The Effect of Massed Versus Spaced Practice on Retention and Problem-Solving in High School Physics[1]

MICHAEL G. GROTE, Department of Education, Ohio Wesleyan University, Delaware, OH 43015

ABSTRACT: **An analysis of the effects of massed versus spaced practice in the study of two topics in physics was performed. Three classes of physics students participated. Students were taught one of the topics followed by massed practice and the other topic followed by spaced practice. Two tests were administered two weeks after the final spaced practice to determine if there was a difference in retention or ability to solve new types of problems. Likert scales were administered to determine if interest in the subject matter was affected by the type of practice. In addition, interview data were collected to ascertain other variables which may have been overlooked. Statistically significant results favoring spaced practice were obtained for both the recall of subject matter (0.37) and the application of information to solve new kinds of problems (0.60). Interest was not significantly affected by the treatments. Interview data generally showed a preference for spaced practice.**

OHIO J. SCI. 95 (3): 243–247, 1995

INTRODUCTION

An educational technique emphasized by the National Science Teachers Association's Scope and Sequence Project is the revisiting of topics throughout the curriculum. One way of examining the effectiveness of revisiting topics is to examine the effects of spacing the practice of skills over a long period of time as opposed to teaching the topic only in an isolated unit of study. Several reviews of the literature published in the 1970s (Melton 1970, Hintzman 1974, Glenberg 1977) support the superiority of spaced practice over massed practice. While there is much clinical evidence to support this phenomenon, the spacing effect receives little attention in teacher training programs and is seldom seen in the repertoire of teachers.

One reason that the past research on the spaced effect has had little impact on teaching is that most research has not been immediately relevant to the classroom. In an article examining the lack of application of the spacing effect, Dempster (1988) concluded "Arguably, the most serious of the plausible impediments to the application of the spacing effect is the paucity of impressive classroom demonstrations of the phenomenon." The bulk of the research in this field has been done in a laboratory setting, generally involving undergraduate psychology students and usually involving psycho-motor learning tasks. The validity of the phenomenon has not been demonstrated in secondary science classrooms.

Some field support comes from the process-product research which was done in the 1970s. Rosenshine and Stevens (1986) summarized the behaviors of exemplary teachers and specifically cited two strategies which are variations of spaced practice that commonly appear among exemplary teachers. First, weekly and monthly reviews enhance learning. Second, college classes with weekly quizzes have higher final exam grades than identical classes without weekly quizzes. Both of these techniques cause students to study/practice material during several different periods of time.

The only experimental study concerning the spaced effect in a science classroom to be located by a literature review was done at the university level in astronomy. This study provided support for the success of the spacing effect in that setting (Lu 1978).

Although spaced practice has well-supported benefits for relatively simple learning (memorizing sentences or paragraphs and performing simple tasks), its superiority over massed practice has not been demonstrated for more complex learning in subjects such as physics (Glover et al. 1990). There has also been no research evaluating the effectiveness of spaced practice in increasing the ability of students to solve new kinds of problems, beyond those specifically practiced.

The present study considers three questions involving spaced practice: 1) Does the long term retention of science facts, concepts, and generalizations benefit more from a) spaced practice or b) an equivalent amount of massed practice; 2) Is there a difference between performance of students who have done spaced practice versus those using massed practice when these students are asked to apply their previously practiced knowledge to new types of problems; and 3) Does interest and/or attention influence the success of one method over the other?

MATERIALS AND METHODS

Two physics topics which were not a part of the normal curriculum were chosen for this study in order to control students' exposure to the subject matter: Hooke's law and a set of four empirical laws dealing with the effects of string characteristics on the vibrating frequency of the strings. The two topics were taught on two successive days. An announced quiz was given the day following the initial teaching of each topic to encourage the students to make a serious effort to learn each. Both topics were initially presented in a learning cycle style lesson beginning with an exploratory activity using real equipment, a demonstration, transparencies to summarize the conclusions of the observations, and initial practice applications including feedback. Fifty-one eleventh-grade students from three physics classes were assigned

[1]Manuscript received 18 July 1994 and in revised form 30 November 1994 (#94-16).

randomly to either group H or group S (Fig. 1). Five students were subsequently eliminated from the study because of absences on test dates. The remaining forty-six students consisted of 23 males and 23 females. The three classes contained a mixture of students in each group. A preliminary study completed in April 1991 (Grote 1995) indicated that topic order had no significant effect on learning the topics.

