JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR **15**, 1–16 (1976)

Monotonic and Nonmonotonic Lag Effects in Paired-Associate and Recognition Memory Paradigms

ARTHUR M. GLENBERG

University of Wisconsin

In Experiment I the relationship between response recall and the spacing of repetitions (lag), as a function of the retention interval, was investigated in the continuous paired-associate paradigm. At the short retention intervals (2 and 8 events) the lag function was nonmonotonic. At the longer retention intervals (32 and 64 events) the lag function increased monotonically. A version of encoding variability theory was used to explain these results. The theory was then tested in Experiments II and III. In the second experiment, using the Brown–Peterson paradigm, the lag function was monotonic for uncued recall, and nonmonotonic for cued recall. In the third experiment, using the continuous recognition memory procedure, the lag function was a nonmonotonic function of the lag interval between the first two presentations when the interval between the second and third presentations was short. Increasing the latter interval produced a lag function that was again monotonic. The results of the experiments support the theory which emphasizes the nature of the cues available for retrieval.

Defining the conditions that determine the effectiveness of a repetition is a major concern for the psychology of learning and memory. In the past five years this concern has been most evident in the study of the intervals between the presentations of to be remembered items. As the interpresentation interval, or lag, increases, recall typically shows a corresponding increase. This lag effect has been the subject of extensive experimentation using the free recall procedure (D'Agostino & DeRemer, 1973; Gartman & Johnson, 1972; Madigan, 1969; Melton, 1970). The characteristic finding in free recall is that the lag function, relating the interpresentation lag to the proportion recalled, is monotonically increasing and negatively accelerated. This form of the lag function is also the predominate result when using the Brown–Peterson distractor method (Peterson, 1963).

The lag effect has also been investigated using the continuous paired-associate (CPA) procedure. The results from the CPA procedure are quite different from those found in free recall. Peterson, Hillner, and Saltzman (1962) reported that items repeated at Lag 0 (zero items between the two presentations) were recalled better than items repeated at Lag 4 after a short (two events) retention interval between the last presentation and the test. However, after a longer retention interval (four events) the Lag 4 items were recalled better than Lag 0 items. This interaction has

The first two experiments are based on a dissertation submitted to the University of Michigan in partial fulfillment for requirements of the Ph.D. Gratitude is expressed to the members of my committee for their aid and guidance: Arthur W. Melton (Chairman), Robert A. Bjork, Louis Jensen, and Robert G. Pachella. That research was completed while the author was supported by a USPHS Traineeship. Experiment I was funded by the Air Force Office of Scientific Research, under Contract No. F44602-C-0038. Experiment II was funded by a grant from the National Graduate Student Dissertation Grant, administered by the Horace H. Rackham Graduate Student Research Fund. Experiment III was supported by USPHS Grant No. 1-R01-MH26643-01 to A. M. Glenberg. Requests for reprints should be sent to Arthur M. Glenberg, Department of Psychology, Charter at Johnson St., University of Wisconsin, Madison, Wisconsin 53706.

yet to be satisfactorily explained. Even at longer retention intervals (eight to 10 events) the lag effect generated by the CPA procedure is apparently quite different from that found in free recall. Peterson, Wampler, Kirkpatrick, and Saltzman (1963) reported that the lag-function was nonmonotonic; increasing the lag beyond eight events led to decrements in recall. This finding was confirmed by Brelsford and Atkinson (reported in Atkinson & Shiffrin, 1968) and by Young (1971). These various interactions seem to provide an excellent testing ground for the theories offered to explain the lag effect.

The intent of Experiment I is to extend our empirical knowledge of the lag effect as found in the CPA paradigm. In particular, the lag function appears to be sensitive to changes in the retention interval. However, the retention interval has been manipulated over a relatively narrow range, from two intervening events to about 10 intervening events. Increasing the retention interval further may produce further changes in the lag function.

At least two theories make definite predictions as to the shape of the lag function as the retention interval is increased. One theory is the Atkinson and Shiffrin (1968) rehearsal buffer theory. In this theory the probability of recalling the response in a paired-associate situation is related to the number of rehearsals afforded the pair. When the pair is first presented it is entered into a limited-capacity rehearsal buffer with a probability of α. Pairs are rehearsed for the duration of their stay in the buffer. If an item is not entered into the buffer it is effectively forgotten immediately. Once an item is in the buffer its representation in Long-Term Store (LTS) is being strengthened by rehearsal. The rehearsal continues until the item is replaced by another item entering the buffer. After an item leaves the buffer its strength in LTS begins to decrease by a constant proportion for each intervening event before the item is tested. On the recall test the response is recalled if the item is still in the buffer (after short retention intervals), or if the item can be retrieved from the LTS. Retrieval from LTS is an exponential function of the item's strength.

When an item is repeated there is additional opportunity for rehearsal which increases the strength of the item in the LTS, increasing its recallability over the items presented just once. The amount of extra strength attributable to the repetition is dependent on the interpresentation lag. At very short lags the item is likely to still be in the buffer from its first presentation. Since the theory does not allow for multiple representations in the buffer, the repetition would be completely ineffective. As the repetition lag increases it becomes increasingly likely that the representation of the item from its first presentation has left the buffer. In this case the second presentation is effective because the item is reentered into the buffer (again with probability of α) and the rehearsal process begins anew. With a very long delay between presentations the item will have left the rehearsal buffer and its representation in the LTS will begin to decay before the second presentation. In effect, the theory predicts recall to be a nonmonotonic function of the repetition lag, with the items repeated at moderate lags being recalled the best. The moderate lag items are recalled better than the short lag items due to the increased number of rehearsals in the lag interval. The moderate lag items are recalled better then the long lag items due to the smaller amount of decay during the lag interval.

After the second presentation, and after the representation of the item has left the buffer (which is independent of the repetition lag), recall is based solely on the items' strengths in LTS. This strength decreases at the same rate for all items. As such, as the retention interval increases, the same nonmonotonic ordering of the lag values is expected. However, since the LTS strength is decreasing to zero, the lag function as a whole becomes flatter. This prediction is presented graphically in Atkinson and Shiffrin (1968, Fig. 12).

