HUMAN LEARNING, VOL. 4, 11–18 (1985)

The Effect of Expanded versus Massed Practice on the Retention of Multiplication Facts and Spelling Lists

CORNELIUS P. REA AND VITO MODIGLIANI*
Department of Psychology, Simon Fraser University, Burnaby, B.C., Canada V5A 1S6

SUMMARY

Distributed practice typically leads to better retention than massed practice. Recent research has shown that distributed practice with intertest intervals of an expanding nature is optimal. In the experiment reported here we used a test series with expanded intervals to teach multiplication facts and spelling lists to 44 grade three students. The hypothesis was that material learned with such a series would be retained better than if presented in a massed series. The results showed that, for multiplication facts, retention in the expanded series condition was almost twice that in the massed series condition. For spelling lists, a significant difference in the same direction was also obtained. These differences were obtained regardless of the level of ability of the students. An important point is that an expanded test series not only engenders effective retention, but also maintains a feeling of success throughout. This should be a prime consideration in teaching young children basic facts that they are sometimes reluctant to tackle. The use of this type of series would therefore have obvious benefit if incorporated into remedial programs or used in learning centres.

Much of what is learned in educational settings involves memory processes. Primary graders, for example, must memorize the alphabet and the correct spelling of thousands of words. They must also master the number series and multiplication tables. Several areas of memory research have implications for classroom activities that foster learning and retention. One of these concerns the spacing effect.

The spacing effect refers to the finding that items are better retained if practice trials are spaced rather than massed (Hintzman, 1976; Melton, 1970; Underwood, 1970). The most typical finding is that as spacing increases, retention also increases. In some cases, however, the relationship between spacing of presentations and retention is not monotonic (e.g. Glenberg, 1976, 1977). A variety of theoretical interpretations of spacing effects have been formulated (Hintzman, 1974; Rea, 1982; see also Cuddy and Jacoby, 1982; Glenberg, 1979; Jacoby, 1978; Wilson, 1976). Despite its theoretical elusiveness, however, the spacing effect is one of the most robust phenomena discovered in memory research. It could be expected, therefore, that there would be a considerable body of research that systematically investigated the spacing of repetitions and rehearsals of to-be-remembered information in applied settings. This, however, is not the case (Bloom and Shuell, 1981).

Most of the technical literature on the spacing effect (see Hintzman, 1974; 1976) is concerned with the spacing of *presentations* of to-be-remembered material. A much smaller body of literature deals with the effects of the spacing of *tests*. In the latter case

* Author to whom reprint requests should be addressed.

0277–6707/85/010011–08$01.00

Received 13 April 1984
Revised 13 July 1984

the typical paradigm includes a single presentation of a to-be-remembered item followed by two or more tests. There are two questions of interest. First, does a test increase later performance; secondly, is the spacing between tests a significant variable in determining later performance?

With regard to the first question, it has been conclusively demonstrated that tests do improve memory (Izawa, 1971, 1978, 1981; Landauer and Eldridge, 1967; Modigliani, 1976; Wenger, Thompson and Bartling, 1980; Whitten and Bjork, 1977; Whitten and Leonard, 1980). This phenomenon has been referred to by Izawa (1971) as the potentiating effect of test trials.

With regard to the second question, Modigliani (1976) showed that a test has a potentiating effect only for items that are successfully recalled. That is, an item that is recalled is strengthened by that recall and is thereby more likely to be recalled again later. An item not recalled on the test, however, has an extremely low probability of being recalled on a later test. Furthermore, the degree to which a recalled item is strengthened *increases* with the length of the distracting interval preceding the recall.

The above findings have important implications for designing a series of tests that are spaced in a maximally effective manner. Specifically, the spacing of intervals between tests should be of an *expanding* nature. In such an arrangement the interval between item presentation and the first test is very short, that between the first and second test is somewhat longer, and so on. In this type of series the probability of a successful recall on the first test is high, since the initial interval is short. A successful recall will strengthen the recalled item and will help to 'carry' it, so to speak, for the second interval which is longer than the first. The probability of a successful recall on the second test is therefore also high and the recall will further strengthen the item, and so on. Thus, in an expanded series, the probability of success is maintained at a high level throughout and, in addition, the potentiating effect of testing increases as the intervals between tests are increased (Modigliani, 1976). By contrast, other types of series are not as effective. For example, in a massed series of tests, where tests follow each other at very short intervals, recall will be high throughout but the potentiating effect will be small, resulting in poor long term retention. On the other hand, with a series in which the first interval is relatively long, the probability of success on the first test is small and the item is unlikely to be recalled on any subsequent test.