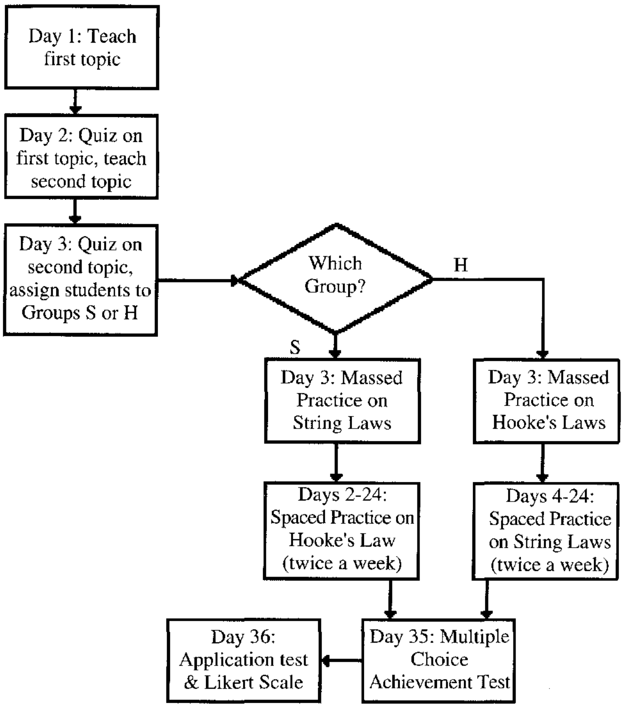

FIGURE 1. Experimental Design.

On the third day, Group H received a massed practice of Hooke's law problems while Group S received a massed practice of string law problems. Both groups were given 80 minutes to complete these independently, with only teacher assistance. Feedback was provided in the form of an answer sheet which students were given after completing each page of the handout.

Twice a week for the next month both groups experienced distributed practice. Group H received a single sheet of 3-4 questions concerning the string laws while group S received a single sheet containing questions on Hooke's law. Students were allowed up to 10 minutes at the beginning of class to complete the practice sheets. Feedback was provided on answer sheets. By the end of the month, both groups had worked the same problems on both Hooke's law and the string laws, the difference being in whether the work had been massed or distributed. Although the time spent working on problems varied somewhat between individuals, the maximum time allotted for both massed and spaced practice was the same. All notes and practice were collected from the students daily to insure that the amount of practice and study was controlled as much as possible.

Six weeks after the initial teaching of the material, the students were given an unannounced achievement test which contained a total of 44 multiple choice questions, 22 questions concerning each topic. This test was used to determine the long-term retention of the content under study. The questions on the achievement test were very similar to the kinds of questions the students used in the application and practice problems they had completed earlier. Due to the experimental design, this test was six weeks after the massed practice and two weeks after the final spaced practice. A previous study by the author using these same topics with a different group of students (Grote 1995) demonstrated no significant difference between achievement test scores at two, four, and six week intervals after completing the study of these specific topics. Problems such as equality of presentation, revealing the possibility of a test, or unnatural lapse of time between instruction and practice precluded a design in which the time between final practice and testing was a constant for the two types of practice.

The following day, students completed a Likert scale attitude inventory to determine if students had a preference for one topic over the other and if they were more attentive to massed practice items or spaced practice items.

In addition, a second test was administered requiring the application of the content to types of problems not found in the original practice. This test was designed to measure the problem-solving skills of the students based on the type of practice they received. These questions required students to reason by analogy or to solve problems requiring several steps. Although these questions were based on the material studied earlier, students had not previously worked with any similar problems. For example, linear strain was studied in the original unit on Hooke's law. A question on the application test asked students to calculate volume strain. The student was forced to build a definition of volume strain based on his/her understanding of linear strain in order to solve the problem. An example of a multi-step problem on the application test follows: "A musician is designing a new stringed instrument which is similar to a harp, but has a distinctive triangular shape. Each string is 2 cm different in length than the strings on either side of it. Eight strings will constitute one octave (which means the 9th string will have twice the frequency of the first). If the first and longest string is 100 cm long and has a frequency of 880 Hz, how will the tension on the 9th string compare with the tension on the first string assuming that the strings are identical except for length and tension?" Although the original lesson involved the relationship of frequency to length and frequency to tension, there were no practice problems involving both variables simultaneously.

