Encoding processes and the spacing effect

FRANCIS S. BELLEZZA, HENRY B. WINKLER, and FRANK ANDRASIK, JR.

Ohio University, Athens, Ohio 45701

The differential encoding theory of the spacing effect was tested utilizing Martin's (1968) encoding variability notion, in which it is hypothesized that low-meaningfulness items are more variable in their encodings than are high-meaningfulness items. In a series of three experiments using a continuous paired associate learning task, it was predicted that the spacing vs. performance curves for CCC items would show a faster improvement in performance than would the curves for high-meaningfulness CVC items. None of the experiments supported this prediction. In addition, it was found that items recognized on their second presentation were more likely to be recalled than were those items not recognized. It was concluded that an item's repetitions are more effective if one code is formed and elaborated with each repetition rather than if more than one code is formed.

The spacing effect refers to the experimental finding that two spaced repetitions of an item lead to better performance on a later test than do two massed presentations. This effect has been found in free recall learning (Glanzer, 1969; Madigan, 1969; Melton, 1967), in paired associate learning (Brelsford, Shiffrin, & Atkinson, 1968; Landauer, 1969; Madigan, 1969; Peterson, Wampler, Kirkpatrick, & Saltzman, 1963; Young, 1971), and in short-term memory experiments (Peterson, 1963; Pollatsek, Note 1). In addition, Bjork and Allen (1970) and Tzeng (1973) have found that in a short-term memory task increasing the difficulty of the interpolated task occurring between the two presentations of an item has the same result as increasing the duration of the interpolated task. A more difficult task led to better performance on a later test. Similar results have been obtained by Robbins and Wise (1972) in paired associate learning. For reasons of parsimony and because of the lack of evidence to the contrary, it seems reasonable to assume that the spacing effects observed in these three types of experiments are the result of the same underlying psychological mechanisms. What these mechanisms are is still in doubt (see Bjork, 1970 for a review).

The most widely accepted, yet rather general, explanation for the spacing effect is differential encoding (Bjork & Allen, 1970). Simply stated, differential encoding predicts better performance with spaced presentations because spaced presentations are more likely to result in the item's being encoded in a different way on the second presentation. Two different encodings of an item means that the item has a better chance of being retrieved from memory during a later

test. Just how and why differences in the encodings occur depends on the particular differential encoding theory consulted. For example, Martin (1968) focuses on the idea that the variability of the encodings of an item is dependent on the meaningfulness of that item. But Melton (1970) points out that the context within which the item is presented changes from one presentation to another. However, there are two problems with a differential encoding explanation of the spacing effect. First, in many forms of the theory, the specific coding and retrieval mechanisms which lead to the spacing effect are not precisely formulated. Consequently, the results of particular experimental manipulations are difficult to predict. Second, differential encoding may be able to explain the initial increase in performance which takes place as the spacing of presentation increases, but it cannot explain why in paired associate learning performance after reaching a maximum declines with further increased spacing.

The other explanation proposed for the spacing effect is based on memory consolidation (Bjork & Allen, 1970). Some theories propose that consolidation of the memory strength of the first presentation is maximized by an appropriately spaced second presentation (Landauer, 1969; Wickelgren, 1970). The optimal spacing is dependent on hypothetical physiological mechanisms in the organism. Consolidation theories based on memory strength seem simply to assume the spacing phenomenon as a basic characteristic of the learning processes of the organism and suggest little to the experimenter as to how he may change the characteristics of the spacing effect. Consolidation may also occur by means of a psychological mechanism such as the rehearsal buffer (Atkinson & Shiffrin, 1968; Brelsford et al., 1968). In this model each presented item is placed in a rehearsal buffer, where it resides for a period of time as information concerning the item is transferred to long-term memory. If a second presentation is made while the first presentation still resides in the rehearsal buffer, the second presentation is

This research was supported in part by Grant OURC 437 to the first author by the Ohio University Research Committee. Portions of this research were presented at the annual meeting of the Midwestern Psychological Association, Chicago, May 1974. The authors would like to thank Marc Foreman, Sandra Gervis, Nancy Glasser, and Daniel Lohr for their assistance in collecting the data.