The second theory which predicts a nonmonotonic lag function at moderate retention intervals is the General Forgetting Theory (GFT) as modified by Sperber, Greenfield, and House (1973). The GFT is a Markov model consisting of four memory states: a long-term state (L) from which recall is perfect and there is no forgetting; a short-term state (S) from which recall is perfect, but from which there is considerable forgetting; a state of item familiarization (F); and an unlearned state (U). In the latter two states recall is at the guessing level. The complete model consists of two transition matrices: a learning matrix which is applied whenever an item is presented, and a forgetting matrix applied during the lag and retention intervals. All items are assumed to start in U. When the item is presented for study it may move to L, S, or F. Upon application of the forgetting Matrix the items in L remain; the items in S may remain or enter F; the items in F may remain or enter U. Items in U remain until another presentation.

When the probability of a transition from S to L is less than the probability of a transition from F to L the model predicts an increasing lag function. Items that are presented, enter S, and remain in S (during a short lag interval) are less likely to enter L on their second presentations than items that originally enter S and then F (during a longer lag interval).

When the probability of a transition from U to L is less than the corresponding probability from F to L the model predicts that moderate lag items are recalled better than items repeated at long lags. The very long lag items may first enter S and then move to F during the lag interval. However, the item may then move to U during the remainder of the lag interval. The moderate lag items are recalled better than the short lag items because of their transition from S to F during the lag interval and the greater probability of moving from F to L than from S to L. The long lag items are recalled less often than the moderate lag items because of their transition from F to U during the long lag interval, and the smaller probability of a transition from U to L than from F to L.

As the retention interval increases the same nonmonotonic ordering of the lag values is predicted by the GFT. In contrast to the rehearsal buffer theory, the GFT predicts a slight accentuation of the curvature in the lag function with increases in the retention interval. After moderate retention intervals the curvature is vitiated by recall from State S (which is independent of the repetition lag). After very long retention intervals, when recall is almost completely from State L, the curvature is most pronounced. When there is no forgetting from State L the nonmonotonicity is predicted to remain at all retention intervals.

Experiment I utilizes the CPA procedure, varying both the repetition lag and the retention interval. The results may be used to discriminate between the GFT and the rehearsal theory. The GFT predicts a nonmonotonic lag effect that becomes more curved as the retention interval increases, while the rehearsal theory predicts that the nonmonotonicity becomes less pronounced.

Experiment I

Method

Subjects. The subjects were 108 women drawn from the University of Michigan's Human Performance Center subject pool. Each subject received $2.00 for participating in the experiment.

Materials and design. Three CPA presentation orders were designed. Each order was composed of 500 events (presentations and tests). The first 20 events were buffer items used to absorb any effects due to the build-up of proactive interference. Afterward, pairs were repeated at lags of 0, 1, 4, 8, 20, and 40 intervening events. These pairs were then tested 2, 8, 32, or 64 events after their second presentation. There were five exemplars at each lag-retention interval combination. The repetition lag was allowed to vary ±2 events for the nominal lag-20 pairs, and ±4 events for the nominal lag-40 pairs. The retention interval was allowed to vary ±3 events for the nominal retention interval of 32 events and ±5 events for the 64-event retention intervals. The items presented once were tested after 1, 2, 8, 16, 32, or 64 events. Each once presented condition was represented by 10 exemplars. Groups of 36 subjects were run on each presentation order. Over the 36 subjects, a given pair was used once in each lag-retention interval combination, and twice in each once presented condition.

The pairs were composed of common four-letter nouns. They were constructed to avoid common pre-experimental associations, rhymes, and orthographic similarities. For a given subject each word served in only one pair.

Procedure. The subjects were tested individually. Each subject was read general instructions, and then was presented with a practice series of 15 events (using five-letter nouns) presented on index cards. Following the practice series the subjects were encouraged to ask any questions they might have had about the procedure. The 500 events in the main series were then presented. The subjects were not informed that the first 20 events were buffer items.

Each event was presented on a cathode ray tube controlled by a PDP-1 computer. Each event was visible for 3 sec. On presentations, the stimulus and response words appeared simultaneously, separated by a hyphen. The stimulus word presented with a question mark signaled a test. The subject responded orally, with the response, during the 3 sec that the stimulus and the question mark were visible. The response was recorded by the experimenter.

Results and Discussion

The proportion recalled of the once presented pairs was .54, .30, .32, .25, and .15 for the 1, 2, 8, 16, 32, and 64 event retention intervals, respectively. The proportion of the repeated words recalled are shown in Fig. 1. The recall of the repeated words was analyzed in a 3 (presentation structures) ×6 (lag intervals) ×4 (retention intervals) analysis of variance, with the last two factors being within subject factors.

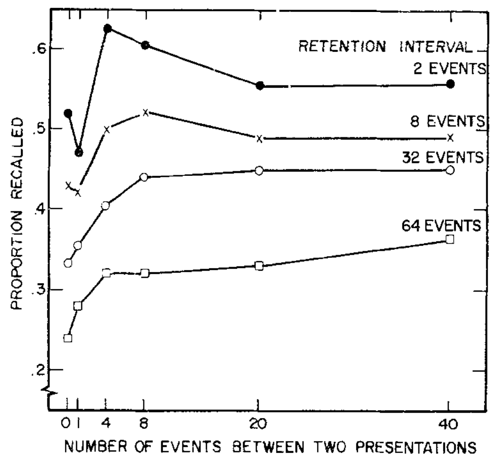

FIG. 1. The proportion recalled of the response terms, as a function of the number of events between the two presentations of a repeated item (lag interval), and the number of events between the second presentation and the test (retention interval).

The effect of interpresentation lag was significant, $F(5, 525) = 22.27$, $p < .005$, as was the retention interval, $F(3, 315) = 133.87$, $p < .005$. The effect of list structure was not significant, nor did it interact with any of the other factors. The critical interaction is the lag by retention interval interaction. This interaction was significant, $F(15, 1575) = 1.68$, $p < .05$. This indicates that the shapes of the lag function change with changes in the retention interval.

