

11-19-2003

# Optimizing Long-Term Retention of Abstract Learning

Danielle Mazur

*University of South Florida*

Follow this and additional works at: <https://scholarcommons.usf.edu/etd>

 Part of the [American Studies Commons](#)

---

## Scholar Commons Citation

Mazur, Danielle, "Optimizing Long-Term Retention of Abstract Learning" (2003). *Graduate Theses and Dissertations*.  
<https://scholarcommons.usf.edu/etd/1427>

This Thesis is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact [scholarcommons@usf.edu](mailto:scholarcommons@usf.edu).

Optimizing Long-Term Retention of Abstract Learning

by  
Danielle Mazur

A thesis submitted in partial fulfillment  
of the requirements for the degree of  
Masters of Arts  
Department of Psychology  
College of Arts and Sciences  
University of South Florida

Major Professor: Doug Rohrer, Ph.D.  
Michael Brannick, Ph.D.  
Darlene DeMarie, Ph.D.

Date of Approval:  
November 19, 2003

Keywords: spacing, distributed practice, massed practice, overlearning, spacing effect

©Copyright 2003, Danielle Mazur

## Table of Contents

List of Figures	ii	
Abstract	iii	
Introduction	1	
Theoretical Accounts of the Spacing Effect	3	
Empirical Demonstrations of the Spacing Effect	6	
Experiment 1	12	
Methods	12	
Participants	12	
Materials	12	
Design	13	
Procedure	13	
Results and Discussion	14	
Experiment 2	16	
Methods	18	
Participants	18	
Materials	18	
Design	18	
Procedure	18	
Results and Discussion	19	
General Discussion	20	
References	24	
Appendices	26	
Appendix A: A sample set of 5 matrix multiplication problems	27	
Appendix B: A sample transparency set	28	
Figure Captions	31	
Figures		
Figure 1	Illustrative Data	32
Figure 2	Experiment 1	33
Figure 3	Experiment 2	34

## List of Figures

<i>Figure 1.</i> Illustrative Data.	32
<i>Figure 2.</i> Experiment 1.	33
<i>Figure 3.</i> Experiment 2.	34

## Optimizing Long-Term Retention of Abstract Learning

Danielle Mazur

### ABSTRACT

In a distributed learning strategy, study time is spread across multiple study sessions, without increasing total study time. The benefits of distributed practice, also known as spaced practice, on learning of rote-memory tasks (e.g., spelling, addition, and cued-recall of word pairs) are well known. However, few researchers have looked at the effects of distributed practice on the learning of abstract materials (e.g., physics problems, logical deductions, and algebra). We examined the effects of distributed practice on learning the abstract task of matrix multiplication. In Experiment 1, we taught participants matrix multiplication in either a massed (i.e., 0-day interstudy gap), or distributed (i.e., 7-day interstudy gap) format and tested students at 2 or 21 days after completion of the last study session. Results showed no significant differences between the massed and spaced groups. However, when only those participants scoring 80% or greater on study session one were included in the analyses, a benefit of spacing was seen at the 21-day retention interval. Although not statistically significant, this leads us to believe that spacing does have benefits for abstract learning when the task is mastered initially.

Experiment 2 looked at overlearning as another learning strategy. In overlearning, all study takes place in one session, but participants continue to study after mastery of material has been achieved. It is commonly accepted that overlearning is a

beneficial strategy, but it is unknown whether the benefits are worth the time invested. We assessed the effects of two levels of massed practice to gauge the benefits of overlearning on long-term retention. Participants completed either 2 or 8 matrix multiplication problems (i.e., low or high massing, respectively) and were tested 1 or 4 weeks after the study session. Results showed a benefit of high massing when analyses included participants who mastered the material (i.e., scored over 50%) during the study session. However, this higher degree of learning was not particularly efficient, because this latter result suggests that overlearning may not be worth the time invested.

