervals longer than 27 sec, as is predicted by model 2.

The prediction of zero asymptotes differentiates models 1 and 2 from model 3. This prediction seems to be an all but fundamental property of unitary trace strength models, if not of multiple copy models; that is, it is not clear that the concept of a trace asymptoting at some nonzero strength has much meaning, though it might be reasonable to assume that some copies of an item in memory are relatively permanent. The latter assumption, though possibly reasonable, creates a considerably more complex model falling somewhere between the present model 2 and model 3 in its properties. At any rate, Hellyer's experiment did not include long enough retention intervals to decide whether an adequate model should predict nonzero asymptotes or not, and, as Bernbach (chapter 4) stoutly asserts, one should only abandon the notion that at any multiple copies in memory are of a single type under intense pressure from the data.

The fact that Hellyer's data support model 2 over models 1 and 3 does not, unfortunately, imply that the multiple copy mechanism is therefore supported without qualification over the trace-strength and short-term, long-term store mechanisms. With a little ego involvement in either models 1 or 3 and a little practice at repairing models with additional assumptions, it is not difficult to modify either model to account for the general shape of the curves in Fig. 1. Assuming that (model 1) trace strength must decrease to some threshold before it begins to affect probability of recall, that (model 3) there is a sensory store that holds items for very brief periods after presentation, or, for that matter, complicating the assumed effects of the intervening activity in various ways would, among other modifications, make models 1 and 3 more comfortable with Hellyer's data.

Thus, the problem of stating implications that hold for all models incorporating a particular repetition or rehearsal mechanism is a very tough one, and I am not going to face up to it completely in this paper. In a sense, not facing up to the problem is demanded at this point by another consideration: existing data do not generally have the power to test any such possible unique implications. Thus, where it seems appropriate, implications of the most simple and representative members of a class will be evaluated against existing data, if data exist.

Why Does the Spacing of Presentations Matter?

There is now an abundance of evidence from Brown–Peterson-type experiments, from paired-associate experiments, from free recall experiments, from recognition memory experiments, and from other verbal learning procedures as well, that the spacing of repetitions of an item has relatively large and relatively clear effects on both latency and frequency measures of performance. The principal results of existing research on the effects of spacing in short-term memory can be summarized by three findings: (1) In general, performance is significantly better following spaced repetitions of an item than performance following massed repetitions of an item. (2) However, there is an interaction: if performance is measured after very short retention intervals, it is better to have massed repetitions. (3) And there is a limit to the improvements in performance with spacing; as the interval between two repetitions of an item is increased, performance improves to a point and then declines.

Evidence for the first of the three summary statements above is available in two experiments of the Brown–Peterson type (Peterson, 1963; Pollatsek, 1969), and in a number of experiments using paired-associate procedures (Greeno, 1964; Peterson, Hillner, and Saltzman, 1962; Bjork, 1966; Rumelhart, 1967; Bjork and Abramowitz, 1968; and others), recognition procedures (Olson, 1969: Hintzman, 1969), and free recall procedures (Melton, Reicher, and Shulman, 1966). The second finding derives from paired-associate experiments by Peterson *et al.* (1962) and Rumelhart (1967). The third finding derives from paired-associate experiments by Peterson, Wampler, Kirkpatrick, and Saltzman (1963) and Young (1966).

To give all this some sense of specificity, consider first the experiments by Peterson (1963) and Pollatsek (1969). Peterson gave subjects a first presentation of a CVC, had them count backwards for a variable period (1, 3, 6, or 11 sec), gave them a second presentation of the CVC, and had them attempt to recall the CVC after a final retention interval of 6 sec during which they also counted backwards. The obtained proportions of correct recalls are shown below.

$P_1 - P_2$ Spacing interval, seconds

1	3	6	11
.66	.67	.74	.77

Performance improves monotonically with the spacing of the two presentations.

The relevant condition in Pollatsek's experiment is somewhat more complicated. Pollatsek's condition can be denoted $P_1R_1I_1P_2R_2I_2T$, where P_1 and P_2 indicate the first and second presentations of a word trigram,

R_1 and R_2 indicate rehearsal periods of 0, 3, or 6 sec, I_1 indicates a first interference period of 6 or 21 sec, I_2 indicates a second interference period of 9 or 21 sec, and, finally, T indicates a test for recall of the trigram. Pollatsek's results are shown below collapsed over the three values of R_2 and the two values of I_2.

	$R_1 = 0$	$R_1 = 3$	$R_1 = 6$
$I_1 = 6$	.70	.80	.82
$I_1 = 21$	.80	.88	.92

The improvement in performance with increased spacing of presentations is more striking in Pollatsek's data than it is in Peterson's data. Furthermore, there is an improvement in performance with increased spacing independent of whether the increase in spacing results from an increase in the intervening rehearsal period (R_1) or an increase in the intervening interference period (I_1).

The interaction between spacing of repetitions and retention interval was first shown by Peterson *et al.* (1962) who presented subjects with a long, continuous string of study trials and test trials on paired associates. The items of interest were given two presentations separated by zero or four trials on other pairs, and were tested after an interpolated interval of one, two, four, or eight trials between the second presentation and the test. The crossed curves in Fig. 2 show that performance following a short retention interval is better if the presentations are massed and that performance following a long interval is better if the presentations are spaced. Rumelhart (1967) found an analogous interaction between the spacing interval and the retention interval using a continuous paired-associate task in which an individual pair was given six anticipation trials separated by various sequences of interpresentation intervals.

