DOCUMENT RESUME

ED 039 222 TE 001 700

AUTHOR	Fishman, Elizabeth Jane; And Others
TITLE	Massed Versus Distributed Practice in Computerized Spelling Drills.
INSTITUTION	American Psychological Association, Washington, D.C.
PUB DATE	Aug 68
NOTE	7p.
JOURNAL CIT	Jounral of Educational Psychology; v59 n4 p290-96 Aug 1968
EDRS PRICE	EDRS Price MF-$0.25 HC Not Available from EDRS.
DESCRIPTORS	*Computer Assisted Instruction, Experimental Programs, Grade 5, *Learning Processes, *Reinforcement, *Retention, *Spelling, Spelling Instruction, Time Factors (Learning)

ABSTRACT

Computerized spelling drills were used to study conditions of massed and distributed practice with each of 29 fifth-grade subjects participating in both conditions. In the distributed practice, two sets of three words each were presented once every other day over a period of 6 days. The drill on six other sets of words was massed so that all practice for that set occurred on the same day. The probability of a correct response for words in the massed condition proved higher than that for the distributed condition during the learning sessions, but on retention tests (given 10 and 20 days later) the words learned under distributed practice were better remembered. Thus, massed repetitions appear to be better on short term performance, but more learning occurs in the long run when repetitions of an item are well distributed. (A mathematical model of the learning process is presented.) (Author/MF)

Journal of Educational Psychology
1968, Vol. 59 No. 4, 290-296

MASSED VERSUS DISTRIBUTED PRACTICE IN COMPUTERIZED SPELLING DRILLS[1]

ELIZABETH JANE FISHMAN, LEO KELLER,[2] AND RICHARD C. ATKINSON
Stanford University

Conditions of massed and distributed practice were studied using a within-Ss design in a situation involving computerized spelling drills. In the distributed condition, 2 sets of 3 words each were presented once every other day over a period of 6 days. The learning trials on 6 other sets of words were massed so that all of the trials for that set occurred on the same day. Ss were 29 5th graders. The probability of a correct response for words in the massed condition was higher than that for the distributed condition during the learning sessions, but on retention tests (given 10 and 20 days later) the words learned under distributed practice were better remembered. A mathematical model of the learning process is presented and shown to provide a fairly adequate account of the experimental data.

Computer-assisted instruction (CAI) refers to an instructional procedure which utilizes a computer to control part, or all, of the selection, sequencing, and evaluation of instructional materials. Over the last 4 years, the Institute for Mathematical Studies in the Social Sciences at Stanford University has been developing a CAI system for regular classroom usage (Atkinson, 1967). One mode of this development is referred to by Suppes (1966) as the "drill and practice systems." These systems are intended to supplement the instruction which occurs in the classroom. They are designed to improve—through practice—the skills and concepts which are introduced by the classroom teacher.

Currently, computer controlled drills are being given to approximately 1,800 students in six schools in five different communities. Some of the students have been receiving daily drills in arithmetic (Suppes, Jerman, & Groen, 1966) while others have been receiving drills in spelling. This study made use of the equipment and students in the school which has been involved in drill and practice in spelling.

In the study to be reported here, the presentation routine for each spelling word was the same: An audio system presented the words, the student typed the word, and the computer evaluated the student's answer. If the response was correct, the computer typed "...C..."; if incorrect, "...X...", followed by the correct spelling of the word. If the response was not given within a predetermined length of time, the message "...TU...", meaning "time is up," was printed. A flow chart summarizing this procedure is given in Figure 1.

These CAI drill and practice systems lend themselves nicely to the study of many experimental variables. One persistent problem in designing instructional systems is the specification of optimal procedures for presenting material. Indeed, the spacing of learning sessions has already received considerable experimental investigation, yet the question of optimal spacing has not been resolved. For example, assume that we have 6 days in which to teach a list of 24 spelling words, and that each daily session is arranged so that 24 presentations can be made. What practice schedule would produce the best results? One might select a different set of four words each day and on that day present each word six times. At the other extreme, one could present each of the 24 words once per day. In both schemes a given word would be presented for study on six different occasions, but in one condition all of the repetitions for a given word would occur on 1 day whereas in the other scheme they would be distributed over 6 days.

[1] This research was supported by an Office of Education contract, No. OE5-10-050, and by a computing grant from the United States Public Health Service, Contract No. MH 06154.