## **Optimizing Long-Term Retention of Abstract Learning**

Long-term retention of material should be a primary goal of educators. However, in today's schools, it is common for students to learn the material, take a test, and forget the material soon after. What is the point of attending school if you forget most of the material?

Therefore, we should consider long-term retention when we evaluate learning strategies. One such strategy is distributed learning, also known as spaced practice. In a distributed learning procedure, total study time is held constant but is spread across multiple study sessions. For example, if a student solves five long division problems on one day and five others one week later, the strategy is distributed. If the student solved all 10 problems on the same day, it would be considered massed practice. At longer retention intervals, distributed practice produces better retention than massed practice, and this phenomenon is termed the spacing effect (e.g., Dempster, 1988; 1989).

Though many teachers are aware of the benefits of a distributed learning format, relatively few instructors make use of distributed practice. There are several possible reasons for this. First, Dempster (1988) cites a lack of empirical evidence for the spacing effect in the classroom. Much of the research has been conducted in laboratories using simple, ecologically invalid learning tasks (e.g., memorizing a list of words). While studies such as these show the merits of spaced practice, these findings do not provide information useful to teachers who have entire lesson plans to prepare, often with too much material as it is. Teachers may feel that it is not a sound decision to rework their

entire curriculum based on studies that show only laboratory findings.

Another possible reason teachers do not make use of spacing is that they may feel distributed practice does not benefit abstract learning. In fact, Ernest Rothkopf once said, “spacing is the friend of memory, but the enemy of induction” (J. Anderson, personal communication to Doug Rohrer). Induction can be thought of as a type of abstract thinking. Therefore, we plan to investigate whether distributed practice shows benefits for abstract tasks. A rote-memory task is a low-level task that involves strict memorization of items. Some examples of rote-memory tasks include spelling, memorization of multiplication tables, or learning an association between two items (e.g., house – casa). A higher-level (or abstract learning) task is one that requires more than a simple association between concepts. Specifically, an abstract task requires one to perform a task with a novel stimulus. For example, such a task might involve the learning of a procedure (such as matrix algebra or physics problems).

Hence, while rote learning requires one to memorize several items, one for each new task (e.g., in learning multiplication tables, knowing  $2 \times 2 = 4$  does not help you with the problem  $3 \times 3 = ?$ ), abstract learning enables one to apply his or her knowledge to novel tasks. For example, once a student learns how to solve a matrix multiplication problem (refer to Appendix A), he or she can apply that knowledge to solving any problem of that type. Moreover, this does not apply to only quantitative tasks. Logical deductions ( $p$  implies  $q$ ;  $p$  is true; therefore,  $q$  is true) and category learning are other types of abstract tasks. An example of category learning includes a student learning the characteristics of a specific group, and then using that knowledge to place novel items



into that group. Items such as these are developing the logical thought process of students, and many teachers would argue that this is one of the most important things you can teach students.

Very few American textbooks make use of distributed practice. One exception are textbooks produced by Saxon Publishers, whose mathematics textbooks use spaced practice (Saxon, 1997a; 1997b). These textbooks employ a spaced-practice approach. Specifically, each lesson includes a few dozen problems, as is true with other textbooks. In the Saxon textbooks, however, only a small subset of problems concern the lesson of the day, with the remaining problems drawn from previously studied topics. This provides continued practice on older topics as well as new material.

We investigated whether spacing improved retention for a higher-level task. Although the spacing effect has been demonstrated in many studies with rote-memory tasks, the few studies demonstrating a spacing effect with abstract tasks appear to be confounded. The relevant literature is described below, beginning with a description of the spacing effect.