The finding that there is a limit to the improvement in performance following a long interval with increased spacing was first shown by Peterson *et al.* (1963) and was replicated by Young (1966). In both experiments, a single test trial followed the second of two study trials after a fixed number of trials (eight in the former experiment, ten in the latter experiment). Figure 3 shows their results: performance improves with spacing and then declines.

I have taken some time to clarify what I consider the most salient effects of the spacing of repetitions because I consider these effects to indicate better than anything else the sophistication required of any

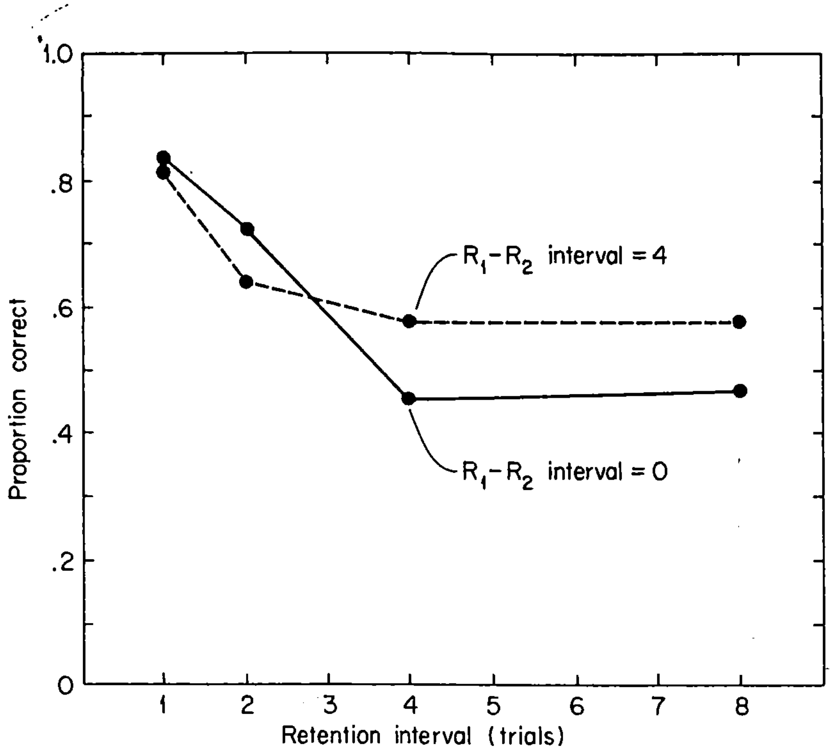

FIG. 2. Mean retention curves for different spacings between a first and second study trial (Peterson *et al.*, 1962).

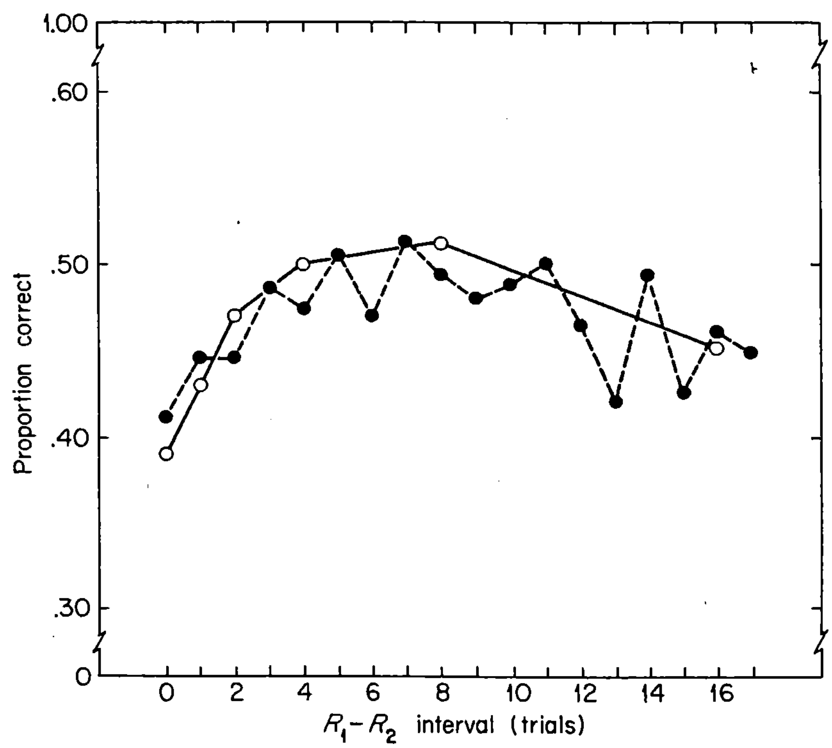

FIG. 3. Proportions of correct responses as a function of the spacing of two study trials (Peterson *et al.*, 1963; Young, 1966). ○—○ Peterson *et al.*; ●----● Young.

repetition mechanism proposed to account in general for existing data. Also, I feel that the data from experiments on spacing of repetitions may have enough power to discriminate among fairly complicated mechanisms. By themselves, the three effects of spacing just summarized rule out a number of otherwise reasonable repetition mechanisms, including, in my view, nearly any mechanism based on the concept of a single memory trace varying in strength along some continuum.