Landauer and Bjork (1978) demonstrated in a laboratory setting that expanded test schedules were more effective than either massed series or series in which intervals were equally spaced. In the present paper we report an experiment in which the effectiveness of an expanded test schedule similar to that used by Landauer and Bjork was verified in a school setting. Generally, in classroom situations there is no systematic manipulation of intervals between the presentation and testing of facts that have to be memorized. Indeed, in the senior author's classroom experience, the most common method of memorizing material is massed practice, as suggested, for example, in the Canadian Spelling Program 3. The present study was an attempt to apply basic research findings as discussed above to help children learn more effectively.

METHOD

Subjects

Subjects were 44 male and female students (average age = 8.5 years) from two Grade 3 classes at Cedar Drive Elementary School, Port Coquitlam, British Columbia. Their teachers were asked to rank order the students separately on mathematical ability and spelling ability. The mathematics ranks were used to form the massed and expanded groups in the mathematics condition (see below) by using the ABBA pattern of assignment. The spelling ranks were similarly used to form the massed and expanded groups in the spelling condition.

Design

The main independent variable involved two between-subject conditions of practice, expanded and massed, for each of two types of material, multiplication facts (mathematics condition) and spelling lists (spelling condition). Half the students received the mathematics condition first and the other half the spelling condition first. Since there was reason to believe that the two classes were not equivalent in either mathematical or spelling ability, a second between-subject variable was class (Class 1, Class 2). A third variable was level. Those students ranked in the top half in each class were designated as Level 1 whereas those ranked in the bottom half constituted Level 2. Class and Level allowed us to examine the effect of massed versus expanded practice on students with different ability.

Materials

In the mathematics condition the to-be-remembered items consisted of five multiplication facts; 8×5 to 8×9. In the spelling condition, subjects learned to spell four words; *brought*, *happily*, *machine* and *police*. These items were selected as the result of a pretest showing that the students were not familiar with their spelling.

On the basis of pilot work, the following method of presentation proved most effective. All events were presented audio-visually using an audio cassette tape in conjunction with a Singer Caramate rear screen slide projector. One track of the tape contained auditory presentation of events and this was synchronized, using pulses on the other track, with slide presentation of the same events. For example, as the auditory message 'eight times five is forty' was presented, a slide displaying '$8 \times 5 = 40$' was shown on the screen.

Sequences

All subjects received a sequence of events which consisted of either a presentation (P) or a test (T) of a target item or a distracting event (D). The distracting events were tests of familiar multiplication facts (e.g. $2 \times 2 = ?$) and words (e.g. cat). In the expanded series condition the following sequence of events was used for each target item: P_1, T_1, P_2, D_1, T_2, P_3, D_2, D_3, T_3, P_4, D_4, D_5, D_6, D_7, T_4, P_5. The subscripts on P and T refer to successive presentations or tests of the same fact. For example, assume that P_1 was the first presentation of the fact $8 \times 5 = 40$; T_1 would then be the first test of that fact ($8 \times 5 = ?$), P_2 a re-presentation of P_1, T_2 a second test identical to T_1, and so on. The subscripts on D refer to different distracting events.

In the above sequence there is no distracting event intervening between the initial presentation and the first test, one between the first and second tests, two between the second and third and four between the third and fourth. Thus, the length of the distracting interval expands between successive tests. The same sequence was used for each of the to-be-remembered items.

The optimal duration of events, established on the basis of a pilot study, was 5 s for presentations, 6 s for tests and 5 s for distractors. A brief interval (10 s in the mathematics condition and 15 s in the spelling condition) followed the last presentation (i.e. P_5) of the sequence for that particular item, after which the sequence for the next item was presented.