#### Theoretical Accounts of the Spacing Effect

Researchers have put forth several theories to explain why the spacing effect exists, but none fully account for the data. One of these is commonly called the voluntary attention hypothesis (Dempster, 1988; 1989). The voluntary attention hypothesis proposes that a subject chooses to pay less attention to massed repetitions and more attention to spaced ones. Individuals may find that massed repetitions become boring, since little or no time passes between repetitions. On the other hand, the time that passes

between spaced repetitions may make repeated lessons seem more interesting. A correlational study by Dempster (1986, as cited in Dempster, 1988) supported this theory. Participants read a passage twice, with either 30 minutes or 5 minutes between readings; and then answered a questionnaire that included questions about the passage, their perceived attention, and interest. He found that participants in the spaced group tended to rate the passage more interesting and also had a higher accuracy rate. Participants in the spaced group also reported paying more attention. Therefore, participants tended to learn more when they paid more attention and found the material more interesting.

Although the voluntary attention hypothesis sounds promising, there are some findings that are inconsistent with this theory. First, the spacing effect has been found with pre-school children who have limited voluntary control over their thought processes (Rea & Modigliani, 1997, as cited in Dempster, 1989). Second, researchers have manipulated conditions to make participants pay more attention to massed presentations, but these studies failed to eliminate the spacing effect (Hintzman, 1976, as cited in Dempster, 1989). Third, the effects of spacing have been observed in incidental learning tasks, where little attention was paid to the task at hand (Rowe & Rose, 1974, as cited in Dempster, 1989). Hence, these findings are not consistent with the voluntary attention hypothesis.

Another theory attributes the spacing effect to encoding variability (Dempster, 1989). There are many ways that information can be encoded, and generally, the more ways it is encoded the better the information is recalled. It is thought that if information is presented in different contexts (i.e., locations, times, methods), there will be more

retrieval routes to access the material within memory. Spaced repetitions may allow for encoding to take place in different contexts, thereby allowing for more than one effective retrieval route. However, this theory is also unable to explain all previous findings. Many studies have purposely manipulated changes in context and have resulted in declined recall rather than the predicted increase (Dempster, 1989). Therefore, the encoding variability theory cannot fully explain the presence of the spacing effect.

A third account of the spacing effect is known as the rehearsal hypothesis (Dempster, 1989), which attributes the inferior utility of massed repetitions to the inability of participants to rehearse the material immediately after presentation. By contrast, with spaced repetitions, the first presentation is followed by a period of time that may allow for rehearsal of the presented material and increase subsequent recall. However, there are some studies that are inconsistent with the rehearsal hypothesis and lead one to question its viability. For example, the spacing effect has been found with very young children who are not known to rehearse material (Rea & Modigliani, 1987, as cited in Dempster, 1989).

In summary, it is important to note that there is no clear or accepted theory regarding why spacing works (Anderson & Schooler, 1991; Baddeley, 1978). Although there are multiple theories, none fully explain all of the findings.

Experiment 1 examined the spacing effect in higher-learning tasks. As discussed below, few researchers have examined higher-learning tasks, but many have demonstrated the spacing effect in rote-memory tasks.

## Empirical Demonstrations of the Spacing Effect

Evidence of the benefits of distributed practice dates back to the 19<sup>th</sup> century (Austin, 1921; Ebbinghaus, 1885/1964; Perkins, 1914). Nellie Perkins looked at the ability of students to learn nonsense syllables. Each pair was presented the same number of total times, in either a massed format (one-day between study sessions) or a distributed format (2-4 days between study sessions). Two weeks after final exposure participants were tested and results showed that the distributed condition yielded the highest percentage of recall.

Austin (1921) also compared distributed practice and massed practice, though she utilized the tasks of free recall and cued recall. Participants were taught material through readings about a specific topic. The results showed no difference between distributed practice and massed practice at a 0-day retention interval (time between last study session and test), but a benefit of spacing was shown at a longer retention interval of 2 weeks. Figure 1A shows this relationship using hypothetical data. Looking at Figure 1, one can see that at the shorter retention interval, there is no difference between massed and spaced practice. However, at the longer retention interval, there is a significant benefit of spaced practice over massed practice.