A consideration of the constraints on an adequate theoretical representation of the effects of repetition imposed by the single finding that performance improves with the spacing of presentations will serve both to defend and to clarify the preceding assertions. Any model predicting that the repetition of two different items operates on their respective memory traces in such a way as to preserve at a fixed time after repetition the order of their respective recall probabilities at the time of repetition is ruled out by the spaced vs. massed practice finding. That is, performance following spaced presentations is better than performance following massed presentations, in spite of the fact that the probability of recall measured at the time of the second massed presentation is considerably higher than performance measured at the time of the second spaced presentation. Thus, an adequate model must admit circumstances under which the repetition of one item can be so effective that the probability of its recall might be greater than that of a second item (which was also repeated), even though the probability of recall of the first at the time of repetition might have been a great deal less than that of the second item. Any model that is strictly order preserving in the sense that it always predicts that a repetition will transform the memory for any two items, I and I', with probabilities of recall $P < P'$ at the time of repetition, into memory states at some subsequent time resulting in probabilities of recall, $r(P) < r(P')$, is in trouble. In particular, any model based on an assumption that a repetition increments the current strength or number of traces of an item by some proportion of the difference between the current strength or number and some maximum strength or number seems untenable.

There has now been considerable work addressed to the question of what kind of model structure is required to predict the effects of spacing (Greeno, 1967; Bjork, 1966; Rumelhart, 1967; Landauer, 1969; Pollatsek, 1969). Although most of this work has dealt with paired-associate data, it appears from existing data, that the same general spacing phenomena occur in the Brown–Peterson paradigm as occur in paired-associate paradigms. I will therefore step out with some confidence on the assumption that the theoretical work mentioned above is directly relevant to memory as exhibited in the Brown–Peterson paradigm.

Without going into all the reasons why other notions have been ruled out, it now appears that there are three extant classes of models that have some chance to account for the effects of spacing. The first is a class of Markov models characterized by their Markov properties and the general assumption that an item can be in one of essentially three states of knowledge: not in memory (a guessing state), short-term memory, or long-term memory (permanent). The individual members of this class are defined by various assumptions about transition probabilities which are in turn interpretations of various notions about coding processes, state properties, attentional processes, and the like.

The second class of models are the buffer models of Atkinson and Shiffrin (1968). These models have at their core the notions that (a) there is a fixed capacity rehearsal buffer which operates as a kind of push down store holding a few items at a time and largely determining the properties of the short-term storage system of which it is a part, (b) there is a long-term store out of which items are retrieved imperfectly due mainly to the operation of the traditional mechanisms of associative interference, and (c) the probability an item is represented in long-term storage and/or the strength of its representation in long-term storage are a function of the time it resides in the rehearsal buffer before it is displaced. The last notion constitutes a mechanism through which memory traces consolidate over time, a notion shared with a number of other models (e.g., Landauer, 1969; Norman and Rumelhart, Chapter 2; Wickelgren, Chapter 3; Judith Reitman, Chapter 5); there is considerable disagreement, however, as to the specific nature of the consolidation process.

The third class of models derives from stimulus sampling theory and is based on the notion of stimulus fluctuation (Estes, 1955; Izawa, 1967). These models assume that an item to be remembered together with the context in which it is presented comprise a set of stimulus elements, some of which are active at any one time (momentarily available to the subject) and some of which are inactive (momentarily unavailable to the subject). Over time, there is assumed to be a random fluctuation of any one stimulus element between the active and inactive states.

Rumelhart (1967) has contributed a very thorough theoretical and empirical analysis of these model classes with respect to a continuous paired-associate task. His principal findings are two: (1) that a model he terms the modified GFT (General Forgetting Theory)—which is a modified version of a model suggested and tested by Bjork (1966), which, in turn, is a generalization of models proposed by Atkinson and Crothers (1964), Greeno (1967), and others—accounts very well for his data; (2) that there are specific versions of both the Buffer models and the Stimulus Fluctuation models that are not distinguishable from the modified GFT in his experiment.

Rumelhart's data, as well as Bjork's (1966), offer general support for the notion first suggested by Greeno (1967) that an item already in the short-term state when presented has a negligible probability of going to the long-term state. This notion, in a nutshell, is why these models predict a spaced practice effect. When items are given an immediate or near-immediate repetition, they are either already in the learned state or in the short-term state at the time of their repetition. Hence, the immediate repetition accomplishes little or nothing compared to a spaced repetition which occurs at a time when the item, if not already in the long-term state, is likely to be in the forgotten state from which the probability of transition to the long-term state is relatively high. Although the notion that the probability of transition from the short-term to the long-term state is very small leads to the prediction of a spaced practice effect, it does not predict the interaction between spacing interval and repetition interval shown in Fig. 2.