Requests for reprints should be sent to Francis S. Bellezza, Department of Psychology, Ohio University, Athens, Ohio 45701.

essentially ignored. Therefore, the second presentation is effective only if presented after some delay. This multiprocess model also has a decay factor in long-term memory which can account for the drop in performance which is found in paired associate learning when the optimal spacing of six to eight items is exceeded. The Atkinson and Shiffrin model does well in explaining the data from paired associate learning, but does not do as well in dealing with data from free recall and short-term memory experiments. For example, Madigan (1969) found in a free recall experiment an optimum spacing of 40 items, which was the largest spacing used. Also, the rehearsal buffer model predicts just the opposite of the results obtained by Bjork and Allen (1970). Increasing the difficulty of a reasonably long interpolated task in a short-term memory experiment should reduce rehearsal of the first presentation during the task and lead to poorer rather than better performance on a later test.

Before the spacing phenomenon can be adequately explained, more has to be known about what factors influence the shape of the spacing vs. performance curve. So far, the type of learning task and the degree of difficulty of the interpolated task have been discussed as two factors which do this. In addition, Tzeng (1973) has shown that increasing the duration of the interpolated task in a short-term memory experiment had a more advantageous effect on low-meaningfulness stimuli than on high-meaningfulness stimuli, although the result was not statistically significant. The effect of meaningfulness on the spacing phenomenon is of particular interest because predictions concerning the shape of the spacing vs. performance curve can be made utilizing Martin's (1968) encoding variability hypothesis as a possible way of influencing differential encoding. Martin argues that items such as low-meaningfulness consonant trigrams tend to be less integrated and, thus, have more possible encodings than do items such as high-meaningfulness consonant trigrams. Applying this notion to the spacing effect, it seems that as the spacing of presentations increases, performance on the low-meaningfulness items should improve faster than performance on the high-meaningfulness items. This is because differential encoding theory predicts that an item with two encodings is more likely to have an encoding retrieved from memory than is an item with only one encoding. This in turn means that in a paired associate learning task, the optimal spacing for low-meaningfulness items should be reached sooner than the optimal spacing for high-meaningfulness items.

It would also be desirable to make predictions about the effects of item meaningfulness from a fairly explicit consolidation model such as that of Atkinson and Shiffrin (1968). However, differences in item meaningfulness in this model could lead to a number of different rehearsal strategies. This makes predictions about the shape of the spacing curve difficult. For example, with low-meaningfulness items, which are more

difficult to learn and may occupy more space in the rehearsal buffer, the subject could choose simply to ignore some of the items and rehearse the others. Therefore, manipulation of stimulus meaningfulness results in a better test of the differential encoding explanation of the spacing effect than it does of the consolidation explanation.

EXPERIMENT I

Method

Subjects. The subjects were 24 volunteers enrolled in introductory courses at Ohio University. The subjects received course credit for their participation.

Lists. Two types of lists were made up. Eight lists had CVC trigrams as stimuli whose association value was 99% or 100% (Archer, 1960). The other eight lists had CCC trigrams as stimuli. The CCC lists were produced by dropping the vowel from each stimulus in the CVC lists and replacing it with a randomly drawn consonant under the constraint that all three consonants were different. Hence, the eight CCC lists and the eight CVC lists were matched in structure. On both types of lists, the responses were randomly drawn letters of the alphabet. There were 64 items in each list; each item was presented twice and then tested (RRT paradigm) in a continuous paired associate learning task. Four different items occurred with each of the following spacings of presentations: 0, 3, 6, . . . , 45. After the second presentation of the item, 10 more intervening items occurred before the test. This wide range of spacings was used because preliminary experiments showed that the optimal spacing occurred when using these materials at a larger value than the seven or eight intervening items found by previous experimenters. The lists were made up following the procedure described by Young (Note 2): A stimulus, a spacing, and a response were randomly drawn and then fit into the list as early in the list as possible. After all 64 items were placed in the list in this manner, filler items, which were similar to the experimental items but not analyzed, were used to fill the empty spaces in the list. Ten additional filler items were used at the beginning of the list to eliminate any primacy effect.