Bloom and Shuell (1981) looked at the learning of a foreign language in a real-life classroom situation. Participants learned a list of vocabulary words in either a distributed or massed format, and a cued-recall test was given immediately following the last exposure and again a week later. For example, after learning the pair house-maison, the test would provide the word house and require the student to write in the word maison.

Performance was equal on the immediate test; yet the spaced-practice students performed substantially better than the massed-practice students when tested 1 week later (as in the hypothetical data shown in Figure 1A). For this study, then, there was no benefit of distributed practice when tested immediately, but benefits were seen in the long-term. Therefore, an interaction can be seen between retention interval and type of practice.

Glenberg and Lehman (1980) studied participants' memory for word pairs. Retention intervals ranged from 0 to 7 days, and a free-recall test was used to test memory. The results showed the classic spacing effect, as illustrated by the hypothetical data in Figure 1B. At the 0-day retention interval, those in the massed practice condition had higher recall than those in the distributed practice condition. However, at the 7-day retention interval, a benefit for distributed practice was shown. Here, retention interval and learning strategy exhibit a crossover interaction.

Rea and Modigliani (1985) examined the ability of 3<sup>rd</sup> graders to learn multiplication facts and spelling lists. Students were taught the material in either a spaced or massed format. An oral test was given one minute following the final presentation of the last fact to be learned. This test was then followed by a written test of the same material. The results showed that retention was increased when the spaced format was employed for both the multiplication and spelling tasks, and the benefits were realized for both the oral and written tests.

Fishman, Keller, and Atkinson (1968) used computerized spelling drills to compare massed and distributed practice. The within-subject study consisted of 5<sup>th</sup> graders. The students learned sets of words, some in a massed format, and others in a

distributed format. All words were presented three times each, with the distributed words presented every other day, and the massed words presented in one day. Ten or 20 days after the end of training, students were asked to type the dictated words. The results showed that distributed practice produced higher retention at both retention intervals.

Though the previous studies utilized brief retention intervals, Bahrick and Phelps (1987) tested students up to 8 years after study. In the task, students studied 50 English-Spanish word pairs (e.g., house-casa). The participants were given a cued-recall test (e.g., house -?) followed by a recognition test. The authors found that a spacing of 30 days between the two study sessions optimized accuracy. In fact, this optimal interstudy gap of 30 days produced a level of accuracy that was two and a half times greater than that of the massed condition (i.e., interstudy gap = 0).

In a remarkable study, Bahrick, Bahrick, Bahrick, and Bahrick (1993) examined long-term retention of foreign language vocabulary after a learning phase lasting 6 months to 4 years. The four authors served as participants in the study, and they learned foreign language vocabulary words. Results again revealed a benefit of spacing at retention intervals up to 5 years, with recall accuracy greatest at the longest interstudy gap of 56 days.

However, while Bahrick and Phelps (1987) and Bahrick et al. (1993) employed remarkably long retention intervals, both studies confounded the distribution of practice. Specifically, during the practice sessions, participants learned to criterion. That is, they cycled through the list of vocabulary words with feedback until 100% accuracy was attained. Naturally, spaced groups took longer to reach criterion than did the massed

groups. Therefore, the spaced groups received more study time than the massed groups, and this additional study time likely contributed to the observed effect of spacing. Keeping study time constant across groups is one of the hallmarks of spacing, and therefore the confound greatly effects the results of these experiments.

In addition to this drawback, few researchers have investigated the effects of spacing on learning of higher-level material. For example, learning the Spanish word for house requires nothing other than a simple association between “house” and “casa,” and such rote learning differs qualitatively from those requiring abstract inferences. For instance, when solving a math problem, students cannot simply memorize all possible questions, as they must learn to generalize their skills to novel problems.