The modified version of the GFT model suggested by Rumelhart predicts both a spaced practice effect and the interaction shown in Fig. 2. The modification consists of assuming that there is a fixed probability on any trial that a subject will not attend to the item presented, in which case the item will end the trial in the same state of knowledge in which it began. This assumption results in there being a higher probability that an item will be in the short-term state immediately following the second of two massed presentations than there is that the item will be in the short-term state immediately following the second of two spaced presentations. That is, even if the subject does not attend to the second of two massed presentations the item is still likely to be in the short-term state from the first of the two presentations, whereas if the subject does not attend to the second of two spaced presentations the chances are negligible that the item is still in the short-term state from the first of the two presentations. Hence, for suitable parameter values, the modified GFT predicts that performance following a short retention interval is better if preceded by massed presentations than if preceded by spaced presentations because the higher probability of being in the short-term state at the time of testing more than compensates for the lower probability of being in the long-term state.

Unfortunately the modified GFT, and the formally equivalent versions of the buffer and stimulus fluctuation models, are significantly less than perfect. There are at least three problems. First, the assumption that the subject does not always attend to the item presented is suspect. The modified GFT underpredicts the performance on a test trial immediately following a single presentation of the tested item: for typical parameter values it predicts that the probability of a correct response will be between .85 and .95, whereas the data are typically between .95 and 1.00. One might add to the model structure the assumption of a very short-term sensory storage system in order to predict essentially perfect performance on any test occurring within a few seconds of a presentation, but it is likely that the required holding time of the assumed sensory store would well exceed any reasonable expectations derived from existing research on such very short-term sensory storage. Another interpretation of Rumelhart's assumption that subjects do not attend to some items as they are presented which seems more reasonable to me, although it does not solve the present problem, is that the subject chooses on any trial whether to attempt to encode the item in some fashion or not. Although, intuitively, I doubt whether subjects ever (for all practical purposes) fail to perceive an item presented in the normal fashion, I think it is quite likely in continuous procedures that subjects choose on some trials not to make any real storage effort with the item presented.

A second problem is that the modified GFT does not predict the finding (Fig. 3) that there is a limit to the improvement in performance with increased spacing of presentations, i.e., beyond some optimal spacing there is a decline in performance. The model predicts that performance improves monotonically with increases in the spacing of presentations. It may be that one should not assume that the long-term state is permanent, but rather, semi-permanent, with a very small probability of loss from storage relative to that of the short-term state.

A third problem, more tentative than the first two, but also, if it holds up, more formidable, derives from an experiment by Bjork and Abramowitz (1968). This experiment employed a continuous paired-associate task in which some pairs had the following sequence of four anticipation trials. The first presentation (P_1) was followed by x intervening trials on other pairs until the second presentation (P_2). The third presentation (P_3) followed P_2 after y intervening trials, and, finally, the fourth presentation (P_4) followed P_3 after z intervening trials. The following values of x, y, and z were used to generate spacing sequences: $(x,y) = (0,20)$, $(2,18)$, $(10,10)$, $(18,2)$, $(20,0)$ and $z = 2, 8, 20$. That is, $x + y$ was always equal to 20 intervening trials. The principal question of interest was whether the x and y interpresentation intervals were commutative in their effects on performance at P_4. For example, the only difference between the spacing sequences $x = 0$, $y = 20$, $z = 2$ and $x = 20$, $y = 0$, $z = 2$ is the order of the first two interpresentation intervals. Some models predict commutivity in such a case, others do not.

In general, Bjork and Abramowitz found commutivity of the x and y intervals in their effect on performance at P_4. Also, the modified GFT gave a very good quantitative account of the observed performance. However, for the parameter values they estimated, the model does not predict the interaction of spacing interval and retention interval shown in Fig. 2. For the parameter values estimated by Rumelhart (1967), who employed the same general procedure as did Bjork and Abramowitz except for the choice of spacing sequences, the modified GFT predicts the interaction shown in Fig. 2 but does not predict commutivity. Thus, the model can predict the interaction in Fig. 2 and it can predict commutivity, but it can not predict both simultaneously.

An experiment designed to include both the spacing conditions necessary to test for commutivity and those necessary to test for an interaction between repetition interval and retention interval could resolve whether the modified GFT is inadequate or whether, in fact, the two effects do not occur together in paired-associate learning. Should it be possible to obtain both effects simultaneously, a considerable challenge is posed to existing models: no current model, as far as I know, predicts both commutivity and an interaction between repetition interval and retention interval.

The effects of the spacing of repetitions have been discussed in some detail because, as mentioned earlier, their complexity and consistency seem to impose severe constraints on an adequate theoretical representation of repetition in models of short-term memory learning. It seems safe to say that the principal empirical effects of the spacing of presentations on performance rule out some simple repetition mechanisms in models of short-term memory; it seems even safer to say that they do not obviously imply any particular repetition mechanism.

The Representation of Rehearsal

Does Rote Rehearsal Strengthen or Maintain?