Both tests were evaluated by a panel consisting of university content experts, university education experts, and secondary school physics teachers for content validity and clarity. The reliabilities of the tests were verified by calculating Cronbach's alpha for each. These were at acceptable levels ranging from 0.6 to 0.8.

Students were interviewed at the conclusion of the testing in an attempt to uncover variables or factors which may not have been anticipated.

A multivariate matched pairs analysis, a special case of a repeated measures multivariate analysis of variance (MANOVA), was used to analyze the data. Since each student was a member of both a massed practice group and a spaced practice group, the scores of the same individual could be matched, removing most within-group variability from the error term (Stevens 1986).

Limitations

1. This study was performed in a single suburban school in southwestern Ohio in the author's physics classes consisting primarily of eleventh graders.
2. It was limited to two topics not normally taught in high school physics.
3. The problem solving test was quite difficult and the scores were rather low. A larger spread of scores would be desirable for providing more convincing support for the results.
4. Since all students completed the multiple choice achievement test before the application test, students, in effect, received a spaced practice before the application test.
5. Although attempts were made to limit exposure to the content outside of the time allocated in the study, it is not possible to control what students think or discuss outside of class time.
6. To model classroom conditions as closely as possible, the time between the massed practice and testing was greater than the time between the spaced practice and testing. Although another study indicated little, if any, difference because of the time difference involved, the current study would have been more convincing if the times had been equal.

RESULTS

Students who used spaced practice in their study of a topic did better on both achievement (Table 1) and the application of the material to new kinds of problems (Table 2). Because of the differences in means and standard deviations of the tests, test scores were converted to z-scores for comparing the students who used massed practice with those using spaced practice for both the achievement test and the application test.

Table 1

Achievement subtest data.

	Hooke			String		
	Spaced	Massed	Total	Spaced	Massed	Total
Mean	9.7	8.4	9.1	9.3	7.7	8.4
Standard Deviation	4.4	3.3	4.0	3.2	2.9	3.1
Range	3-18	3-15	3-18	4-17	3-16	3-17
Number	24	22	46	22	24	46

Significance: 0.37

Table 2

Application subtest data.

	Hooke			String		
	Spaced	Massed	Total	Spaced	Massed	Total
Mean	15.5	9.0	12.4	7.4	4.8	6.1
Standard Deviation	8.2	5.8	7.8	6.0	3.9	5.1
Range	0-28	0-20	0-28	0-24	0-16	0-24
Number	24	22	46	22	24	46

Significance: 0.60

The results of the multivariate tests of significance indicated multivariate significance at the 0.05 level. To determine which of the variables were affected by the type of practice, a series of multiple correlated *t*-tests were performed using the Bonferroni approach to keep the overall alpha level under control; a 0.017 alpha level (0.05/3) was used to test for significance at the 0.05 level overall.

Achievement was influenced to a statistically significant degree by the type of practice. The massed mean z-score was -0.18 while the spaced practice mean was +0.18, yielding an effect size of 0.37, a small to moderate effect.

The correlated *t*-test for the application data was also significant indicating that there was a statistically significant difference between types of practice and the students' ability to apply information to new situations. The effect size here was 0.66, a fairly large effect for an educational study, demonstrating both practical as well as statistical significance.

The results of the Likert interest survey were not statistically significant (Table 3). At least as far as these two topics are concerned, the type of practice did not have an influence on student interest in the topics.

In addition to the statistical tests, students were interviewed concerning the two types of practice (Table 4). Generally, most students believed that they learned the material and remembered it better with spaced practice. They seem to be evenly divided in their attentiveness to the material relative to the two kinds of practice.