Apparatus. The lists were presented using a Lafayette IBM memory drum. Three Hunter interval timers were used in tandem to turn the drum and start a Hunter KlocKounter at the same time. On test trials subjects responded vocally, closing a Grason-Stadler voice-operated relay. The subject's response and response latency were recorded by the experimenter.

Procedure. The subjects were run individually for two 50-min sessions, with each session on a different day. During each session, the list used had either CCCs or CVCs as stimuli, and the order of presentation was counterbalanced over subjects. The two lists used with each subject were always based on different randomizations. Both the presentation trials and test trials were 7 sec in duration, and the subjects were not forced to guess if they did not know the answer.

Analysis. The data consisting of the proportion of items correct were evaluated using a 2 by 2 by 16 analysis of variance, with the specific factors being list order, stimulus meaningfulness, and spacing of presentations. Only list order was a between-subjects factor. Because the subjects were not required to respond on each test trial, the Ns in each cell for the latency data were unequal. Therefore, the latency data were analyzed using a 2 by 16 least-squares solution analysis of variance, with the specific factors being stimulus meaningfulness and spacing.

Results

Analysis of the proportion of items correct showed

Figure 1. Proportion of CCC and CVC items correct at each spacing of presentations in Experiment I.

that the main effect of stimulus meaningfulness was significant [$F(1,22) = 79$, $p < .001$], as was the main effect of spacing [$F(15,330) = 2.47$, $p < .005$]. The only significant interaction was Stimulus Meaningfulness by List Order [$F(1,22) = 14.64$, $p < .001$]. If CVC items were presented during the first session, the subjects gave .66 correct responses to them. If the CVC items were presented during the second session, the subjects gave .85 correct responses. However, for the CCC items the proportions were .45 and .50, respectively. The Stimulus Meaningfulness by Spacing interaction was not significant [$F(15,330) = 1.18$, $p > .25$]. As can be seen in Figure 1, the performance curves for CCC and CVC items are roughly parallel, with the best performance occurring at a spacing of about 15 intervening items. Because there was no Stimulus Meaningfulness by Spacing interaction, the CCC and CVC items were combined and a trend analysis was performed (Kirk, 1968). To simplify the computations, performance values for adjacent spacings 0 and 3, 6 and 9, etc., were combined so that the trend analysis was done on eight points. The results of this analysis indicated that only the cubic component of the best-fit polynomial equation was significant [$F(1,330) = 17.05$, $p < .001$]. It accounted for 77% of the sums of squares of the spacing effect. The shape of the two curves predicted from the best-fit polynomial equation is shown in Figure 1. The rise that occurs at the end of the predicted curve is a characteristic of a cubic equation. The empirical curves show only slight evidence of a second maximum.

A separate analysis of variance was done on the latencies of incorrect and correct responses. The analysis of incorrect responses showed no significant effects. The analysis of correct responses showed stimulus meaningfulness to be significant [$F(1,1856) = 201$,

$p < .001$], as was spacing [$F(15,1856) = 2.97$, $p < .001$]. Figure 2 shows the mean response latencies in seconds for CCC and CVC items. As with the proportion of items recalled, the Stimulus Meaningfulness by Spacing interaction was not significant [$F(15,1856) = 1.10$, $p > .25$]. The number of observations for each data point ranged from 35 to 84. A trend analysis showed both the quadratic component [$F(1,1856) = 16.45$, $p < .001$] and the cubic component [$F(1,1856) = 5.66$, $p < .05$] to be significant. The two components accounted for 54% and 19% of the spacing sum of squares, respectively. The best-fit polynomial equation for the CCC and CVC items combined is also shown in Figure 2. It indicates an optimal spacing of from 15 to 18 intervening items and is almost a mirror image of the spacing curve formed by the proportion correct shown in Figure 1. The predicted line is closer to the CVC curve because it contained the greater number of observations.