There are certainly some among the readers of this paper who are a bit put off by my implicit distinction between rote rehearsal and rehearsal that consists primarily of a search for an effective mnemonic of the to-be-remembered item. Not only do some researchers think that such a distinction has little meaning, but some (e.g., Neisser, 1967, p. 239) feel that it may be misleading to think of rehearsal activity as formally distinct from interfering verbal activities of various kinds. Beyond mentioning the issue, I want to ignore it and to assume that rote rehearsal, i.e., repeating an item over and over to oneself, is a distinct, well-defined activity that subjects find reasonable and commonly do. I do so because (a) it agrees with my introspections and, more importantly, (b) instructions to subjects "repeat the item over and over to yourself" and "try to associate or relate or image the item with or to something you already know" have differential effects on performance.

Brown (1958) seemed to assert, though he has since said that he didn't mean it (personal communication), that the main function of rehearsal is to maintain items in short-term memory and not to somehow strengthen the memory trace of items. In order to investigate this issue, I tested subjects' memory for five two-digit numbers under the four conditions schematized in Fig. 4. Conditions I and III were designed to

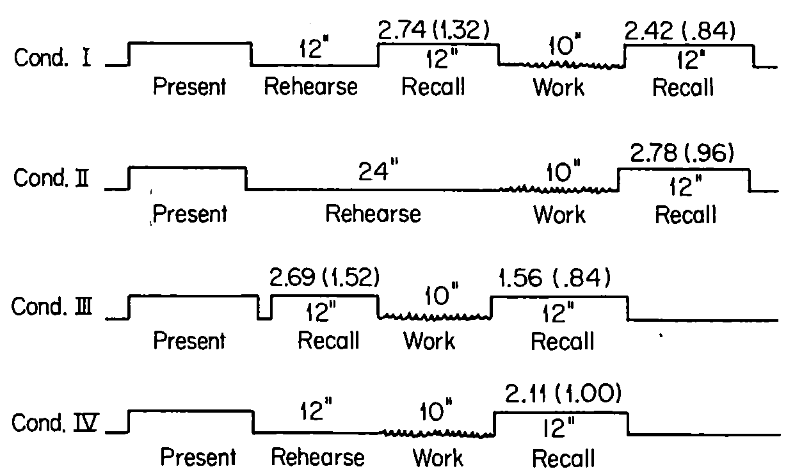

FIG. 4. Mean number of two-digit items recalled correctly in the correct position out of five possible. (The values in parentheses are from a replication of the experiment conducted at the University of Michigan by D. Raymond and L. Hannah in which the subjects were children and four rather than five two-digit numbers were presented on each trial.)

assess by comparison with each other whether a 12-sec rehearsal period prior to a first test for recall (condition I) would make the items recalled on the first test more resistant to the interfering activity interpolated between the first test for recall and the second test for recall relative to the case (condition III) in which there was no such opportunity to rehearse.

A comparison of performance on the final recall between conditions I and III and between conditions II and IV (Fig. 4) illustrates the strengthening effect of the rehearsal periods in conditions I and II. Also in Fig. 5, the higher conditional probability of a correct recall on test two given a correct recall on test one in condition I as compared to condition III illustrates clearly that items recalled on the first test following the 12-sec rehearsal period in condition I are more resistant to the interfering activity than are items recalled on the first test immediately after presentation in condition III. Conditions II and IV

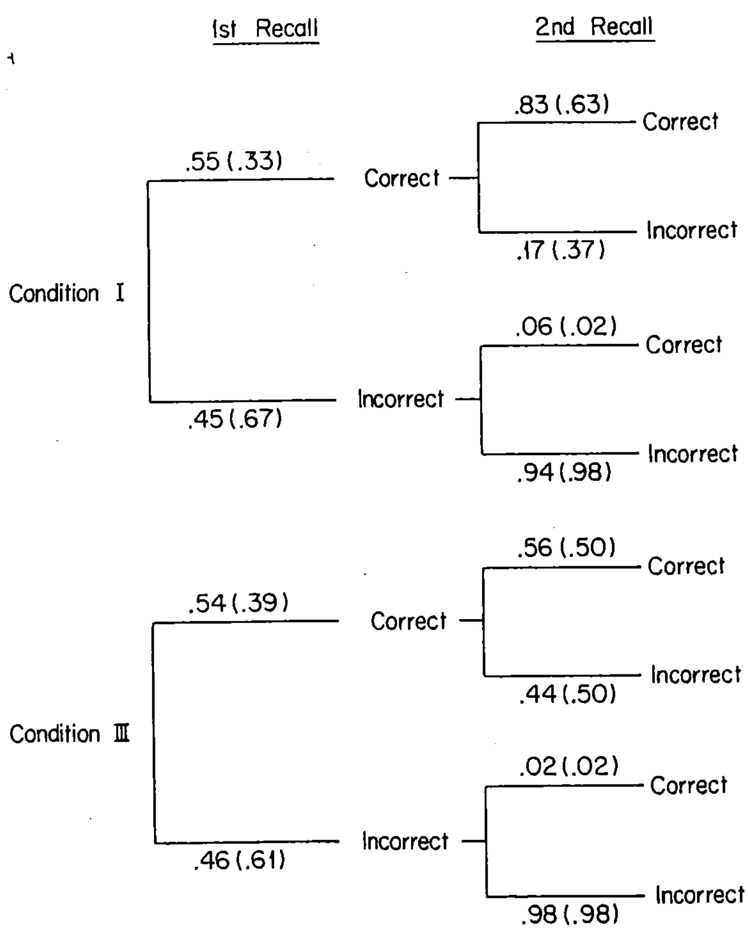

FIG. 5. Conditional probabilities of a recall or nonrecall on the second test given a recall or nonrecall on the first test in Conditions I and III, Fig. 4.

were designed to permit, by comparison with conditions I and III, respectively, an assessment of the relative efficiencies of covert rehearsal and overt recalls in making items resistant to interference, but that, though important, is a different problem (see, e.g., Brelsford and Atkinson, 1968).