Grote (1995) did one of the few studies of the spacing effect with abstract learning material. Grote examined high school physics, and his students learned two topics in either a massed or distributed format. Tests were given at 2, 4, and 6 weeks. Grote’s results showed a definite effect of spacing, with the distributed-practice group outperforming the massed practice group at all test intervals.

Although the Grote (1995) study is suggestive, the results are confounded. The distributed-practice students were tested 2 weeks after their last exposure to the material, while there was approximately 6 weeks between exposure and test for the massed-practice students. This confound is likely to have affected the results of the study, because retention declines with the passing of time.

Gay (1973) also examined the effects of distributed practice on the learning of abstract materials. Participants were taught algebra and geometry in either a massed

format (1 day between study sessions) or spaced format (1 or 2 weeks between study sessions). Though results showed no significant differences between groups, the study was confounded in the same manner as the Grote (1995) study described above. Specifically, the test was given 3 weeks after completion of the first study session, not the most recent study session. Hence, the retention interval was not constant across groups, as the interval between the second study session and the test ranged from 7 days for the spaced group to 20 days for the massed group. Since retention declines with the passing of time, the results of Gay's study do not address the spacing effect.

Saxon Publishers (Saxon Publishers, 2001) cite considerable data showing the improved performance of students who have used the Saxon math books. While impressive, these data are confounded, because the students were not randomly assigned. Each study compared the performance of students using a non-Saxon textbook with the performance of students using the Saxon textbook during a different school year. While each study looked at only one school, there were many group differences between the Saxon and non-Saxon students. The performance of one group of students was compared with a completely different group of students who may have had different math abilities. There may also have been different teachers, different teaching methods or changes in curriculum introduced during that time. Thus, it is impossible to be sure the improvement seen was due solely to the use of these textbooks. Other evidence presented by Saxon only included data of students while using the Saxon textbooks. Without any comparison data it is impossible to know if the performance was indeed affected by the Saxon-textbook.



In summary, the prior research reveals the benefits of spacing at sufficiently long retention intervals, but this research is generally characterized by three limitations. First, the majority of research has focused on low-level rote memory tasks such as spelling, addition, and memory of word pairs (e.g., Bloom & Shuell, 1981; Fishman, Keller, & Atkinson, 1968; Rea & Modigliani, 1985). It is important to know if there is an effect of spacing in higher-learning tasks like math, physics, and category learning. This would have huge implications for the classroom. This is because a large portion of what is learned in school is of an abstract nature. If the effects of distributed practice are seen for these types of tasks, that means that the method would be extremely useful in improving the learning of these abstract tasks. A second limitation of prior research is that many previous studies used a fairly short retention interval (e.g., Glenberg & Lehman, 1980; Rea & Modigliani, 1985). It is important to know if the effects of spacing are long lasting. If not, then distributed practice is not particularly beneficial to long-term learning. A final limitation of previous studies is a reliance on ecologically invalid tasks. For example, the recall of a word list (Glenberg & Lehman, 1980) is a task unlike any encountered in everyday life. More valid tasks include, for example, cued-recall (e.g., casa – house) and solving of math problems.

Experiment 1 addressed the limitations discussed regarding spacing studies. Specifically, we used the task of matrix multiplication which is an abstract, ecologically valid task. We also used a longer retention interval than many of the previous studies. The results of Experiment 1 led us to examine a learning strategy known as overlearning in Experiment 2.

## Experiment 1

In this study, participants learned the higher-learning task of matrix multiplication. We chose matrix multiplication for two reasons. First, we believed that few people are familiar with the task. Second, the task requires only the prerequisite skills of multiplication and addition, which are skills that most people possess.

We employed a 2 x 2 design for the matrix task. One independent variable was study gap (0 days, 7 days), which equals the time period between the two study sessions. The second independent variable was retention interval (2 days, 21 days), the time between the last study session and the test. We examined whether there exists spacing superiority at the longer retention interval.