Discussion

The results indicate that the notion of differential encoding utilizing Martin's (1968) encoding variability hypothesis is inadequate as an explanation for the spacing effect in paired associate learning. The CCC items were certainly more difficult to learn. The overall proportion of CCC items correct was .48, whereas the proportion correct of CVC items was .75. Similarly, the mean response latency for correct CCC items was 2.71 sec, but for correct CVC items it was 1.96 sec. However, the two spacing vs. performance curves were parallel. According to differential encoding theory, increased spacing of the two presentations should have improved performance at a faster rate for the CCC items than for the CVC items (Tzeng, 1973), and the optimal spacing for CCC items should have been smaller. It might also be noted that the results do not correspond with what would be predicted from the work of Bjork and Allen (1970). Because the CCC items are more difficult to learn, the rate of improvement with spacing for the CCC items is again predicted to be greater.

Figure 2. Mean latency of correct responses to CCC and CVC items at each spacing of presentations in Experiment I.

EXPERIMENT II

The optimum spacing of 15 items found in Experiment I is inconsistent with the results of previous research in continuous paired associate learning. Two similar experiments (Peterson et al., 1963; Young, 1971) found an optimum spacing of seven or eight intervening items. One possible difference in procedure was that a 7-sec presentation time was used in Experiment I, whereas Peterson et al. (1963) used a 2-sec presentation time and Young (1971) used a 1-sec presentation time followed by a 3-sec blank. Therefore, Experiment II was done using the same procedure as Experiment I, but using a 4-sec presentation time for each item. Also, it was thought that a 4-sec presentation time might amplify the differences in encoding variability between the CCC and high-meaningfulness CVC items. This could result in a significant Stimulus Meaningfulness by Spacing interaction, as predicted by the differential encoding theory.

Method

The procedure used was the same as in Experiment I, except that the presentation time was 4 sec/item. The same lists were used, but latency data was not collected.

Results

The results of Experiment II are shown in Figure 3. The main effect of stimulus meaningfulness was again significant [$F(1,22) = 74, p < .001$], as was the spacing effect [$F(15,330) = 2.48, p < .005$]. The Stimulus Meaningfulness by List Order interaction was not significant, as it was in Experiment I, nor was the Stimulus Meaningfulness by Spacing interaction. As in Experiment I, the optimal spacing was about 15 items. A trend analysis showed only the cubic component to be significant [$F(1,330) = 3.94, p < .05$]. The best-fit polynomial equation is plotted in Figure 3 and accounts for 22% of the spacing sum of squares.

Discussion

The data from Experiment II also do not support the differential encoding hypothesis. Neither do they support the hypothesis that presentation rate influences the shape of the spacing curve, although presentation rate does affect the overall level of performance. The invariance of the shape of the spacing curve with respect to presentation rate has also been found in free recall learning by Melton (1967) and by Glanzer (1969).

EXPERIMENT III

The results of Experiments I and II indicate that either the differential encoding hypothesis is not a satisfactory explanation for the spacing effect or that something is wrong with Martin's (1968) proposed relationship between stimulus meaningfulness and encoding variability. In Experiment III, a more direct

Figure 3. Proportion of CCC and CVC items correct at each spacing of presentations in Experiment II.

test of the differential encoding theory was made, which was independent of the differences in stimulus meaningfulness. A key idea in differential encoding theory is that multiple encodings lead to a higher probability of recall at the time of test (Bjork & Allen, 1970; Madigan, 1969; Melton, 1970; Robbins & Wise, 1972; Tzeng, 1973). Martin (1968) maintains that a stimulus is encoded differently on its second presentation if it is not recognized. Does this mean that a stimulus will be encoded in the same way if it is recognized on its second presentation? It seems reasonable to assume that, if the first encoding provides the basis for recognition, then the encoding on the recognized second presentation will be the same. Tzeng (1973), for example, states that only by making the first encoded retrieval cue unavailable at the time of the second presentation of the to-be-remembered item will a different encoding be possible. In Experiment III a recognition task was added to the continuous paired associate learning task. Differential encoding predicts that items which are not recognized on their second presentation will be recalled better at test because they will have two encodings rather than only one.