The lag effect at each retention interval was further analyzed in a one-way analysis of variance on the six lag values. At the two-event retention interval the overall effect was significant, $F(5, 535) = 8.34$, $p < .01$, as well as the orthogonal quadratic, $F(1, 107) = 5.52$, $p < .05$, and cubic, $F(1, 107) = 16.60$, $p < .01$, trends. The linear component was not significant, $F(1, 107) = 1.61$, $p > .05$. At the eight-event retention interval the linear, quadratic, and cubic orthogonal components were all significant, $Fs(1, 107) = 5.98$, 6.35, and 10.73, respectively, all $ps < .05$. For the 32-event retention interval the linear, quadratic, and cubic orthogonal components were significant, $Fs\,(1,107) = 17.79$, 10.43, and 5.96, respectively, all $ps < .05$. For the 64-event retention interval only the linear component was significant, $F(1, 107) = 28.85, p < .01$.

The results from the two- and eight-event retention intervals are in good agreement with previous research. At the shortest retention intervals the Lag 0 items are recalled better than the Lag 1 items. This is similar to Peterson et al. (1962). This result is transitory, however, in the sense that items repeated at slightly greater lags are eventually recalled better than the Lag 0 items. To complicate matters even further, with even longer lags, recall again decreases. Ignoring the Lag 0 items, the lag function after a two-event retention interval is similar in shape to the lag function after an eight-event retention interval. The lag function produced by an eight-event retention interval in the present study is similar in absolute level and shape to that reported by Peterson et al. (1963), and by Young (1971).

The most surprising results were found at the two longer retention intervals. At both retention intervals the lag effect is monotonic, and shows the negative acceleration characteristics of the lag effect as found in both free recall and the Brown–Peterson paradigm.

The results of Experiment I are inconsistent with both the GFT and the rehearsal theory. Neither theory predicts a shift in the lag function from nonmonotonicity to monotonically increasing, with increases in the retention interval. Nor does it appear likely that either theory can be simply modified to predict this change. However, a version of encoding variability theory does seem capable of predicting at least the gross results of Experiment I, that is, that the lag function is nonmonotonic at short retention intervals, and tends towards monotonicity as the retention interval increases.

Encoding variability theory is based on the notions of stimulus encoding variability as developed by Estes in his stimulus sampling theories (1955), and as modified by Martin (1968) and Bower (1972). The major assumption behind these types of theories is that stimuli may be variably encoded, that is, the memory representation stored or accessed (functional stimulus) may change, although the same physical stimulus (nominal stimulus) is presented. Changes in the encoding are due to changes in the context. The concept of context is multidimensional and refers to such aspects of the experimental situation as the other stimuli, the experimental task, the experimental room, the subject's strategies, and the subject's internal physiological and psychological states. As the context changes the subject's perception or elaboration of the stimulus changes, resulting in changes in the functional stimulus. It is also assumed that the amount of change in the context is positively correlated with the amount of change in the functional stimulus.

In the paired-associate task, at the time of presentation the subject must store a representation of the stimulus-response pair (Greeno, 1970). The representation that is stored (if any) will depend on the encoding of the stimulus and the response at the time of presentation, in other words, on the input context. At the time of testing, the subject's task is to retrieve the stored representation when given the nominal stimulus as a retrieval cue. The subject's success at the test will depend on (*a*) whether a representation was stored at the presentation, and (*b*) how closely the encoded functional stimulus at the time of the test matches the encoded version of the stimulus that was used at input. The probability of recalling the response depends on the degree to which the two functional stimuli correspond.

When a pair is repeated the effectiveness of the second presentation depends on the interpresentation lag. At very short lags there will have been no change in the context, and no change in the encoding. Whatever was stored on the first presentation will be duplicated with no additional learning. At the time of testing, the subject will be able to generate the response only if he can generate an encoding similar to the one encoding used at input. As the interpresentation lag increases there will be corresponding changes in the context which introduce changes in the functional stimulus. At the second presentation the subject will be able to store a different representation of the stimulus-response pair. In this case, at the time of testing, the subject will have a greater opportunity to retrieve the response. The response will be recalled if the subject's encoding of the stimulus is similar to the encoding used at the first or the second presentation. Thus as the lag increases the context changes which induces changes in the functional stimuli which, in turn, increases the probability of retrieval which results in the lag effect.

The theory, as outlined, will predict a monotonically increasing lag function which asymptotes after some lag representing maximal context change. One additional consideration allows the theory to predict the results of Experiment I. The functional stimulus that is employed for retrieval will depend on the context present at the time of the test. After a short retention interval the testing context will be very similar to the context at the second presentation. If the context at the second presentation (and the test) is very different from the context at the first presentation, then the representation stored at the first presentation will play a minor role in recall, that is, it is unlikely to be similar to the functional encoding used at the test. Therefore, items presented at long interpresentation lags will be functionally similar to items repeated at short lags (where there is only a single representation stored). However, the items repeated at moderate intervals will have two representations that are similar to each other, and also similar to the functional stimulus on the test. These items should be recalled more often than either the short or long lag items.

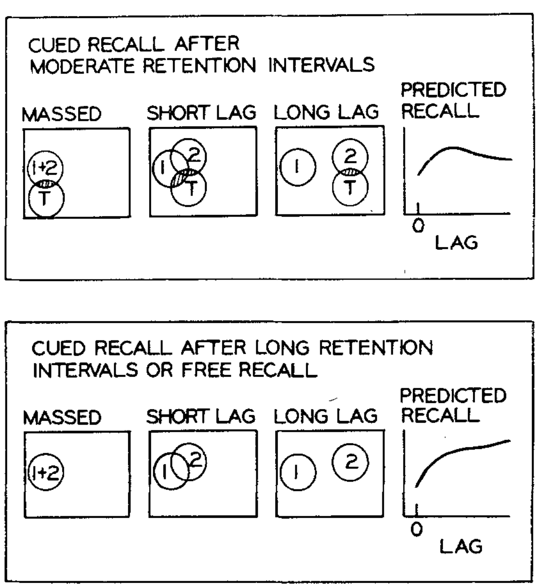

FIG. 2. Upper panel: Predicted probability of recall in the CPA task, at moderate retention intervals, as a function of the number of events between two presentations of the item. Lower panel: Predicted probability of recall in the CPA task with long retention intervals as a function of the number of events between the two presentations of an item.

The upper panel of Fig. 2 displays this prediction graphically. The three small panels represent the space upon which a single stimulus can be encoded. The circles represent the encoding that was realized on its first presentation, its second presentation, and on the test. As the interpresentation lag increases (from left to right), the encodings obtained on the presentations become less similar. At a constant short retention interval (the upper panels), the encoding on the test bears the same relationship to each of the last presentations. However, the degree of similarity between the encoding on the test and the input encodings (as represented by the cross hatching) is, on the average, a nonmonotonic function of lag.