[2] Now at the University of California, Irvine.

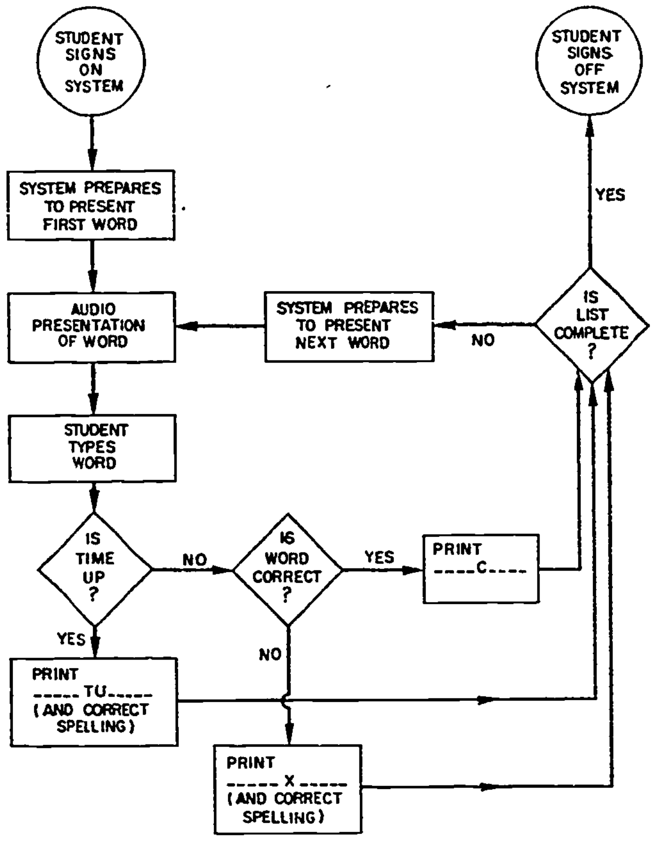

FIG. 1. Flow chart for presentation routine.

The two extremes could be called, respectively, massed and distributed practice, although this terminology is somewhat at variance with the classical usage of these terms. The preponderance of experimental evidence indicates that, for the same amount of practice, learning is better when practice is distributed rather than massed, although there are exceptions to the generalization. The purpose of the present study is to investigate this problem further and to evaluate optimum procedures for distributing instructional material in computer-based spelling drills.

METHOD

Subjects

The *Ss* were 29 students from a fifth-grade class in an East Palo Alto school. Approximately 50% of these students scored below grade level on standardized reading tests; 20% were reading at the second and third grade level.

The Computer System and Terminals

The computer which controlled the student terminals was a modified PDP-1 digital computer located at Stanford University. It was a time-sharing computer capable of handling over 30 different users simultaneously from a variety of input devices. The audio system for the spelling drills was controlled by a Westinghouse P-50 computer which, in turn, was linked to the PDP-1.

The four student terminals were located at an East Palo Alto school in a converted storeroom a short distance from the child's classroom. Each terminal consisted of a standard teletype machine and a set of earphones; both were linked to the computer at Stanford by telephone lines.

All four terminals were controlled by a single program on the PDP-1; each student user was serviced sequentially in a round-robin cycle. Due to the extremely rapid speed of the computer, the student received the impression that he was getting "full-time" service, although actually the computer was devoting only a small fraction of its running time to any one individual.

Daily Operation

A full-time monitor was on duty whenever the children were using the teletypes. Her presence was primarily a precautionary measure so that an adult would be available in case of an emergency. The actual check-in, presentation and evaluation of the drill, and the sign-out were all handled by the CAI system and occurred as follows.

The student entered the room, sat down at a free terminal, and put on his earphones. The machine printed out, "Please type your number." (This whole routine had been explained to the students during a 2-week orientatioin session.) After the student typed in his identification number and depressed the space bar—the latter operation was used as a termination signal for all student responses—the computer printed the student's name and the program was set in operation. The message, "If you hear the audio, please type an 'a' and a space," was then heard over the earphones. If the instructions were followed, the lesson began and each word was presented according to the sequence given in Figure 1.