The massed condition group received exactly the same events as the expanded condition group, differing only in the sequence in which presentations, tests and distractors were given. All seven distractors occurred first and were then followed by consecutive presentations and tests as follows: $D_1, D_2, D_3, D_4, D_5, D_6, D_7, P_1, T_1, P_2, T_2, P_3, T_3, P_4, T_4, P_5$. This sequence was used for each of the to-be-remembered items. The duration of each event was the same as in the expanded condition. Thus, the expanded and massed conditions were identical in every way except in the sequencing of the events.

Tests

In the mathematics condition, following the final presentation of the last to-be-remembered fact, there was a one-minute distraction interval during which the experimenter engaged the subject in conversation while the tape and slide tray were changed. At the end of this retention interval, an oral test of all five multiplication facts was given using the same format as the learning session. In this test the student was allowed 10 s to answer each test question and all answers were tape recorded for later analysis. Upon completion of the oral test, subjects were given a second untimed written test the purpose of which was to allow an additional opportunity to respond without time constraints. The subject was given a sheet containing five questions, one for each of the facts, and asked to write down the answers. Subjects were instructed to take their time and attempt to remember as many facts as they could without resorting to computing the answers.

In the spelling condition the same testing procedure was used. A taped oral test was given one minute after the final presentation of the last to-be-remembered word. The subject had 12 s in which to respond by spelling the word aloud and responses were tape recorded for later analysis. This test was followed by an untimed written test in which each test word was again presented on an audio tape and the subject's task was to write his or her response. Subjects were instructed to take their time. The next test word was not presented until the previous question had been answered or an attempt made.

Procedure

Each subject was seen individually in a room set aside for this purpose and was read a standard set of instructions. The experimental sequences were preceded by a brief practice section in which known multiplication facts or words were used.

Mathematical facts and spelling words were presented to all the subjects in the same order, but two orders of testing were used, one the reverse of the other, counterbalanced across subjects. Subjects who received Order 1 in the oral test received the same order in the written test, and likewise for Order 2.

RESULTS

The main results are presented in Figure 1. As can be seen, the performance levels of the