### Methods

*Participants.* One hundred ten University of South Florida undergraduates participated for course credit. Nineteen males and 91 females comprised the sample, with an average age of 20.8 years.

*Materials.* Materials consisted of two booklets, one for each study session. A two-page sheet was used for study session one of the matrix multiplication task. The first page requested some basic information of the subject (e.g., name, gender, and knowledge of matrix multiplication). The front side of the second page consisted of 10 matrix multiplication problems. The reverse side had 5 additional problems. Study Session Two used the same format, though with different problems. The 15 problems in each study session were divided into sets of 5, and each set included 5 types of matrix multiplication problems. These five include row x column, row x square, square x

column, square x square, and column x row, and an example of each is shown in Appendix A. The answers to all problems consisted of matrices with single digit entries only, and none of the problems were repeated.

*Design.* We had two levels of interstudy gap (0 and 7 days) and two retention intervals (2 and 21 days), and both factors were between-participants manipulations. Thus, each participant was randomly assigned to one of four groups.

*Procedure.* The first study session included 15 problems and took place on the same day for all participants. First, participants were instructed to complete the cover sheet. Participants were then taught how to perform matrix multiplication. Using transparencies (Appendix B), the instructor illustrated each step for 5 sample problems (4 minutes), while participants followed along in their packet. Participants were then instructed to attempt 5 problems on their own. After attempting each problem, transparencies were displayed that outlined the correct way to solve the problem, without any oral feedback. This gave the participants time to check their work. Participants had 30 seconds to attempt the problem, and 30 seconds to view the transparency. After completing these five problems, the students were instructed to turn over the sheet and attempt 5 problems without feedback (3 minutes). Booklets were then collected.

The second study session took place immediately following the first for the 0-day interstudy gap group and 7 days later for the 7-day interstudy gap group. Participants attempted 10 problems, and received feedback via transparency as outlined above, with no instructor feedback. Again, participants had 30 seconds to try each problem and 30 seconds to view the transparency. Participants then turned over their sheet and attempted

to complete 5 problems on their own without feedback (3 minutes). Booklets were then collected.

A test consisting of 10 matrix multiplication problems was given either 2 days or 21 days after the second study session.

On the last day of class, all participants were debriefed about the experiment and results were presented to the class.

### Results and Discussion

The results of all participants are shown in Figure 2A. Analysis of the results through ANOVA revealed a main effect of retention interval ( $F(1,106) = 5.76, p < 0.05$ ). However, the main effect of gap was not statistically significant ( $F(1,106) = 0.24, p > 0.05$ ). There also was no significant interaction between gap and retention interval ( $F(1,106) = 0.057, p > 0.05$ ). Specifically, there was a slight benefit of massing at the 2-day retention interval, although not statistically significant. At the 21-day retention interval, the performance of the massers fell off to almost equal that of the spaced group.

However, this initial analysis included all participants, including those participants who performed poorly on study session one. Naturally, if the task was not learned during the study session, participants cannot be expected to perform well on the test. Therefore, we conducted a second analysis on these data for participants who correctly answered 80% of the problems in study session one (as shown in Figure 2B). This included 48 of the 110 participants. This analysis is perhaps the most telling, because these participants are the ones who truly learned the material during the learning session. This figure shows a small advantage for the massed group at 2 days and a small

advantage for the spaced group at 21 days, but this interaction is not statistically significant. Nevertheless, the trend of the data suggests that spacing is beneficial at longer retention intervals.