The audio system presented a word, used the word in a sentence, and then repeated the word again. As soon as the audio was through, the machine typed a dash (—). This was the student's signal to begin his response. When he finished typing his answer, he depressed the space bar, and the computer evaluated the answer. A correct response was followed by the typed message, "...C...". An incorrect response was indicated by the message, "...X...," followed by several spaces and a correct spelling of the word. If a response was not given in 40 seconds, the message, "...TU..." was printed. As on an incorrect answer, this message was followed by several spaces and the correct spelling of the word. Following his response the student was given 6 seconds to study the correct answer before the next item was presented. Each time a new item was presented, all previous items were covered.

In the training sessions of this study, a "list" consisted of 12 such presentations; in the test sessions, 24 presentations. When the entire list had been presented, the machine printed out the following information for the student: his list number for the next session, the date and ending time, and the number of words he spelled correctly on the day's session. The drills were collected by the monitor and at no time was the student given a copy of the words to study on his own.

Words

The words used in the experiment were taken from the New Iowa Spelling Scale (Greene, 1954). This scale is the product of the testing of some 238,000 pupils throughout the country in the early 1950s to determine the percentage of students that could spell a word correctly at each grade level. A list of the actual words used in the experiment can be found elsewhere (Fishman, 1967).

Experimental Design

The experiment involved a within-*S*s design, (i.e., each *S* participated in all conditions). The two main conditions were those of massed (M) and distributed (D) practice. There were eight sets of words: six of them were massed, designated M_1, M_2, M_3, M_4, M_5, and M_6; and two were distributed, designated D_1 and D_2. Each of these eight sets contained three words. Thus a total of $8 \times 3 = 24$ words were used in the experiment for a given *S*. Training sessions ran for 6 consecutive days. Each session used one of the M sets and one of the D sets. The M words were presented three times within a session, whereas the D words were presented once. Thus, there were $3 \times 3 = 9$ presentations of M items plus 3 presentations of D items yielding a total of 12 presentations in any one session. Words from a different M set were presented in each session and all the learning trials for the set occurred on the same day. Words from a given D set were presented on alternating days. Table 1 summarizes the daily presentations.

The arrangement of the list for the first training session (Day 1) illustrates the procedure used for the entire training sequence. The first four items of the day's list consisted of the three words in M_1 plus a randomly chosen word from D_1. The second four items consisted of the three M_1 words plus a second randomly chosen D_1 word. The last four items consisted of all three M_1 words plus the remaining word from D_1. In other words, the 12 presentations to an *S* on any day were given in three blocks with four words in a block. Each block contained all three M words and a randomly chosen D word. The order of the words within a block was randomly determined. Further, the assignment of words to M and D sets was completely counterbalanced over *S*s, so that every word appeared equally often in the various M and D conditions.

Tests were administered 10 and 20 days after the end of the training sequence. The students did not receive any computerized drill between the training and test days. The basic test procedure consisted of presenting the complete list of 24 words. The order of the words for each *S* was randomly determined, and each word was presented once using the procedure of Figure 1. As during

TABLE 1

SUMMARY OF THE WORD SETS USED DURING THE SIX TRAINING SESSIONS

Condition	1	2	3	4	5	6
Massed (M)	M_1	M_2	M_3	M_4	M_5	M_6
Distributed (D)	D_1	D_2	D_1	D_2	D_1	D_2

the training sessions, the student was told whether or not his response was correct, and was then given 6 seconds to study the correct answer before the next item was presented.

RESULTS

Figure 2 presents the proportion of correct responses over successive presentations of M and D items. For example, on Day 1, the M_1 items were each presented three times; the proportions correct for each of the three presentations were averaged over *S*s and plotted successively above Training Session 1. The D_1 items were each presented once; the mean proportion correct for these items is also plotted above Training Session 1. This was done for the data from each of the six training sessions. Approximately 2 minutes elapsed between two presentations of a massed item, whereas 2 days elapsed between any two presentations of a distributed item.

The tests were given on Days 16 and 26. The test results are also presented in Figure 2. The six massed curves are similar in form; they all rise sharply, then drop off by the time of the administration of the first test. In contrast, the two distributed curves rise more gradually but do not show a drop off at the time of the first test.