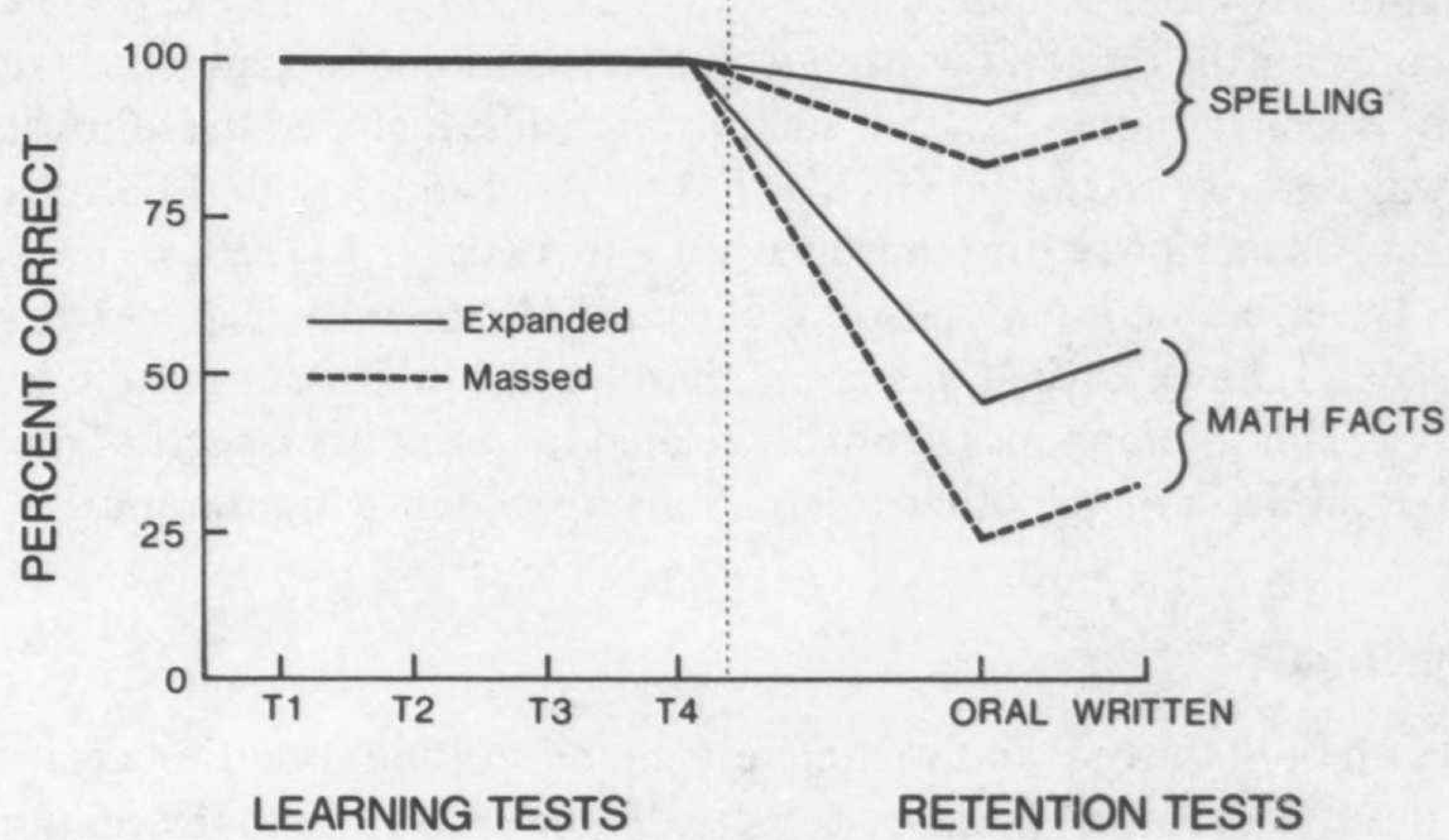

Figure 1. Mean percentage recall in the learning and retention tests as a function of type of practice in the learning session

groups during the *learning* session were all at the 100 per cent level. There were, however, major differences between the two conditions on the *retention* tests, as shown in Table 1. Retention performance was analysed separately for the mathematics and spelling conditions.

Mathematics condition

Mean scores on both the oral and written tests are presented in Table 1. It can be seen that performance ranged from 13 to 76 per cent correct, the overall average being 36 per cent. A 2 (Practice) × 2 (Class) × 2 (Level) × 2 (Type of final test) analysis of variance revealed that the practice variable was highly significant, $F(1, 36) = 24.35$, $p < 0.01$, with expanded practice being superior to massed practice. Performance on the written

Table 1. Mean percentages of correct responses for multiplication facts and spelling lists as a function of practice (expanded vs. massed), test (oral vs. written), Class (1 vs. 2) and Level (1 vs. 2).

		Oral test		Written test		
Class	Level	Expanded	Massed	Expanded	Massed	Means
Multiplication facts						
1	1	23	17	30	16	20
	2	36	13	40	23	28
2	1	64	43	76	57	60
	2	40	17	67	16	34
	Means	41	20	53	28	36
Spelling lists						
1	1	96	90	100	95	95
	2	70	67	85	88	77
2	1	100	92	100	92	96
	2	92	70	96	75	83
	Means	90	78	95	85	87

test was significantly better than on the oral test, $F(1, 36) = 11.38$, $p < 0.01$. Mean retention scores for the oral and written tests as a function of expanded versus massed practice can be seen in Figure 1. Other significant results included the following. Recall in Class 2 was superior to that in Class 1, $F(1, 36) = 23.65$, $p < 0.01$. Level 1 subjects gave marginally superior performance to those in Level 2, $F(1, 36) = 3.97$, $p < 0.06$. The Class × Level interaction was also significant, $F(1, 36) = 13.53$, $p < 0.01$. As can be seen from Table 1, Level 2 recall was better than Level 1 in Class 2, but the reverse was true in Class 1. This interaction is probably related to unknown aspects of the teachers' ranking criteria. None of the other interactions approached significance.