Multiple linear regression was used to determine which of the independent variables were significant predictors of test score. Performance on study session one, interstudy gap, and retention interval were all used as possible predictors. The overall model was significant,  $F(3,106) = 25.40, p < 0.001$ . This model predicted approximately 42% of the variability in test scores (adjusted  $R^2 = 0.40$ ). The regression equation that resulted was  $\text{test} = 0.391 + 0.00164 * \text{gap} - 0.00891 * \text{RI} + 0.642 * \text{SS1}$ . Gap was study gap; the time that passed between study session one and study session two. RI was retention interval; the time that passed between study session two and the test. SS1 was performance on the five problems students attempted at the end of study session one. Retention interval and study session one performance were significant predictors of test score. The b-weight for retention interval was  $-0.00891, t(1) = -3.805, p < 0.001$ . This shows that as retention interval increased, there was a corresponding decrease in test performance. The b-weight for study session one performance was  $0.642, t(1) = 8.122, p < 0.001$ , which shows that as performance on study session one increased, performance on the test increased as well. The b-weight for interstudy gap was not significant and therefore is not a significant predictor of test score.

One limitation of the study was that many participants were previously exposed to matrix multiplication. By self-report, only 42% of the participants had never seen matrix multiplication before. Therefore, many of the massers (63%) had been previously

exposed to the task and were effectively spacers, because they learned matrix multiplication in a previous situation. Hence, during session one, they were exposed a second time. This could have elevated the scores of the massed groups at both retention intervals.

Another limitation was that the retention interval may not have been long enough, because the benefits of spacing are known to increase with retention interval. It is possible that the 21-day retention interval may have been too short, and a benefit of spaced learning may have been seen if the retention interval was longer.

A third limitation was the difficulty of the task. It was expected that matrix multiplication would be fairly easy for all participants to learn, but this was not the case. There was quite a bit of variability in participants' abilities. Some participants learned the material easily and excelled at the task. However, many participants failed to master the task during the first study session. Therefore, this may have reduced the performance of the spacers, because the absence of learning during study session one means there was no spacing.

In summary, although the results were not significant, the trend shown in Figure 2B is promising. We therefore believe that spaced learning is beneficial for participants who master an abstract task.

## Experiment 2

The results of Experiment 1 led us to some additional questions. While massing, participants often master the material before the end of the study session. Any continuation of study beyond mastery in the same session is known as overlearning. For

example, suppose a student has 5 spelling words to learn and attempts all 5 words during 10 trials, with the correct answer supplied after each attempt. If the student correctly spells all 5 words after 5 trials, the remaining 5 trials are considered overlearning. By this definition, many of the massers in Experiment 1 were overlearning. This brought us to the question of how beneficial is overlearning.

Previous research with overlearning has revealed a benefit of the study strategy. A meta-analysis of 15 studies by Driskell, Willis, and Cooper (1992) found that overlearning produced higher levels of retention. This is commonly accepted and therefore overlearning is often used as a learning strategy. For example, the majority of mathematics textbooks have problem sets that consist of many problems all of the same type. Many times, the student masters the problem type after a few tries, and thereafter is overlearning. Another example involves foreign language learning. When learning vocabulary, students may be drilled repeatedly within the same session and long after the vocabulary words are mastered. The question at hand is not only whether overlearning is effective, but also whether or not overlearning is worth the time that is invested. For example, a doubling of study time may produce less than twice the retention.

In Experiment 2, we looked at the benefits of low versus high massing with matrix multiplication. We wanted to see if the benefits of overlearning were worth the extra study time the participants invested. The low and high learners completed 2 or 8 matrix multiplication problems, respectively, and were tested at retention intervals of 1 or 4 weeks.

## Methods

*Participants.* One hundred undergraduate students at the University of South Florida participated in the study. Fifteen males and 85 females participated in the study, and the mean age of the participants was 20.1 years.

*Materials.* Materials consisted of two sheets, one for each of the two study times (low or high). The front side consisted of 4 sample matrix multiplication problems, all of the square-by-square type, as shown in Appendix A. The reverse side consisted of either 2 or 8 problems of the square-by-square type. The answers to all problems consisted of matrices with single digit entries only, and no problems were repeated.

*Design.* There were two levels of study time (low and high) and two retention levels (1 or 4 weeks), and both factors were between-participants manipulations. Study time was defined as low or high based upon number of problems attempted (2 and 8, respectively). Each participant was randomly assigned to one of four groups.