Method

Subjects. The subjects were 64 volunteers from introductory psychology courses at Ohio University. They received course credit for their participation.

Lists. Eight randomized lists were made up using the procedure described in Experiment I, except that each list contained a mixture of 32 CCC and 32 CVC items rather than 64 of only one type. This was done so that both CCC and CVC items were judged with respect to the same set of distractor items. Recognition performance on the two types of items could then be compared (Banks, 1970).

Procedure. The subjects were run individually for one 50-min session on one of the eight lists. Each item was presented for 7 sec. Thirty-two of the subjects were instructed to study the items and give a response whenever a test item appeared. The

Figure 4. Proportion of CCC and CVC items correct at each spacing of presentations in Experiment III.

other 32 subjects were instructed to indicate whether or not they recognized each item as it appeared by saying "yes, yes," which was scored 4 and meant that the subject was sure that he recognized the item; "yes," which was scored 3 and meant that he recognized the item but was not completely sure about it; "no," which was scored 2 and meant unsure nonrecognition; and "no, no," which was scored 1 and meant sure nonrecognition. These subjects were requested to give the recognition response to both presentation and test items and, in addition, were to give the correct response on any test trials.

Analysis. The proportion of items correct was analyzed using a 2 by 8 by 2 by 16 analysis of variance, with the specific factors being group task, randomization form, stimulus meaningfulness, and spacing. Group task and randomization form were between-subjects factors. An 8 by 2 by 16 analysis of variance was done on the recognition scores of the second presentation to determine if CVC items were recognized more easily than were CCC items, as predicted by encoding variability theory. The factors were randomization form, stimulus meaningfulness, and spacing, respectively. To determine if items recognized on their second presentation were recalled better or worse than those not recognized, a point biserial correlation (McNemar, 1962) was computed between the recognition score of the second presentation and the recall response of the test trial, which was scored 0 or 1.

Results

Analysis of the proportion of items correct showed that the group task was significant [$F(1,48) = 32$, $p < .001$], with the control group recalling .65 of the items correctly and the recognition group .42. Stimulus meaningfulness [$F(1,48) = 255$, $p < .001$] and spacing [$F(15,720) = 2.65$, $p < .001$] were also significant. The recall performance curves for CCC and CVC items for the control and recognition groups combined are shown in Figure 4. The only significant interaction was the Stimulus Meaningfulness by Spacing interaction [$F(15,720) = 1.80$, $p < .05$]. This interaction appeared to be caused by a slight rise in the CVC curve at spacings of 27 to 45. Because the effect seemed to be a small one, a trend analysis was performed on the CCC and CVC

items combined. The quadratic component [$F(1,720) = 4.18$, $p < .05$] and the cubic component [$F(1,720) = 4.46$, $p < .05$] were found to be significant. The quadratic component accounted for 24% of the spacing sum of squares, and the cubic 25%. The best-fit polynomial equation is also shown in Figure 4.

The analysis of the recognition data from the second presentation showed that stimulus meaningfulness [$F(1,24) = 184$, $p < .001$], spacing [$F(15,360) = 15$, $p < .001$], and the Stimulus Meaningfulness by Spacing interaction [$F(15,360) = 3.25$, $p < .001$] were significant. At a spacing of 0, both CCC and CVC items were recognized almost perfectly, with a mean recognition score of 3.9. With spacing values of 9 to 15 intervening items, the mean recognition score for CCC items was 3.15 and for CVC items was 3.73. With a spacing of 39 to 45 items, the values were 2.54 and 3.30, respectively. The value of 2.50 represents chance performance in recognition. A greater decrease in recognition performance with spacing for low-meaningfulness items has also been reported by Martin and Melton (1970).