All items were presented three times during the training sequence and once on each of the test days. Figure 3 gives the proportion correct on each presentation averaged separately over M and D items. During the training sequence, the proportion correct for the M items increased from about .31 on the first presentation to .77 on the third presentation, whereas the D items correspondingly increased from about .25 to .57. The difference between the average proportion correct on the first presentation of M items and the first presentation of D

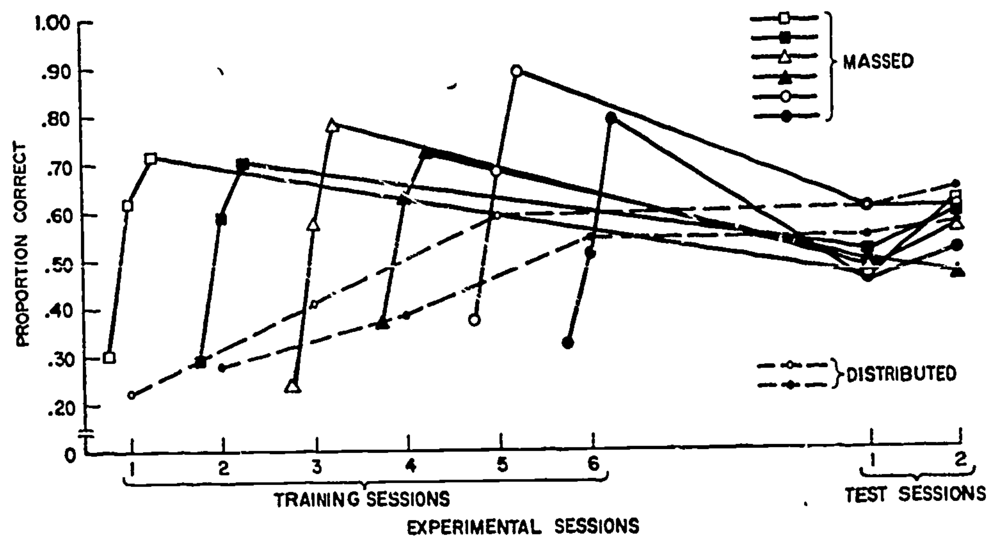

Fig. 2. Proportion of correct responses for massed and distributed items on both training and test trials.

items was not significant at the .05 level using a paired t test, $t = 1.58$, $df = 28$. However, there is no reason to expect equality when it is noted that the data point for the mean of the massed first presentations came from all six training sessions whereas the data point for the mean of the distributed first presentations came from the first two training sessions. In contrast, as indicated in Figure 3, there were significantly more correct responses on the second and third presentations of the M items than on the corresponding presentations of D items.

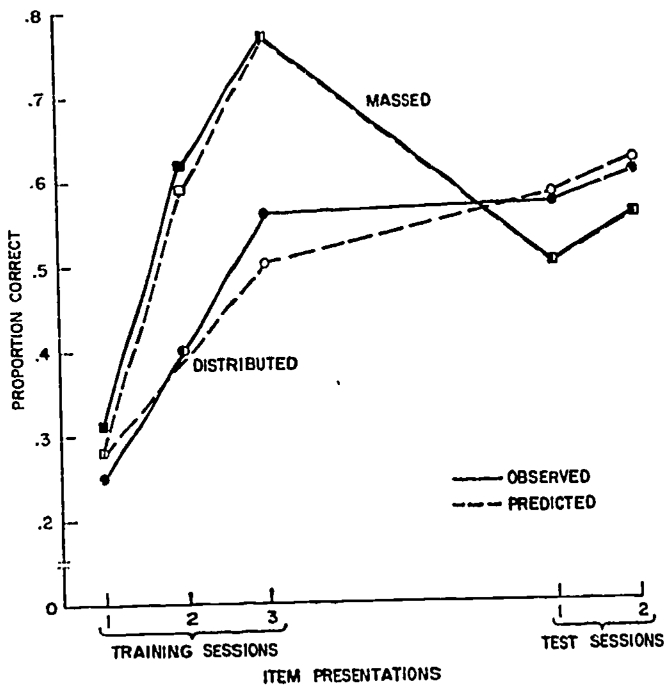

FIG. 3. Observed and predicted values for the massed and distributed conditions.

A paired t test on the combined data from the posttraining tests yielded $t = 2.44$, $df = 28$ which was significant at the .025 level, indicating that distributed practice resulted in better performance than massed practice.

DISCUSSION

The major results of this experiment were: (*a*) the massed condition was superior to the distributed condition on the second and third presentations of the training sequence and (*b*) the distributed condition was superior on both of the test sessions. Thus, it appears that the massed repetitions are better if one looks at short-term performance, but in the long run more learning occurs when repetitions of an item are well distributed.