There are other relevant data that might be offered to support the contention that rote rehearsal strengthens something, but there doesn't seem to be much question about that, and the question of what is strengthened and how is a good deal more difficult.

Rehearsal: What Happens?

There are several reasonable ways in which rehearsal might act to increase the resistance of a memory trace to interference; they are roughly the same as for repetition. Rehearsal might increase the strength of a single memory trace, it might increase the number of traces of the item, or it might accomplish a change from short-term storage to long-term storage in the state of the item.

Although the ways in which rehearsal and repetition might operate on memory seem similar, it is unlikely that rehearsal and repetition operate in exactly the same fashion to increase the resistance of a memory trace to interference. For example, although rehearsal is less "perfect" than repetition (an item might be lost from memory during the rehearsal period) and, therefore, one might expect that n one-second repetitions would result in better short-term memory than a single presentation followed by an n-second rehearsal period, it might also be the case that a rehearsal period permits more freedom for the subject to achieve a long-term encoding of some kind than does a series of repetitions and, hence, might result in better performance over long retention intervals. In fact, exactly that difference between forced repetitions and covert rehearsal is strongly suggested in the results of a free recall experiment by Glanzer and Meinzer (1967). In this experiment, there was a 3.2-sec interval between the presentations of any two successive words in the list during which, in one condition, subjects were required to give the word just presented six vocal repetitions and, in another condition, subjects were free to rehearse covertly in any fashion they wished. Glanzer and Meinzer found that the latter condition resulted in clearly better recall performance for words presented at the beginning and middle of the list (long retention intervals) than did the former, but that there was no difference in the recall of the items presented at the end of the list (short retention intervals).

There are a number of other possible differences between rehearsal and repetition processes that might be elaborated, but given that there has been no really systematic experimentation comparing the two, there seems little reward in such speculation.

Some hints as to the nature of the rehearsal process as well as some indications of possible differences between the operation of rehearsal and the operation of repetition are provided in the experiment by Pollatsek (1969) discussed earlier. One condition of Pollatsek's experiment was designed to investigate the role of rehearsal in a manner analogous to that employed by Hellyer in the study of repetition effects (Fig. 1) reported earlier. Pollatsek gave subjects a single presentation of a word trigram followed by rehearsal periods of 0, 3, 6, or 9 sec and tested for recall after interference (counting backwards) periods of 0, 3, 6, 9, 15, or 21 sec following the rehearsal. His results are shown in Fig. 6.

The retention curves in Fig. 6 appear to become increasingly sigmoid-shaped with increasing rehearsal and seem to be decreasing to non-zero asymptotes, though longer interference intervals would be necessary to say for sure. In comparison to Hellyer's results (Fig. 1), (a) performance is better overall in Hellyer's experiment (Hellyer used consonant trigrams and Pollatsek used word trigrams) and (b) performance at short

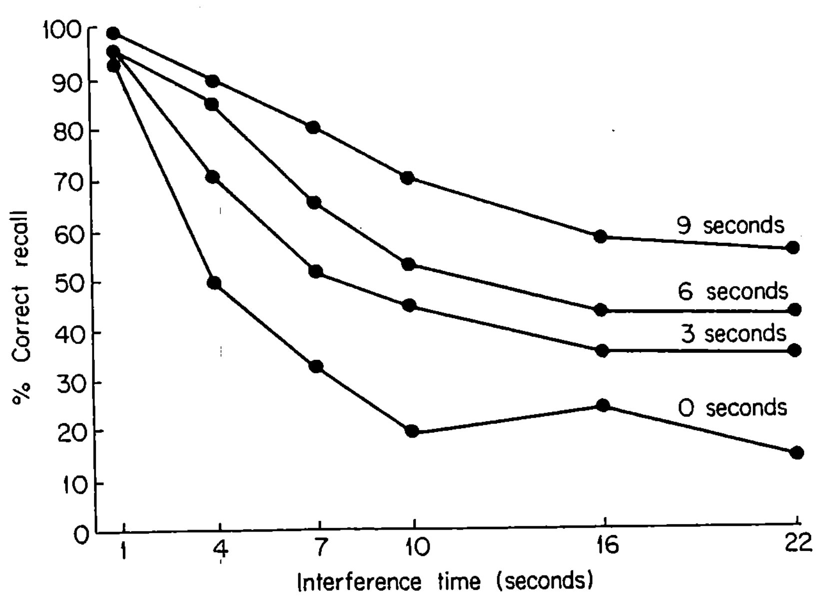

FIG. 6. Proportions of word trigrams recalled completely correctly as a function of rehearsal time (Pollatsek, 1969).

retention intervals in Pollatsek's experiment appears relatively worse and performance at long intervals relatively better than in Hellyer's experiment. The second comparison provides some support for the notion mentioned above that a rehearsal period might be less beneficial on the short term and more beneficial on the long term than are multiple repetitions.