Spelling condition

Mean scores on both the oral and written tests in the spelling condition are also shown in Table 1. It can be seen that in this task, performance was close to the ceiling, ranging from 67 to 100 per cent correct, with an overall average of 87 per cent. A $2 \times 2 \times 2 \times 2$ analysis of variance similar to that used for the mathematics condition showed expanded practice to be superior to massed practice, $F(1, 36) = 4.55$, $p < 0.05$. As was the case for the mathematics condition, performance on the written test was better than that on the oral test, $F(1, 36) = 15.78$, $p < 0.01$. Mean retention scores for the oral and written tests as a function of expanded versus massed practice can be seen in Figure 1. Recall was not significantly different in Classes 1 and 2, $F(1, 36) < 1$. Level 1 was superior to level 2, $F(1, 36) = 13.61$, $p < 0.01$. Two of the interactions, Test × Class and Test × Level, were significant, $F(1, 36) = 6.91$, $p < 0.05$ in both cases. The Test × Class interaction reflects the fact that Class 2 performed significantly better than Class 1 on the oral test but on the written test Class 1 did better than Class 2, although this difference was not significant. Likewise, the Test × Level interaction pattern showed a non-significant difference between Level 1 and Level 2 on the written test, but Level 1 performance was significantly better than Level 2 on the oral test. Both interactions are artefacts of the ceiling effect that was present in the written test but not in the oral test.

DISCUSSION

The main hypothesis of this experiment, that expanded practice is superior to massed practice in a classroom situation, was confirmed. A sizable and striking effect was achieved with multiplication facts and a significant difference was obtained, despite performance being close to ceiling levels, with spelling lists. Not surprisingly, performance was better on the untimed written test than on the timed oral test.

The lack of any significant interactions with the variable of interest, condition of practice, is important. As can be seen from Table 1, differences between classes and between levels were quite substantial. The lack of interactions indicates that expanded practice is equally beneficial regardless of the ability level of the students. This result confirms laboratory findings in demonstrating, once more, the robustness of the spacing effect.

An expanded series of tests is a special case of distributed practice. It has characteristics that are, in our opinion, especially important in teaching young children. Perhaps the most important of these is that during the learning session there should be a high probability of success throughout the series. That this was indeed the case is evident in Figure 1 (left side): performance remained at the 100 per cent level, not only in the massed series where that is expected, but in the expanded series as well. The important difference, is that, whereas an expanded series also maintains performance at a high level during learning, it engenders much better retention than the massed series.

In other types of distributed practice in which the intervals between tests are equal, the probability of success on the initial test may be low if the first interval is too long. Thus, in an equal spacing condition, a child who is unsuccessful on the first test might become discouraged. But in an expanded type of series, the child, by experiencing success throughout the learning session, is less likely to find the task aversive.

It would seem to be especially important to provide this type of positive learning experience for children who are having problems in subjects such as mathematics or spelling. A direct application of an expanded series as a remedial aid for children with learning difficulties would undoubtedly be beneficial. If children can be helped to experience success in learning basic facts such as those in the multiplication tables, this will have a facilitative effect on learning more difficult concepts such as long multiplication and division. Additionally, this type of series could easily be adapted for learning centres or as part of mathematics drills used in computer assisted learning.

In the present experiment, we taught the children a unit of five mathematical facts in one session and a unit of four word spellings in another. Mean retention in the final tests was 36 per cent in the mathematics condition and 87 per cent in the spelling condition. Such retention differences raise the question of the optimal unit size. Given a total number of to-be-remembered facts, one could divide them into a relatively large number of units of small size or into a smaller number of units of larger size. In the first case, the child might have an easier time learning each unit but the units would not be well integrated into the context of the whole. In the second case, the child might have difficulty coping with too much information at once. The issue of whether it is better to learn parts first or tackle the whole from the outset has long been debated but not settled (e.g. Breen and Jurek, 1975; Higbee, 1977). Our results with multiplication facts and spelling lists suggest that the optimal unit size may vary depending on the type of material involved. It is obvious that more research is needed to determine the most efficient approach for different kinds of material encountered in school learning. Regardless of the size of the units the robustness of the spacing effect, as shown in the laboratory and now in the classroom, indicates that the use of expanded test series, as an efficient teaching device that is not aversive to the child, is a technique that has obvious practical advantages.