Some of the responses of students to the survey questions are useful in interpreting the numbers. Students who preferred spaced practice gave these kinds of reasons: "The spaced practice was short and easy to concentrate on," "I liked doing the problems over several days because it helps one learn it and remember it rather than study it and forget it," "I like spaced . . . you don't get bored and start cheating or just not trying." Of the 19 students who preferred massed practice, these were typical reasons: "I prefer to do it all at one time because you don't have to remember the stuff . . . When I do it all spaced out, I have to keep referring to my notes and think about what I'm doing [sic]," "I would prefer to do it all in one day so that you can get it over with," "I preferred this (massed) because I didn't have to remember

TABLE 3

Likert interest scale results.

	Hooke			String		
	Spaced	Massed	Total	Spaced	Massed	Total
Mean	0.1	-1.4	-0.06	-0.8	0.8	0.0
Standard Deviation	5.1	6.5	5.8	6.5	5.1	5.8
Range	-9–+9	-10–+14	-10–+14	-12–+11	-12–+8	-12–+11
Number	24	22	46	22	24	46

things day after day." Although not firm evidence, there seems to be a trend in these answers which indicates that students who are interested in mastering the material prefer spaced practice, while students who just want to "complete the course" prefer the massed practice.

Fifteen of the students felt that they learned better using massed practice. Their reasons seemed to center around the perception that the material was taken more seriously when there was a lot of it at once. The 29 who preferred spaced practice felt it was more effective because it forced them to recall information many times, did not require them to absorb all the information in one sitting, and provided time to think about the subject matter.

A Post Study Analysis

During the interview process, it became apparent that massed practice benefited considerably from a Hawthorne and a novelty effect. As the study continued, many students became less inclined to give their full effort. It was hypothesized that the effect size between massed and spaced practice would actually be greater if the Hawthorne and novelty effect were not present. To test this hypothesis, students were asked to place themselves into one of two groups by checking a statement on a survey. The statements were: 1) I gave approximately the same effort to all sheets. I did not consciously let down my effort to do or understand the problems by any significant amount, 2) The novelty of the study wore off after a while. I did not apply myself to the same degree as I did initially to the problems at the end of the study. Students were instructed that to qualify for the second group, they must have made a conscious decision not to try. Seventeen students identified themselves as belonging to the first group. In an attempt to examine the effect of controlling effort expended on the practice material, a multivariate, matched pairs analysis was run on the seventeen students who claimed equal effort throughout the study. Their data yielded an F-value of over 13 and significance at less than the 0.001 level. Like the original study, the multivariate test was followed by matched pairs t-tests using a reduced level of significance. The results show an effect size of 1.0 for the achievement test and 1.1 for the application test, both in favor of spaced practice. Although such post-study analysis should be viewed with a degree of skepticism, the possibility of effect sizes of this magnitude certainly demands further study.

DISCUSSION

There is a need to identify and test teaching strategies which can be useful in increasing students' long term knowledge. As curriculum modifications strive to include effective strategies, solid evidence from actual classrooms is needed to convince teachers of their effectiveness. In the preponderance of psychological studies exactly the same items are practiced (Wild and Payne 1983), while in this study the content was practiced through different questions and problems. In addition, psychological studies may use an intervening time of seconds as opposed to days as used in the present study.

Although no research had been previously completed on whether spaced practice is effective for more complex tasks like physics (Glover et al. 1990), the present study indicated that students who used spaced practice were better able to solve new kinds of problems. The application test which measured this effect, however, was very difficult, resulting in very low scores. This result should be studied further with a less difficult test.

To address the practical concerns of classroom teachers, massed and spaced practice were given meaning in the context of classroom teaching. Massed practice, being a

TABLE 4

Interview results.

Question	Spaced Practice	Massed Practice	No Opinion
Which type of practice did you prefer?	30	19	0
Which type of practice helped you learn better?	29	15	5
Which type of practice helped you remember the material better?	30	10	9
Did you pay more attention to material in one type of practice?	23	22	4
Which type of practice was the most boring?	10	29	10

large amount of practice done at one time, is the common assignment after each topic is taught. Spaced practice involves teaching the topic and then providing shorter practice opportunities or revisits to the topic over a period of several weeks or more. Practice was provided on the facts, concepts, and generalizations of the material that was taught, and it was not simply the same type problem or question over and over again. Thus the practice in this study closely paralleled the type of practice found in most textbooks and assigned by most teachers.