The predictions of the theory in regard to long retention intervals are displayed in the lower panels of Fig. 2. As with the short retention intervals, increasing the repetition lag results in a decrease in the similarity of the input encodings. The test encoding is not represented in the lower panels. Since there is a long delay between the second (and for that matter the first) presentation and the test, the context at the time of the test will no longer be related to the input contexts except through random fluctuation. Therefore, the encoding at the time of the test will not be consistently related to either the first or the second input encoding. The circle representing the encoding at the test may be placed anywhere in the space. On the average, the encoding at the test is most likely to be similar to *some* input encoding when the input encodings are as dissimilar as possible (at the longest lags). In this case recall is a monotonic function of the interpresentation lag. A mathematical base for these predictions is presented in the Appendix.

Experiment II

The results of Experiment I were explained by assuming that increases in the retention interval decreased the similarity of the stimulus as encoded at the time of the test, to the stimulus as encoded at the two presentations. If this general theory is correct, it should be possible to produce monotonic and nonmonotonic lag effects by directly manipulating this similarity, while holding the retention interval constant. Providing the subject with retrieval cues having the relationship to the encoded information depicted in the upper panel of Fig. 2 should result in an inverted U-shaped spacing function regardless of the retention interval. Providing the subject with retrieval cues of a more amorphous sort should result in a monotonic spacing function, regardless of the retention interval.

This prediction was tested in a modified Brown–Peterson paradigm. The subjects were presented with three pairs of words to learn. After the retention interval the subjects were asked to recall the response terms of the three pairs. This recall was either cued by the three stimulus terms, or was uncued. It was assumed that, after a moderate retention interval, the cued recall would reflect the same processes that underlie the recall in the CPA task, after moderate retention intervals. Thus cueing should produce the nonmonotonic spacing function. In the uncued tests the subjects can only use the general context of the trial to aid recall. This retrieval cue should bear no consistent relationship to the encoded information.[1] The results in this condition should be similar to free recall and the long retention interval conditions in Experiment I.

[1] In Experiment I it was assumed that the context at the test is related to the context at the second presentation. The test context then influences the encoding of the stimulus which acts as the retrieval cue for the response. Predictions for the cued condition of Experiment II follow the identical reasoning. In the uncued condition it is assumed (due to lack of concrete evidence (ef. Falkenberg, 1972)) that it is something in the context itself which acts as the retrieval cue. Whereas the context at the time of the test is similar to the context at the time of the second presentation, the relationship between the context at the test (the retrieval cue) and the encoded information would appear to be very tenuous, as required in the lower panel of Fig. 2.

Method

Subjects. A total of 35 male and female subjects, drawn from the University of Michigan Human Performance Center paid subject pool, was used in the experiment. Eleven of these subjects were eliminated due to equipment failures. Each subject was paid \$2.00 for participating in the experiment.

Materials and design. A single trial of a repeated item consisted of a ready signal presented for 2 sec, the simultaneous presentation of three pairs of words for 5 sec, and a lag interval of 0, 10, or 30 sec. Following the lag interval was a second presentation of the three pairs for 5 sec, a 10-sec retention interval, and an 8-sec recall interval. During the lag and retention intervals the subject was required to read aloud a four-digit number, digit by digit, with a different number being presented each second. When the pairs were presented once, the trial had the same basic structure; however, the recall interval always came 10 sec after the presentation. On cued recall trials the subject was presented with the three stimulus words, simultaneously, for the entire 8-sec recall interval. The uncued recall intervals were indicated by three groups of four Xs in the window of the memory drum.

Two presentation sequences were constructed. Each sequence was composed of 64 trials. The first five trials were buffer trials to absorb the build-up of proactive interference. The next 54 trials contained the main experimental events. There were 12 trials on which the pairs were repeated after each lag of 0, 10, 30 sec of digit naming. Six of these trials were cued, and six were uncued. The 18 single presentation trials were distributed throughout the middle 54 trials. Nine of these trials were cued, and nine were uncued. The last five trials were buffer trials. The two sequences were exactly alike except that the cued trials were changed to uncued trials, and the uncued trials changed to cued trials from Sequence 1 to Sequence 2. Each sequence was used with 12 subjects. Over the 12 subjects a given pair appeared in each lag-cueing condition equally often.

The stimuli were the pairs used in Experiment I. Care was taken to avoid any obvious semantic, acoustic, and orthographic relations between the three pairs presented together on a trial.

Procedure. The subjects were tested individually. The instructions emphasized that the subject should try to learn the material as three pairs, not as six words. They were told that their responses on the cued recall test would be counted as correct only if they remembered the response and the correct pairing with the stimulus. The subjects were also told that the pairs would be reordered on each presentation, and the stimuli would be reordered on the cued recall test. During the cued recall interval the subjects were to pronounce the stimulus and then the response that was paired with that stimulus. The experimenter recorded the response and the pairing. Over 95% of the responses recalled on the cued recall tests were paired with the proper stimulus. On the uncued tests the subjects were instructed that the task required only the recall of the three response terms. However, if they were uncertain as to whether a word was a stimulus or a response, they were encouraged to guess.

Following the instructions the subjects were given three practice trials. After any questions were answered the series of 64 trials was presented. The subjects were not informed that the first and last five trials were buffer trials.

The events were presented to the subject in the window of a high-speed memory drum (change time less than .05 sec). The events were timed by a paper tape reader which controlled the memory drum.

Results and Discussion

The proportions recalled are given in Table 1. These proportions were analyzed in a 2 (presentation sequence) $\times$ 2 (cued v. uncued) $\times$ 3 (lag) analysis of variance, the last two

TABLE 1
MEAN PERCENT RECALL IN THE MAIN CONDITIONS OF EXPERIMENT II

Condition	Repeated pairs Lag 0	10	30	Once presented pairs
Cued	65.4	81.0	77.1	44.9
Uncued	45.1	62.5	69.9	29.6

being within subject factors. The effects of cueing, $F(1, 22) = 50.42$, lag, $F(2, 44) = 51.63$, and the critical cueing by lag interaction, $F(2, 44) = 7.91$, were all significant, all $ps < .005$. The only other significant effect was the cueing by presentation sequence interaction, $F(1, 22) = 4.74$, $p < .05$. While cued recall was better than uncued for both presentation sequences, the effect was larger in one sequence than the other.