ACKNOWLEDGEMENTS

Preparation of this article was supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada to Vito Modigliani.

We gratefully acknowledge the support and cooperation of the principal and teachers at Cedar Drive Elementary School.

REFERENCES

Bloom, K. F. and Shuell, T. J. Effects of massed and distributed practice on the learning and retention of second-language vocabulary. *Journal of Educational Research*. 1981, 74, 245–248.

Breen, M. J. and Jurek, J. M. Serial learning as a function of age and part versus whole learning procedures. *Psychological Reports*. 1975, **36**, 767–773.

Cuddy, L. J. and Jacoby, L. L. When forgetting helps memory: An analysis of repetition effects. *Journal of Verbal Learning and Verbal Behaviour*. 1982, **21**, 451–467.

Glenberg, A. M. Monotonic and nonmonotonic lag effects in paired-associate and recognition memory paradigms. *Journal of Verbal Learning and Verbal Behavior*. 1976, **15**, 1–15.

Glenberg, A. M. Influences of retrieval processes on the spacing effect in free recall. *Journal of Experimental Psychology: Human Learning and Memory*. 1977, **3**, 282–294.

Glenberg, A. M. Component-level theory of the effects of spacing of repetitions and recall and recognition. *Memory and Cognition*. 1979, **7**, 95–112.

Higbee, K. L. *Your memory: How it works and how to improve it*. New York: Prentice-Hall, 1977.

Hintzman, D. L. Theoretical implications of the spacing effect. In R. L. Solso (Ed.), *Theories of cognitive psychology: The Loyola symposium*. New York: Erlbaum, 1974.

Hintzman, D. L. Repetition and memory. In G. H. Bower (Ed.), *The psychology of learning and motivation*. New York: Academic Press, 1976.

Izawa, C. The test trial potentiating model. *Journal of Mathematical Psychology*. 1971, **8**, 200–224.

Izawa, C. Theoretical and empirical performance differences between anticipation and study-test methods: retention interval vs. spaced practice effects. *Scandinavian Journal of Psychology*. 1978, **19**, 151–158.

Izawa, C. Quantitative tests of the retention interval model. *Journal of General Psychology*. 1981, **105**, 273–292.

Jacoby, L. L. On interpreting the effects of repetition: solving a problem versus remembering a solution. *Journal of Verbal Learning and Verbal Behavior*. 1978, **17**, 649–667.

Landauer, T. K. and Bjork, R. A. Optimum rehearsal patterns and name learning. In M. M. Gruneberg, P. E. Morris and R. N. Sykes (Eds), *Practical aspects of memory*. New York: Academic Press, 1978.

Landauer, T. K. and Eldridge, L. Effects of tests without feedback and presentation-test interval in paired-associate learning. *Journal of Experimental Psychology*. 1967, **75**, 290–298.

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.

Modigliani, V. Effects on a later recall by delaying initial recall. *Journal of Experimental Psychology: Human Learning and Memory*. 1976, **2**, 609–622.

Rea, C. P. *A practical application of the spacing effect to classroom learning*. Unpublished master's thesis. Simon Fraser University, Burnaby, B. C., Canada, 1982.

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.

Wenger, S. K., Thompson, C. P. and Bartling, C. A. Recall facilitates subsequent recognition. *Journal of Experimental Psychology: Human Learning and Memory*. 1980, **6**, 135–144.

Whitten, W. B. II and Bjork, R. A. Learning from tests: Effects of spacing. *Journal of Verbal Learning and Verbal Behavior*. 1977, **16**, 465–478.

Whitten, W. J. II and Leonard, J. M. Learning from tests: Facilitation of delayed recall by initial recognition alternatives. *Journal of Experimental Psychology: Human Learning and Memory*. 1980, **6**, 127–134.

Wilson, W. P. Developmental changes in the lag effect: an encoding hypothesis for repeated word recall. *Journal of Experimental Child Psychology*. 1976, **22**, 113–122.