To the degree that the curves in Fig. 6 are in fact approaching non-zero asymptotes, the notion that rehearsal benefits memory by increasing the chances that the item is processed into a relatively permanent storage state is supported. However, if that were the only effect of rehearsal, one would expect that the curves would go to different asymptotes but would otherwise have the same general shape. The systematic change in the shape of the retention curves in Fig. 6 with increasing rehearsal argues that rehearsal affects the short-term storage of an item as well as the long-term storage. As was pointed out earlier for Hellyer's curves, the fact that the curves become progressively more sigmoid-shaped is consistent with the notion that rehearsal adds additional copies of the memory trace to short-term memory. A system combining the multiple-copy and short-term to long-term storage rehearsal mechanisms would give rise to curves such as those in Fig. 6. It might be that during rehearsal the subject can both build up the number of copies of the item in short-term memory (rote rehearsal) or try to achieve a long-term encoding of the item (nonrote rehearsal), and that he does some of each.

Concluding Comments

In contrast to most of the papers in this volume, it has not been the goal of this paper to suggest or test any particular theoretical model of human memory. Rather, this paper represents an attempt to formulate some questions about the formal representation of repetition and rehearsal in models of human memory and to consider the likely alternative representations with respect to those questions in the light of existing data.

At some points during this endeavor, I have felt quite sure that it was overdue and quite worthwhile; at other points, I have been equally convinced that it was premature and frustrating. On the one hand, the current intense theoretical activity seems to be sprouting a number of attractive models (the chapters in this book are good examples). There are large, reliable effects on performance in several different memory paradigms with variations in the conditions of presentation or rehearsal, and it seems only a matter of looking at existing data and theories very carefully to see the general shape of an adequate theoretical representation of repetition and rehearsal. On the other hand, it also seems that for every experimental finding that clarifies some aspect of human memory there is one that questions the integrity of some generally accepted view of another aspect. It is often difficult to derive differential predictions from alternative models, even when they have apparently quite different underlying psychological assumptions. Even when alternative models do make differential predictions, the crucial data are not available.

Depending on how one looks at it, one can see a picture of the human memory system developing bit-by-bit or, one can see a confusing, complex collage of experimental findings, theoretical notions, and unanswered questions. I feel that both views are at least in part veridical and that they together characterize the present state of model building in the area of human memory.

In the remainder of this section, I want to summarize and conclude this paper by stating and elaborating three assertions with respect to the theoretical representation of rehearsal and repetition.

(1) Increases in either the number of presentations of an item or in the time allotted for rehearsal of the item effect both the short-term storage of the item and the long-term storage of the item. Any reasonable doubts that repetition or rehearsal of an item build up the memory storage of the item should be dispelled by Figs. 1 and 6. However, it is at least logically possible that only the short-term storage or only the long-term storage of an item but not both would be strengthened by repetition or rehearsal. Although such systems are logically possible, they appear quite unlikely on the basis of the families of curves in Figs. 1 and 6. In both figures, there appear to be systematic changes in the shapes of the retention curves over short retention intervals as a function of increasing repetitions or rehearsal time, as well as substantial differences in performance at long retention intervals.

(2) The clear but complex effects on performance of variations in the spacing of repetitions of an item indicate the considerable sophistication required of an adequate theoretical representation of repetition. As mentioned earlier, however, the data from spacing of practice experiments may also have enough power to allow one to choose among fairly

sophisticated alternative representations. It would be an impressive achievement to formulate a model that would provide a satisfactory quantitative account of the complex effects of spacing.

As pointed out earlier in this chapter, there are several models that hold some promise of ultimately providing an adequate quantitative account of the spacing of practice phenomena. In particular, the basic Markov structure (termed the General Forgetting Theory) accounts quite well for the first order quantitative effects of spacing during paired-associate learning (Bjork, 1966; Rumelhart, 1967). It might be that modifying the General Forgetting Theory by complicating the short- and long-term state representations within the theory would result in an improved quantitative correspondence between the theory and spacing data that would justify the increase in complexity. For example, one might assume that retrieval from long-term memory is less than perfect or that it is possible to have multiple copies of an item in short-term memory. Although I don't agree with Bernbach (chapter 4) that his single-memory multiple-copy model is any more parsimonious than multiple-memory single-copy models of the General Forgetting Theory type—it seems only a matter of where one chooses to assume complexity within the model structure—his work may well indicate that Markov models have oversimplified the representation of an item within any one type of memory store.

Wickelgren, in Chapter 3, indicates in a general way how his multi-trace strength theory could account for each of the principal effects of spacing summarized earlier in this chapter. Whether Wickelgren's multitrace system could account quantitatively and simultaneously for the effects of spacing remains to be seen.

(3) Although rehearsal and repetition probably operate in the same general way to increase the resistance of a memory trace to interference, they probably do not operate in exactly the same way.