The results of Experiment II have three major implications. First, they provide a strong empirical bridge between the paired-associate learning paradigms and the Brown–Peterson paradigm. Specifically, the inverted U-shaped spacing function was produced in the Brown–Peterson paradigm. Second, the results implicate the retrieval process as a major component in the production of the spacing effect. Since the subjects never knew whether recall was to be cued or uncued there can be no storage, consolidation, rehearsal, or habituation differences that can account for the results of the experiment. The only difference between the cued and uncued conditions was the relationship of the retrieval cues to the encoded information. Third, the results provide strong support for the theoretical conceptualization presented in Fig. 2.

EXPERIMENT III[2]

The theory, as outlined above, predicts the lag effect with the assumption of encoding variability. The important processes are those concerned with stimulus encoding; learning the response plays a minimal theoretical role. In Martin's (1968) terms, we can think of paired-associate learning as involving two stages. The first stage, the encoding stage, establishes the functional stimulus. The second stage involves associating the response to this stimulus. According to the theory, the lag effect and changes in the lag function are due to the first stage of learning. Martin (1968, 1972) has also suggested that this first stage may be tapped by a recognition test. If the encoded stimulus changes from the first to the second presentation, then the learner should be less likely to recognize the stimulus on the second presentation.

Applying this reasoning, the theory developed in terms of paired-associate learning can be extended to the recognition paradigms. As the lag between the first and second presentations increases, the subject should be less likely to recognize the stimulus on the second presentation. Of course, this prediction is based on the reasoning that as the lag increases the context changes, inducing changes in the encoded functional stimulus. Recognition on the third presentation will be dependent on the lag between the first and second presentation (lag interval), and the interval between the second and third presentation (retention interval). Increasing the lag interval decreases the similarity between the representations stored or tagged at the two presentations. After a short retention interval the lag function should be nonmonotonic, while after a long retention interval the lag function should be monotonically increasing. These predictions are derived from the theory in exactly the same manner as for Experiment I. According to the theory the processes that produce the lag effect are exactly the same in both the CPA and recognition memory paradigms.

The effect of the spacing of repetitions has been investigated in the domain of recognition memory. These investigations have produced

[2] Many thanks are due Mark Wolf who constructed the presentation sequences, ran the subjects, and keypunched the data.

results that are, at least, not contradictory to the present theory. Kintsch (1966) varied the lag between repetitions of the items in a continuous recognition memory task (CRM, Shepard & Teghtsoonian, 1961). The probability of recognition on the second presentation was inversely related to the lag interval. On the latter presentations the probability of correct recognition was positively related to the lag interval (when the retention intervals were held constant across the lag conditions). Hintzman (1969) reported an analogous finding using recognition time as the dependent variable. The correct recognition time increased with lag on the second presentation, while the correct recognition time decreased, with increases in the lag, on the third presentation, after a constant retention interval. Neither of these studies reported nonmonotonic lag effects, even though relatively short retention intervals were used. However, this is not damaging to the theory. In Kintsch's experiment only two lag intervals were used so that nonmonotonicity could not be detected. In Hintzman's experiment the lag intervals ranged from 0 to 16 items with a retention interval of 16 items. Extrapolating from the results of Experiment I it is very possible that this range of lag values was too narrow to detect any nonmonotonicity.

Ciccone and Brelsford (1974) employed the CRM paradigm using lags of 0 to 50 items, with a constant 25-item retention interval. Their results indicated that the lag function was nonmonotonic, peaking at a lag of 12 items. While this result is completely in accord with the theoretical predictions it remains slightly suspect. In the experiment the subjects were required to overtly rehearse the items during their presentation. The nonmonotonic lag effect was only found in their variable rehearsal condition where the subjects were required to rehearse two items, the one being shown and the other of their choice. No lag effect was found in the constant rehearsal condition where the subjects were only permitted to rehearse the item being presented. Aside from this interaction, the results are difficult to interpret. The overt rehearsal requirement has been shown to interact with other independent variables (Fischler, Rundus, & Atkinson, 1970), and to be an imperfect indicator of processing (Craik & Watkins, 1973; Einstein, Pellegrino, Mondani, & Battig, 1974). In light of these considerations, it seemed advisable to provide additional tests of the theory in the CRM paradigm.

Method

Subjects. The subjects were 60 men and women drawn from the Introductory Psychology classes at the University of Wisconsin-Madison. Each subject received course credit for participating in the experiment.

Materials and design. The stimuli were 255 consonant trigrams having Witmer association values of 0–25% (as reported in Underwood & Schulz, 1960). Of these, 240 were used in the main experimental sequence. The range of frequency of use of the 19 first letters was 7–16. For the initial bigrams, 83 were used once, 45 twice, 14 three times, 5 four times, and 1 (CX) was used five times. The 240 trigrams were divided into two sets of 120, having comparable distributions of association value and first letter frequencies.

The main sequence of events consisted of the presentation of 480 stimuli. The main sequence was constructed from two subsequences of 240 events each. Of these 240 events, 60 were items that were presented once, while 60 items were presented three times each. In total, half the events were new (the subject had never seen them before in the experiment) and half the events were old. The second presentations of the repeated items occurred at lags of 0, 1, 8, 20, and 40. These events were then given a third presentation at retention intervals of 8, 32, and 64 events. The long lag and retention intervals were allowed to vary within the same limits as in Experiment I. Each subsequence contained four observations at each lag-retention interval combination, while the total experiment contained eight observations, per subject, at each combination.

One half of the subjects were presented with one subsequence before the other subsequence while the assignment was reversed for the other half of the subjects. For one half of the subjects in each of these groups, one of the sets of 120 trigrams was assigned to the once presented conditions, and the other set was assigned to the repeated conditions. The remaining subjects received the reverse assignment. Finally within each group of 15 subjects, each subject received a different assignment of trigrams to the 15 lag-retention interval conditions, and a different assignment of the once presented trigrams to serial locations within the list. The result was that each subject received a unique list, with the assignment of trigrams to conditions completely counterbalanced.

Each trigram was typed on a separate standard size punched card. Next to each trigram were the numbers 1–6 which the subject used to indicate his judgment of old or new, and his confidence in that judgment. The stimulus cards were held in place by a specially built card holder.