The point biserial correlation between the recognition response on the second presentation and the recall response on the test trial was $r = .24$, $p < .001$ for the CCC items and $r = .27$, $p < .001$ for the CVC items. This indicates that those items better recognized on their second presentation were more likely to be recalled.

Discussion

The recall results of Experiment III are much like those of Experiments I and II, except that the Stimulus by Meaningfulness interaction was significant. This appeared to be caused not by a sharper increase or by a smaller optimal spacing for the CCC items, as predicted by differential encoding theory, but by a somewhat steeper drop in performance for the CCC items compared to the CVC items after both curves reached a maximum at about 15 to 18 intervening items. The recognition group had a more difficult task than did the control group because the subjects in this group had to make one or two overt responses on every trial. However, the expected shortening of the optimal spacing did not take place in the recognition group (Bjork & Allen, 1970). The main purpose of Experiment III was to determine if a new encoding of the second presentation of an item resulted in better recall. Just the opposite was found, if it can be assumed that recognition of the item on the second presentation meant that the same encoding was used. Encoding nonvariability led to better performance, which is not what is predicted from the multiple-retrieval-cue notion of differential encoding theory.

GENERAL DISCUSSION

Differential encoding theory when used in

conjunction with encoding variability theory is inadequate to explain the spacing effect in paired associate learning. In all three experiments reported here, the rate of improvement in recall performance with spacing and the optimal spacing were the same for CCC and CVC items. Differential encoding predicts that the rate of improvement should be greater for CCC items than for CVC items, and the optimal spacing should be less. If the decline in performance after optimal spacing has been reached is the result of interference in long-term memory, then the decrement in performance after the optimum should probably be greater for CCC items. Of these three predicted phenomena, only the third was found, and then in only one of the three experiments. However, the recognition data support Martin's (1968) notion of encoding variability. The recognition performance for CCC items dropped faster with spacing than did performance for CVC items. For spacing of 33 to 45 items, the recognition performance of CCC items was essentially at the chance value of 2.5, whereas the mean recognition score for CVC items in this range was 3.30 on a scale from 1.0 to 4.0. Therefore, it is difficult to determine why the spacing vs. performance curves are so similar.

The results of Experiment III indicate that the multiple-retrieval-cue notion of differential encoding theory, which assumes that the cues are unconnected in memory, cannot explain the repetition effect. The repetition effect is the observed improvement in performance that results from multiple presentations (Bjork, 1970). If the second presentation is recognized, then the item is more likely to be recalled correctly on a later test trial. This result has also been found by Melton (1967) in free recall learning. If variability is to be beneficial, then the variability should not lead to a second code for the item but should lead, perhaps, to an elaboration of retrieval cues for the first code. Other research supports this position: Hintzman and Block (1973) found that on a delayed test subjects were able to judge the spacing of two occurrences of the same word far better than they could judge the spacing of single occurrences of two different words. Nelson (1971) found that in order for a repetition to be effective, it must be a repetition of remembered information.

In Experiments I, II, and III, the optimal spacing under all experimental conditions was always in the range of 15 to 18 intervening items. This conflicts with the spacing results of other paired associate learning experiments (Brelsford et al., 1968; Landauer, 1969; Madigan, 1969; Peterson et al., 1963; Young, 1971), all of which show an optimum spacing of six to eight intervening items. The method of list construction used here is one described by Young (Note 2). Young's analysis of his list showed no relationship between order of first appearance in the list and spacing. In an analysis of the eight lists used in Experiments I and II, however, a rank correlation of $-.14$, $p < .001$, was found between first

appearance in the list and spacing. It is not difficult to see why this negative correlation occurs. A stimulus, a response, and a spacing are randomly drawn, and the item is fit into the list as early in the list as possible. This means that the beginning of the list is built up sooner than any other part. An item with a large spacing between its first and second presentation can be more easily fit into the densely populated early part of the list than can an item with its first and second presentation close together. Also, a point biserial correlation of $-.25$, $p < .001$, was found over all items between order of first occurrence in the list and correct recall. Thus, it appears that the lists are biased somewhat toward correct responding for items with larger spacings. This may explain why the optimum spacing occurred at 15 intervening items rather than at 6 to 8. This result does not, however, negate the finding that the optimum spacings for CCC and CVC items were the same. Another interesting finding, which mitigates somewhat those findings mentioned above, is that an analysis of covariance of the data from Experiments I and II which used the order of first appearance of the item in the list as the covariate did not affect either the significance of the spacing effect or change the optimal spacing from 15 items. All things considered, however, the optimal spacing value of 15 items must be interpreted with caution.