The possible experimental conditions of rehearsal and repetition during a fixed time interval following a presentation of an item vary considerably in terms of how much the subject's behavior is under the experimenter's control. In the case that the subject is free to rehearse covertly, there is very little control; in the case that the subject is required to shadow multiple, rapid repetitions overtly during the interval, there is considerable control. It may well be that the covert rehearsal case provides a better opportunity for the subject to achieve a relatively long-term encoding of the item than does the multiple repetition case which, in a sense, intrudes on the subject. Hence, a period of covert rehearsal is likely to result in a relatively better performance after long retention intervals than is a period of multiple repetitions. The converse, however, is likely to be true for performance following short retention intervals.

The whole problem of formulating theoretical representations of rehearsal and repetition and characterizing any formal differences between the two should be helped considerably by the various current experimental efforts to specify the role of encoding strategies, imagery, grouping strategies, retrieval strategies, etc. in human memory. Hopefully, if this chapter were to be written several years from now, it would be possible to feel quite confident that slowly but surely a picture of the human memory system was developing.

References

ATKINSON, R. C., AND CROTHERS, E. J. A comparison of paired-associate learning models having different acquisition and retention axioms. *Journal of Mathematical Psychology*, 1964, 1, 285–315.

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

BJORK, R. A. Learning and short-term retention of paired-associates in relation to specific sequences of interpresentation intervals. Technical Report No. 106, Institute for Mathematical Studies in the Social Sciences, Stanford University, 1966.

BJORK, R. A., AND ABRAMOWITZ, R. L. The optimality and commutivity of successive interpresentation intervals in short-term memory. Paper presented at the meeting of the Midwest Psychological Association, Chicago, May, 1968.

BRELSFORD, J. W., JR., AND ATKINSON, R. C. Short-term memory as a function of overt and covert rehearsal procedures. *Journal of Verbal Learning and Verbal Behavior*, 1968, 7, 730–736.

BROWN, J. The nature of set to learn and of intra-material interferences in immediate memory. *Quarterly Journal of Experimental Psychology*, 1954, 6, 141–148.

BROWN, J. Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology*, 1958, 10, 12–21.

ESTES, W. K. Statistical theory of spontaneous recovery and regression. *Psychological Review*, 155, 62, 145–154.

GLANZER, M., AND MEINZER, A. The effects of intralist activity on free recall. *Journal of Verbal Learning and Verbal Behavior*, 1967, 6, 928–935.

GREENO, J. G. Paired-associate learning with massed and distributed repetitions of items. *Journal of Experimental Psychology*, 1964, 67, 286–295.

GREENO, J. G. Paired-associate learning with short-term retention: Mathematical analysis and data regarding identification of parameters. *Journal of Mathematical Psychology*, 1967, 4, 430–472.

HELLYER, S. Supplementary report: Frequency of stimulus presentation and short-term decrement in recall. *Journal of Experimental Psychology*, 1962, **64**, 650.

HINTZMAN, D. L. Recognition time: effects of recency, frequency, and the spacing of repetitions. *Journal of Experimental Psychology*, 1969, **79**, 192–194.

IZAWA, C. Function of test trials in paired-associate learning. *Journal of Experimental Psychology*, 1967, **75**, 194–209.

LANDAUER, T. K. Reinforcement as consolidation. *Psychological Review*, 1969, **76**, 82–96.

MELTON, A. W., REICHER, G. M., AND SHULMAN, H. G. A distributed practice effect on probability of recall in free recall of words. Paper read at the meeting of the Psychonomic Society, St. Louis, October 1966.

NEISSER, U. *Cognitive psychology*. New York: Appelton-Century-Crofts, 1967.

OLSON, G. M. Learning and retention in a continuous recognition task. *Journal of Experimental Psychology*, 1969, **81**, 381–384.

PETERSON, L. R. Immediate memory: Data and theory. In C. N. Cofer and B. S. Musgrave (Eds.), *Verbal behavior and learning: Problems and processes*. New York: McGraw-Hill, 1963.

PETERSON, L. R., HILLNER, K., and SALTZMAN, D. Time between pairings and short-term retention. *Journal of Experimental Psychology*, 1962, **64**, 550–551.

PETERSON, L. R., AND PETERSON, M. J. Short-term retention of individual verbal items. *Journal of Experimental Psychology*, 1959, **58**, 193–198.

PETERSON, L. R., WAMPLER, R., KIRKPATRICK, M., AND SALTZMAN, D. Effect of spacing presentations on retention of a paired-associate over short intervals. *Journal of Experimental Psychology*, 1963, **66**, 206–209.

POLLATSEK, A. Rehearsal, interference, and spacing of practice in short-term memory. Technical Report, Human Performance Center, University of Michigan, 1969.

RUMELHART, D. E. The effects of interpresentation interval on performance in a continuous paired-associate task. Technical Report No. 116, Institute for Mathematical Studies in the Social Sciences, Stanford University, 1967.

WICKELGREN, W. A., AND NORMAN, D. A. Strength models and serial position in short-term recognition memory. *Journal of Mathematical Psychology*, 1966, 3, 316–347.

YOUNG, J. L. Effects of intervals between reinforcements and test trials in paired-associate learning. Technical Report No. 101, Institute for Mathematical Studies in the Social Sciences, Stanford University, 1966.