Procedure. The subjects were run in groups of one to three. Upon entering the experimental room the subjects were seated in individual, partially enclosed booths. The presentation decks were randomly assigned to the subjects. The subjects were then instructed that every 4 sec they were to turn over a stimulus card, examine the next trigram, and then to make a judgment as to the trigram's oldness. If the subject believed the trigram was new he was to circle a number on the card from 1 to 3, 1 indicating greatest confidence new. If he believed the trigram to be old he was to circle a number from 4 to 6, 6 indicating the greatest confidence old. The 4-sec intervals were indicated to the subject of a sequence of .5 sec tones previously tape-recorded.

The session began with a 10-event (five new and five old) practice deck using trigrams. Following a brief, second instruction period, the subjects began to work on a sequence of 500 events. The first 20 events in this sequence (10 new and 10 old) were not scored. These events were used to absorb the initial buildup of false positives.

Results and Discussion

Discounting the first 20 events there was a total of 14 400 first presentations. The false positive rate was .41, with an average confidence of 3.1. The false positive rate increased over the first three sets of eight new items (.30, .36, .45) and was relatively stable thereafter. A 2 (assignment of sets of 120 trigrams) × 2 (subsequence order) × 5 (lags) × 3 (retention intervals) analysis of variance was performed of the subject's responses to the second presentations. The dependent variable was the number of old (4, 5, or 6) responses out of a total possible of eight. Sequence order was significant, $F(1, 56) = 10.15, p < .01$. Sequence order did not interact with any other variable and will not be considered further. The only other significant effect was lag, $F(4, 224) = 117.13, p < .01$. The probability of saying old on the second occurrences were .95, .83, .73, .67, and .62 for lags 0, 1, 8, 20, and 40, respectively. The absence of any significant effects involving the retention interval indicates that the items within a given lag value were essentially identical at the time of the second presentation.

Two similar analyses of the second occurrences were performed using the arcsine transformation of the percent old responses and the average confidence ratings as dependent variables. The pattern of results was identical in all three analyses, with the exception that the main effect of sequence order was not significant in the analysis of the confidence ratings, $F(1, 56) = 3.17, p > .05$. The average confidences for the second presentations were 5.73, 4.98, 4.45, 4.19, and 3.91 for lags 0, 1, 8, 20, and 40, respectively.

The proportion old for the third presentation lag-retention interval combinations is presented in Fig. 3. Three analyses of variance,

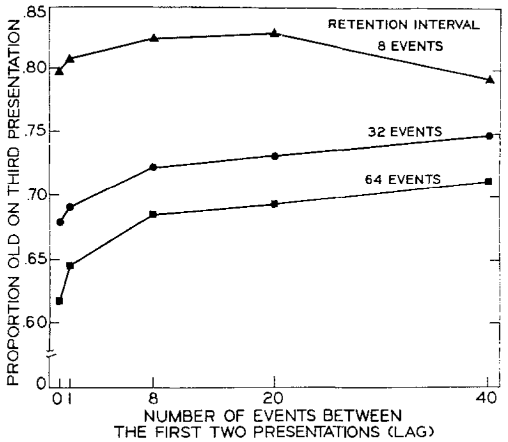

FIG. 3. The proportion of items called old as a function of the number of events between the first two presentations (lag interval) and the number of events between the second and third presentations (retention interval).

identical to those used for the second presentations, were performed. Since the pattern of results was similar, unless indicated, only the results from the number old analysis will be reported. Again, sequence order was significant, $F(1, 56) = 5.54$, $p < .05$; however, it did not interact with any other variable. The lag variable was significant, $F(4, 224) = 6.65$, $p < .01$, as was the retention interval, $F(2, 112) = 76.10$, $p < .01$. The critical term is the lag by retention interval interaction. As can be seen in Fig. 3 the lag functions displayed the predicted curvatures. However, the overall interaction was not significant, $F(8, 448) = 1.24$, $p > .05$. The eight degrees of freedom in the interaction sums of squares (12.66) were further partitioned by the use of orthogonal contrasts, specifically the linear and quadratic components of the interaction were tested against their appropriate error terms. The weights were derived and tested as suggested by Meyers (1972). The linear component of the interaction was significant $F(2, 112) = 4.21$, $p < .025$, while the remainder of the sums of squares, 1.02, distributed with six degrees of freedom was not significant. Similar results were found for the arcsin analysis, $F(2, 112) = 3.81$, $p < .025$, accounting for 90% of the interaction sums of squares, and for the analysis of the confidence ratings, $F(2, 112) = 3.92$, $p < .025$, accounting for 82% of the interaction sums of squares.

Clearly the monotonic lag effect is present at the two longer retention intervals. At the eight-event retention interval the greatest difference (between Lags 20 and 40) is only 3.5%. A one-way analysis of variance on the five lags at the eight-event retention interval was performed. Only the quadratic component of the lag effect was significant for the confidence ratings, $F(1, 56) = 5.47$, $p < .05$, accounting for 93% of the lag sums of squares, and for the arcsin transformation of the proportion old $F(1, 56) = 4.30$, $p < .05$, accounting for 98% of the sums of squares. For the analysis of the number old the F-ratio was not quite significant, $F(1, 56) = 3.13$, $p < .1$, accounting for 92% of the lag sums of squares.

The results of Experiment III are supportive of the theory outlined above. At the short retention intervals the lag function is nonmonotonic, while at longer retention intervals the function is monotonically increasing. The results are also in accord with two more general propositions. First, the lag effect seems to be solely dependent on stimulus encoding which in turn effects the probability of retrieving a previously encoded functional stimulus (and its response in the CPA paradigm). In addition, the results present clear evidence for equating the processes used in the CRM paradigm with the first, stimulus encoding stage of paired associate learning (Martin, 1972). Of course, there are some differences between the results of Experiments I and III. At the eight-event retention interval the CPA lag function peaked at Lag 8, while the CRM lag function peaked at Lag 20. At the 32-event retention interval the CPA lag function appeared to asymptote at Lag 20, while the CRM lag function increased to Lag 40. Because of the many differences in the experiments (procedurally and in the stimuli), and the relatively small size of the effects, it would not appear to be fruitful to try to explicate these differences at this time.