It is important to emphasize that the results of the present study should not be interpreted to mean that simply providing spaced practice will provide students with adequate means to construct their knowledge base. The subject matter in this study was originally taught using a learning cycle style involving exploration, formulation of generalizations, and application of the generalizations. This study has sought to examine ways that knowledge, once acquired, can be readily accessed at distant, future dates. A structured revisiting of topics throughout a course, in this instance through spaced practice, seems to increase the ability of students to not only recall the material, but to use the material with much greater facility in the solving of new types of problems.

The Likert Scale measuring interest level was included in the study to test the hypothesis that the reason that spaced practice is more effective is that it is more interesting (Ausubel 1966). This study did not verify that hypothesis as the results of the Likert survey were equivocal. Although no interest advantage was found, neither was it found that spaced practice decreased interest. Teachers can use spaced practice without fear of alienating students. To control as many variables as possible, the practice used in the present study was quite sterile, basically a worksheet. If the results of the study are applied in classrooms, every effort should be made to make the spaced practices interesting and provide variety. For example, either in addition to or in place of the practice sheets, various techniques such as group work, homework, board work, discussions, computer programs, and/or lab activities can be used to deliver the spaced practice. Coupled with other effective teaching techniques, spaced practice can be a part of a popular course with high enrollments and high achievement.

Although a particular mechanism for the success of spaced practice was not isolated in this study, some speculation is possible. It is possible that the multiple exposures afforded by spaced practice allow for more connections to other topics. Students think about and study different topics on different days. The more connections that exist in memory, the easier it could be to locate information that has been stored in long term memory (Houston 1976). This possibility also helps to explain superior problem solving skills. The ability to solve problems seems to depend on one's ability to connect different topics and procedures in different ways. Perhaps these hypothetical additional pathways created in spaced practice allow this to happen more easily. Of course, just having the information available may be all that is necessary to enhance problem solving ability.

Another possibility is that students get more feedback on their success of encoding information into long term memory in spaced practice. In massed practice, insufficient time has elapsed to test the retrieval of information in long term memory. Once students know that the retrieval of information from long term memory for a particular item has failed, they can try to re-encode the material in a more effective way (Farr 1987).

LITERATURE CITED

Ausubel, D. P. 1966 Early versus delayed review in meaningful learning. Psychol. in the Sch. 3: 195-198.

Dempster, F. N. 1988 The spacing effect: A case study in the failure to apply the results of psychological research. Amer. Psychol. 43: 627-634.

Farr, M. J. 1987 The Long Term Retention of Knowledge and Skills. Springer-Verlag, New York, NY. 116 pp.

Glenberg, A. M. 1977 Influences of retrieval processes on the spacing effect in free recall. J. Exp. Psy. 3: 282-294.

Glover, J. A., R. R. Ronning, and R. H. Bruning 1990 Cognitive Psychology for Teachers. Macmillan, New York, NY. 416 pp.

Grote, M. G. 1995 Distributed versus massed practice in high school physics. Sch. Sci. Math. 95: 97-101.

Hintzman, D. L. 1974 Theoretical implications of the spacing effect. *In*: R. L. Solso (ed.), Theories in Cognitive Psychology: The Loyola Symposium. Lawrence Erlbaum, Potomac, MD. 386 pp.

Houston, J. P. 1976 Fundamentals of Learning. Academic Press, New York, NY. 364 pp.

Lu, P. K. 1978 Three integrative models of kinetic structure in teaching astronomy. J. Res. Sci. Teach. 15: 249-255.

Melton, A. W. 1970 The situation with respect to the spacing of repetitions and memory. J. of Verb. Learn. and Verb. Beh. 9: 596-606.

Rosenshine, B. and R. Stevens 1986 Teaching functions. *In*: M. C. Wittrock (ed.), Handbook of Research on Teaching, Third Edition. Macmillan Publishing Company, New York, NY. pp. 376-391.

Stevens, J. 1986 Applied Multivariate Statistics for the Social Sciences. Lawrence Erlbaum Associates, Hillsdale, NJ. 515 pp.

Wild, H. M. and R. B. Payne 1983 An experimental test of two theories of spaced practice effects. Percept. and Motor Skills 56: 267-274.