In this section, these data are analyzed in terms of a model that has been proposed to account for paired-associate learning. The model is a variation of the trial-dependent-forgetting model presented in recent articles by Atkinson and Crothers (1964) and Calfee and Atkinson (1965). The learning of a list of spelling words can be said to resemble the learning of a list of paired-associate items; no assumption is made that the two tasks are identical, yet there are variables in paired-associate learning that clearly are relevant to the spelling task.

In the model, S is assumed to be in one of three learning states with respect to a stimulus item: (*a*) state U is an unlearned state, in which S responds at random from the set of response alternatives, (*b*) state S is a short-term memory state, and (*c*) state L is a long-term state. The S will always give a correct response to an item if it is in either state S or state L. However, it is possible for an item in state S to be forgotten, that is, to return to state U, whereas once an item moves to state L it is learned in the sense that it will remain in state L for the remainder of the experiment. In this model, forgetting involves a return from the short-term memory state, S, to state U, and the probability of this return is postulated to be a function of the time interval between successive presentations of an item.

More specifically, two types of events are assumed to produce transitions from one state to another: (*a*) the occurrence of a reinforcement, that is, the paired presentation of the stimulus item together with the correct response, and (*b*) the occurrence of a time interval between successive presentations of a particular item. The associative effect of a reinforcement is described by the following transition matrix:

$$\begin{array}{c c} & \begin{array}{ccc} L & S & U \end{array} \\ \begin{array}{c} L \\ S \\ U \end{array} & \begin{bmatrix} 1 & 0 & 0 \\ a & 1-a & 0 \\ bx & (1-b)x & 1-x \end{bmatrix} \end{array}$$

Thus, if an item is in state U and the correct response is shown to S, then with probability $(1 - x)$ the item stays in state U, and with probability x the item moves into state S or L: If it moves, then with probability b it moves into L and with probability $(1 - b)$ into S. Similarly, if an item is in state S and the correct response is shown, then with probability a the item moves to state L, and with probability $1 - a$ the item stays in state S. Finally, if an item is in state L, then it remains there with probability 1. The parameter x is assumed to vary as a function of the familiarity of the items in the list being studied. Thus, during the test sessions involving 24 familiar items, x will be larger than during the initial study sessions involving 12 items, many of which are presented for the first time.

From one presentation of an item to its next presentation, a transition can occur as described by the following matrix:

$$\begin{array}{c c} & \begin{array}{ccc} L & S & U \end{array} \\ \begin{array}{c} L \\ S \\ U \end{array} & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1-f_t & f_t \\ 0 & 0 & 1 \end{bmatrix} \end{array}$$

The parameter, f_t, depends on the time interval between successive presentations of the same item. If a given item is in state S, a time interval t between successive presentations may result in forgetting of the item (i.e., transition to state U) with probability f_t. Otherwise there is no change in state. For simplicity, we assume $f_t = 0$ for short time intervals within the range of a given training session. When the time interval is a day or greater, then we assume $f_t = 1$. In essence, no forgetting occurs from the short-term state within a given training session, but from one day to the next no information is retained in short-term store. Furthermore, the above transition matrices imply that L is an absorbing state; once an item enters state L it remains there. The model makes the additional assumption that at the start of the experiment an item is already known (state L) with probability p, or not known (state U) with probability $1 - p$.

For this model, the difference between the M and D items on the second and third presentations is due to a difference in the probability that an item is in short-term memory (state S). The parameter a characterizes the probability of going from state S to state L. This parameter can operate only for the massed items, since it is impossible for a distributed item to be in state S when a reinforcement occurs. A distributed item could go into state S immediately after its presentation, but from one presentation to its next, it would have been forgotten. The probability of being correct on an item that is in state S is one; thus the massed curves should be higher for the second and third presentations.

The assumption that $f_t = 1$ when the time interval is a day or longer, means that short-term memory has been wiped out completely by the time the first test is given. Thus, superiority of the D items over the M items in the test data indicates differences in the number of items in state L. This in turn implies that the parameter b must be larger than the parameter a. If b were smaller than a, one would expect the M condition to do better than the D condition during both the training and test sessions, whereas if b were equal to a, one would expect a difference during the training sessions in favor of the M condition, but none in the test sessions.