General Discussion

Empirically the results are clear. Across three different verbal learning paradigms, utilizing both words and consonant trigrams, the shape of the lag function has depended on the conditions at the time of testing. In Experiments I and III the lag function was nonmonotonic at short retention intervals, and monotonic at long retention intervals. In Experiment II the lag function was nonmonotonic after cued recall, while monotonic after a relatively uncued recall. In all three experiments the nonmonotonicity was on the order of 3 to 5%. However, the result was consistent across experiments and statistically significant in each experiment.

While the degree of nonmonotonicity is small, it is not as important, theoretically, as the finding that the lag function is monotonic at the longer retention intervals. At the 64-item retention intervals the increase in performance from Lag 0 to Lag 40 was on the order of 12% in Experiments I and III. In Experiment II the improvement from Lag 0 to Lag 30 was close to 25% in the uncued conditions. These results may be used to reject the explanations of the CPA lag effect based on the GFT and the rehearsal theory, two theories which have enjoyed a considerable amount of success in this area of investigation.

For the present, the most satisfactory explanation of these results is a variant of encoding variability theory. The major assumptions are as follows. Over time the context changes. These changes are not random in the sense that similarities between contexts are an orderly function of time. The context at time $n+2$ is more similar to the context at time $n+1$ than to the context at time n. In other words, what a person is thinking and feeling at time $n+2$ is more closely associated with whatever he was thinking and feeling at time $n+1$ than time n. The encoding of stimuli, and their resultant functional stimuli, are determined by the nominal stimulus and the context. Changes in the context produce changes in the functional stimuli although the nominal stimuli are invariant. Similarities in contexts are reflected by similarities in functional stimuli.

At the time of testing the subject will be able to recognize the nominal stimulus (call it old), or utilize the nominal stimulus to retrieve a response, only to the extent that the encoded version of the stimulus at the time of the test is similar to the encoded versions stored at the input. It is this similarity function that determines the shape of the lag function. At short retention intervals, or when the cueing context is similar to the input context, the average similarity is a nonmonotonic function of lag (see Fig. 2). At longer retention intervals, when the cueing context is only asymptotically related to the input contexts, the average similarity is a monotonic function of the interpresentation lag.

This theory is a modification of encoding variability theory as proposed by Estes' (1955) Stimulus Fluctuation Theory, and its recent formulations (e.g., Bower, 1972). The major difference lies in the conceptualization of the mechanism of context change. For both Estes and Bower stimulus elements, or functional stimuli, are assumed to fluctuate between available and unavailable states, during the interpresentation intervals. Consider two stimulus elements *a* and *b*. Element *a* was available at time n, and became unavailable at time $n+1$. Element *b* was never in the available set, and remains unavailable at time $n+1$. For Estes and Bower, both of these stimulus elements have an equal probability of becoming available at time $n+2$. This assumption is mathematically tractable, but psychologically unreasonable. Stimulus elements do not fluctuate during interpresentation intervals; it is the context which fluctuates. If at the next presentation, the context has changed, then the functional stimulus will be different. As can be seen from an inspection of Fig. 2, if elements *a* and *b* are unavailable (i.e., would not be encoded) at time $n+1$, element *a* would still be more likely to be sampled at time $n + 2$. Once this change in conceptualization has been made, the prediction of nonmonotonic lag effects at short retention intervals and monotonic effects at longer retention intervals becomes obvious.

Perhaps the most important aspect of this research is its ability to lay some suspicions about the lag effect to rest. For example, Underwood (1970) was made uneasy by the ubiquity of the monotonic lag effect which seemed to suggest that the effect was psychologically trivial. This research demonstrates that the monotonic lag effect is not ubiquitous, but is critically dependent on the nature of the stimulating conditions at both input and retrieval.

APPENDIX

The purpose of this appendix is not to derive a mathematical model. The mathematics will be used to demonstrate that a relatively straightforward mathematical application of the processes diagrammed in Fig. 2 will predict the gross results of Experiment I. These results are (*a*) the main effect of the retention interval, (*b*) the main effect of the interpresentation lag, (*c*) the nonmonotonic lag effect at short retention intervals, and (*d*) the monotonic lag effect at long retention intervals.

These effects should be accounted for using the following principles:

1. The experimental context changes as a function of time and trials.

2. The encoding of the stimulus is influenced by the context at the time of presentation. The greater the change in the context from t_0 to t_1, the greater the change in the encoding from t_0 to t_1.

3. The greater the similarity between the stimulus as encoded at the test and the stimulus as encoded at the presentations, the greater the probability that the response is recalled.

Define a function $f(t)$ as the probability that the context at time t is similar enough to the context at time t_0 to reinstate the same encoding. Therefore, $f(L)$ is the probability that the same encoding of a stimulus is elicited on two presentations separated by a lag of L items. Likewise, $f(R)$ is the probability that the same encoding will be obtained after a retention interval of R items.

The only constraints on the form of $f(t)$ are that $f(0)$ be unity, $f(n) > f(n+1)$, and that $f(\infty)$ be an asymptotic value greater than zero. This last constraint provides that there is some small asymptotic probability that two encodings will be identical.

One possible formulation for the probability of recalling the response is given in Eq. (1):

$$P(\text{recall}) = f(L)\,f(R) + [1 - f(L)]\,[f(R) + [1 - f(R)]\,f(L+R)]. \quad (1)$$

In words, the probability of recalling the response is equal to the probability that the same encoding of the stimulus is obtained on both presentations and is also obtained on the test, or that two different encodings are obtained on the presentations and that either one is reinstated on the test. When R is small, $f(R)$ will be large, and retrieval is biased towards the information encoded at the second presentation. When R is large the information at both presentations contribute equally to the probability of recall (i.e., $f(R) \simeq f(L+R)$).

Clearly, as R increases $f(R)$ decreases and recall decreases. The question remaining is whether or not Equation 1 describes a nonmonotonic lag function at short retention intervals, and a monotonic function at long retention intervals. To answer this question the derivative of Eq. (1) with respect to L is set equal to zero and solved for values of L. These values of L are then substituted into the equation representing the second derivative of Eq. (1). For those values of L and R for which the derivative is negative, the lag function is concave downward.