In summary, a differential encoding theory which utilizes Martin's (1968) encoding variability hypothesis was not supported as an explanation of the spacing effect. The recognition data of Experiment III did support the idea that low-meaningfulness items are more variable in their encodings than are high-meaningfulness items. However, in all three experiments there was very little difference in the shapes of the spacing vs. performance curves for the two types of items. Also, the notion that multiple encoding improves recall was not supported by the data. Multiple cues may facilitate recall, but only if the retrieval cues are connected to the same memory code.

REFERENCES

- Archer, E. J. A re-evaluation of the meaningfulness of all possible CVC trigrams. *Psychological Monographs*, 1960, 74(10, Whole No. 497).
- Atkinson, R. C., & 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. 2. New York: Academic Press, 1968.
- Banks, W. P. Signal detection theory and human learning. *Psychological Bulletin*, 1970, 74, 81-99.
- Bjork, R. A. Repetition and rehearsal mechanisms in models for short-term memory. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970.
- Bjork, R. A., & Allen, T. W. The spacing effect: Consolidation or differential encoding? *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 567-572.
- Brelsford, J. M., Shiffrin, R. M., & Atkinson, R. C. Multiple reinforcement effects in short-term memory. *British Journal of Mathematical and Statistical Psychology*, 1968, 21, 1-19.
- Glanzer, M. Distance between related words in free recall: Trace of the STS. *Journal of Verbal Learning and Verbal Behavior*, 1969, 8, 105-111.

- Hintzman, D. L., & Block, R. A. Memory for the spacing of repetitions. *Journal of Experimental Psychology*, 1973, 99, 70-74.
- Kirk, R. E. *Experimental design procedures for the behavioral sciences*. Belmont, Calif: Wadsworth, 1968.
- Landauer, T. K. Reinforcement as consolidation. *Psychological Review*, 1969, 76, 82-96.
- Madigan, S. A. Intraserial repetition and coding processes in free recall. *Journal of Verbal Learning and Verbal 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., & Melton, A. W. Meaningfulness and trigram recognition. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 126-135.
- McNemar, Q. *Psychological statistics*. 3rd ed. New York: Wiley, 1962.
- Melton, A. W. Repetition and retrieval from memory. *Science*, 1967, 158, 532.
- 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.
- Nelson, T. O. Transfer of serial information from short-term memory to long-term memory: Toward the locus of the repetition effect. *Psychonomic Science*, 1971, 25, 344-345.
- 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., 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.
- Robbins, D., & Wise, P. S. Encoding variability and imagery: Evidence for a spacing-type effect without spacing. *Journal of Experimental Psychology*, 1972, 95, 229-230.
- Tzeng, O. J. L. Stimulus meaningfulness, encoding variability, and the spacing effect. *Journal of Experimental Psychology*, 1973, 99, 162-166.
- Young, J. L. Reinforcement-test intervals in paired-associate learning. *Journal of Mathematical Psychology*, 1971, 8, 58-81.
- Wickelgren, W. A. Multitrac strength theory. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970.

NOTES

1. Pollatsek, A. W. Rehearsal, interference, and spacing of practice in short-term memory. (Tech Rep 16) Ann Arbor: University of Michigan, Human Performance Center, 1969.
2. Young, J. L. Effects of intervals between reinforcements and test trials in paired-associate learning. (Tech Rep 101) Stanford: Stanford University, Institute for Mathematical Studies in the Social Sciences, 1966.

(Received for publication September 7, 1974;
accepted September 17, 1974.)