Parameter estimates for the model were obtained by methods described in Atkinson and Crothers (1964). The values which yielded the best fit between observed and predicted proportions were:

$$p = .28$$
$$a = 0$$
$$b = .38$$
$$x \text{ (for training sessions)} = .45$$
$$x \text{ (for test sessions)} = .74$$

These estimates were consistent with the notion that b should be larger than a. The model proposed here is similar to Greeno's (1964) model for paired-associate learning in which he explicitly requires the parameter a to be zero. The present findings for this more complex task indicate that his theory and related research on paired-associate learning are relevant to the effect of repeated presentations of spelling items. Figure 3 presents the fit between the observed and predicted proportions using the above parameter estimates. Inspection of this figure indicates that the model gave an adequate account of the results of the experiment.

To check the validity of these results, the same S's were run 2 weeks later using precisely the same procedure but with a new set of words. Figure 4 presents learning curves for this replication comparable to those presented in Figure 3. Application of the model to this data yielded the following set of parameter estimates:

$$p = .32$$
$$a = 0$$
$$b = .33$$
$$x \text{ (for training sessions)} = .60$$
$$x \text{ (for test sessions)} = .72$$

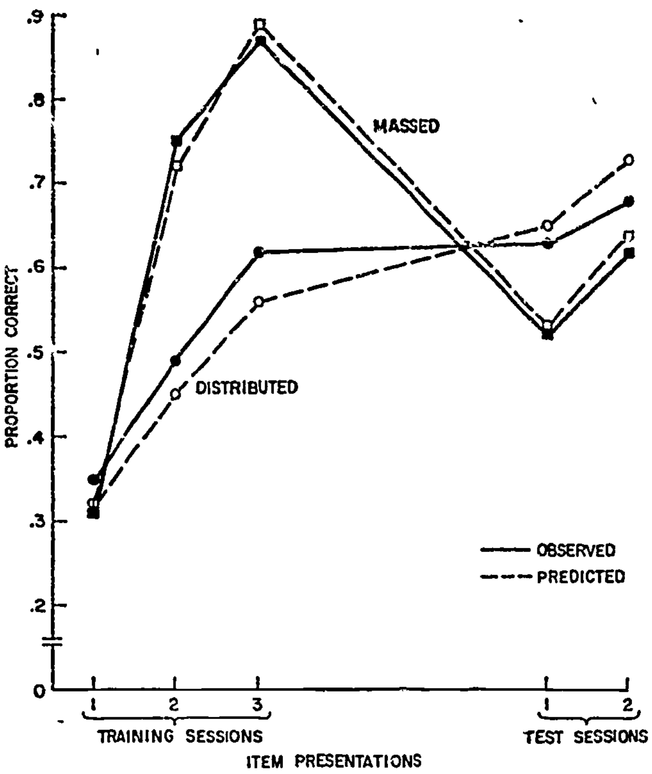

FIG. 4. Observed and predicted values for the replication experiment.

Once again, the estimate of a is zero confirming our earlier result. Also, in general, performance is superior in the second experiment, suggesting that some form of learning-to-learn may be operating in this situation.

The authors have not carried out analyses that bear on some of the more detailed features of the model. In fact, in view of the stimulus material used, it seems unlikely that these features would be verified. What clearly needs to be done is to generalize the paired-associate model to take account of the linguistic constraints imposed by the spelling task. Some of the present results and those of Knutson (1967) suggest guidelines for such a model but the authors are not prepared to be more specific at this time. Hopefully such a model would provide a more definitive answer to the problem of optimizing the instructional sequence in spelling drills.

REFERENCES

ATKINSON, R. C. Instruction in initial reading under computer control: The Stanford Project, *Journal of Educational Data Processing*, 1967, **4**, 175–192.

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

CALFEE, R., & ATKINSON, R. C. Paired-associate models and the effects of list length. *Journal of Mathematical Psychology*, 1965, **2**, 254-265.

FISHMAN, E. Massed vs. distributed practice in computerized spelling drills. Unpublished master's thesis, Stanford University, 1967.

GREENE, H. A. *The New Iowa Spelling Scale*. Iowa City: State University of Iowa, 1954.

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

KNUTSON, J. M. Spelling drills using a computer-assisted instructional system. Technical Report No. 112, Institute for Mathematical Studies in the Social Sciences, Stanford University, 1967.

SUPPES, P. The uses of computers in education. *Scientific American*, 1966, **215**, 207-220.

SUPPES, P., JERMAN, M., & GROEN, G. Arithmetic drills and review on a computer-based teletype. *Arithmetic Teacher*, 1966, April, 303-308.

(Received August 5, 1967)