Before proceeding, $f(t)$ must be specified. This function is the heart of the model, yet there are an infinite number of functions that will meet the constraints listed above. While the choice is psychologically arbitrary, an exponential decay, because of its mathematical tractability, was selected to represent $f(t)$. The specific function is given Eq. (2), where a represents the asymptotic value of the function, and λ controls the rate of decay.

$$f(t) = a + (1 - a)\exp(-t\lambda). \qquad (2)$$

The derivative of Eq. (1) is given in Eq. (3).

$$\frac{d}{dL} = [1 - f(R)]\,[f'(L + R)\,[1 - f(L)] - f'(L)\,f(L + R)]. \qquad (3)$$

Substituting the right-hand part of Eq. (2) into Eq. (3), differentiating, setting the derivative equal to zero, and simplifying results in Eq. (4).

$$\begin{aligned}
&(a + (1 - a)\exp[-(L + R)\lambda])' \\
&\quad (1 - a - (1 - a)\exp(-L\lambda) = \\
&\quad (a + (1 - a)\exp(-L\lambda))' \\
&\quad (a + (1 - a)\exp[-(L + R)\lambda]) \\
&\quad (1 - a)(1 - 2\exp(-L\lambda) = a\exp(R\lambda) \\
&\quad L = -(\lambda)^{-1}\ln[(1/2) - a\exp(R\lambda)/2(1 - a)].
\end{aligned} \qquad (4)$$

For values of $R < (\lambda)^{-1}\ln[(1 - a)/a]$, Eq. (4) can be solved for L, indicating that the lag function is nonmonotonic. It can also be seen that as R increases L must also increase to maintain the equality of Eq. (4).

The second derivative of Eq. (1) is:

$$\frac{d^2}{dL} = [1 - f(R)]\lambda^2(1 - a)\exp(-L\lambda)\,[(1 - a)\exp(-R\lambda) - 4(1 - a)\exp[-(L + R)\lambda] - a].$$

The derivative is negative when the bracketed portion of the right-hand side is less than zero. Equivalently, upon substitution of the right-hand side of Eq. (4) for L, and simplification, the second derivative is negative when $R < (\lambda)^{-1}\ln[(1 - a)/a]$. The lag function is concave downward when R satisfies this inequality, and monotonic for larger values of R[3].

[3] After this paper was completed, I had the opportunity to read the article by Landauer (1975) describing a theory of memory with a geometric representation similar to that presented in Fig. 2. Landauer uses his theory to account for the interaction between the lag interval and retention interval in a manner essentially identical to that proposed in this paper. The theories may be brought into closer correspondence by assuming that (a) the major structure in Landauer's Fig. 1 represents the range of possible encodings of a single item, as opposed to the whole of memory, and (b) the data entered into the structure represent specific encodings of a specific item, as opposed to specific encodings of any item. With this reformulation, the random walk of Landauer's pointer and the parameter r serve a function analogous to $f(t)$ in this paper.

REFERENCES

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

BOWER, G. H. Stimulus-sampling theory of encoding variability. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory*. Washington, D.C.: V. H. Winston & Sons, 1972.

CICCONE, D. S., & BRELSFORD, J. W. Interpresentation lag and rehearsal mode in recognition memory. *Journal of Experimental Psychology*, 1974, **103**, 900–906.

CRAIK, F. I. M., & WATKINS, M. J. The role of rehearsal in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 599–607.

D'AGOSTINO, P. R., & DEREMER, P. Repetition effects as a function of rehearsal and encoding variability. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 108–113.

EINSTEIN, G. O., PELLEGRINO, J. W., MONDANI, M. S., & BATTIG, W. F. Free recall performance as a function of overt rehearsal frequency. *Journal of Experimental Psychology*, 1974, **103**, 440–449.

ESTES, W. K. Statistical theory of spontaneous recovery and regression. *Psychological Review*, 1955, **62**, 145–154.

FALKENBERG, P. E. Recall improves in short-term memory the more recall context resembles learning context. *Journal of Experimental Psychology*, 1972, **95**, 39–47.

FISCHLER, I., RUNDUS, D., & ATKINSON, R. C. Effects of overt rehearsal procedures on free recall. *Psychonomic Science*, 1970, **19**, 249–250.

GARTMAN, L. M., & Johnson, N. F. Massed versus distributed repetitions of homographs: A test of the differential-encoding hypothesis. *Journal of Verbal Learning and Verbal Behavior*, 1972, **11**, 801–808.

GREENO, J. G. How associations are memorized. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970.

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

KINTSCH, W. Recognition learning as a function of the length of the retention interval and changes in the retention interval. *Journal of Mathematical Psychology*, 1966, **3**, 412–433.

LANDAUER, T. K. Memory without organization: Properties of a model with random storage and undirected retrieval. *Cognitive Psychology*, 1975, **7**, 495–531.

MADIGAN, S. A. Intraserial repetition and coding processes in free recall. *Journal of Verbal Learning Behavior*, 1969, **8**, 828–835.

MARTIN, E. Stimulus meaningfulness and paired-associate transfer: An encoding variability hypothesis. *Psychological Review*, 1968, **75**, 421–441.

MARTIN, E. Stimulus encoding in learning and transfer. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory*. Washington: V. H. Winston & Sons, 1972.

MELTON, A. W. The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 596–606.

MEYERS, J. L. *Fundamentals of experimental design*. (2nd Ed.). Boston: Allyn and Bacon, Inc., 1972.

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., & SALTZMAN, T. Supplementary report: Time between pairings and short-term retention. *Journal of Experimental Psychology*, 1962, **64**, 550–551.

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

SHEPARD, R. N., & TEGHTSOONIAN, M. Retention of information under conditions approaching a steady state. *Journal of Experimental Psychology*, 1961, **62**, 302–309.

SPERBER, R. D., GREENFIELD, D. B., & HOUSE, B. J. A non-monotonic effect of distribution of trials in retardate learning and memory. *Journal of Experimental Psychology*, 1973, **99**, 186–198.

UNDERWOOD, B. J. A breakdown of the total-time law in free-recall learning. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 573–580.

UNDERWOOD, B. J., & SCHULZ, R. W. *Meaningfulness and verbal learning*. Chicago: J. B. Lippincott, 1960.

YOUNG, J. L. Reinforcement-test intervals in paired-associate learning. *Journal of Mathematical Psychology*, 1971, **8**, 58–81.

(Received July 21, 1975)