*Procedure.* All participants learned matrix multiplication in the same session. Using the same procedure as in Experiment 1, the instructor taught the task of matrix multiplication. The students were then instructed to attempt each of the four problems on the front side of the sheet. After attempting each problem, transparencies were displayed that outlined the correct way to solve each problem (Appendix B). Students had 30 seconds to attempt the problem and 30 seconds to view the transparency. After completing these 4 problems, students then turned over the sheet and attempted 2 or 8 problems on their own without feedback. The sheets were then collected.

A test consisting of 20 novel matrix multiplication problems was given either 1 or



4 weeks later. Students had 20 minutes to complete the test.

## Results and Discussion

We analyzed the data of all participants (shown in Figure 3A) and an analysis of variance revealed a main effect of retention interval ( $F(1,96) = 4.69, p < 0.05$ ). Thus, as the retention interval increased, test performance decreased. The high-massers recalled more than the low-massers, but this difference was not statistically significant ( $F(1, 96) = 0.54, p > 0.05$ ). In addition, there was no reliable interaction between study time and retention interval ( $F(1,96) = 0.25, p > 0.05$ ).

As in Experiment 1, however, many participants failed to master matrix multiplication during the study session. Hence, these participants could not be expected to recall a skill they had not actually learned. Therefore we conducted an analysis including those participants who scored at least 50% during the study session (at least one out of two, or 4 out of 8). These results are shown in Figure 3B. This analysis included 85 of the 100 participants, and although the high-massing participants remembered more than the low-massing participants at both retention intervals, the difference was not significantly significant. Fifty-eight of the 100 subjects received a perfect score on the study session. This shows that a little over half of the subjects completely mastered the material.

In this study, the greater time put in at study did not pay dividends proportional to the time invested. In fact, the higher study groups studied four times as much as the lower study groups (8 vs. 2 problems, respectively), yet this extra study time resulted in performance not even twice that of the low study group at either retention interval. In

other words, four times the study yielded less than twice the accuracy of retention. Therefore, this result suggests that overlearning is an inefficient learning strategy.

### General Discussion

The results of Experiments 1 and 2 led us to some conclusions regarding how to optimize long-term retention of abstract material. While the spaced learners did not recall significantly more than the massers in Experiment 1, the results were still in the direction of a spacing effect, especially at the longer retention interval. Previous research with spacing has focused primarily on rote learning tasks, and we investigated the benefits of spacing for abstract learning tasks.

Experiment 2 examined overlearning as a learning strategy. While often employed in the classroom, particularly in mathematics, the benefits of overlearning are not proportional to the total study time. Hence, students may put in extra hours of studying only to see a small long-term increase in retention.

The majority of teachers do not utilize spacing in their lesson plans. The massing commonly employed often leads to limited long-term retention. For example, I can remember learning spelling lists in elementary school. My homework consisted of writing the words over and over again for a week, and never seeing the words again until it was time for a subsequent test. Studies of the spacing effect suggest that it would have been more effective to space the practice by studying both new and old words each week. Moreover, this kind of spaced practice could have been less boring. If a teacher lectured continuously on the same subject, my concentration would wane within a half an hour. Had the lessons been spaced, my attention may have remained better focused.

Although the benefits of spacing fell short of significance in the present study, we still believe that spaced practice is worth implementing in the classroom. When looking at students who mastered the material, there was a small non-significant benefit of spacing. Had there been more participants that truly learned the material, we may have seen more of a spacing effect. Therefore, for the teaching of abstract tasks, we believe that a spaced approach would lead to increased long-term retention. Because this should be a primary goal of education, long-term retention should also be a primary concern.

The results of Experiment 2 suggest that the benefits of overlearning may be overrated, because the long-term benefits of overlearning were not proportional to the time invested by the student. The time spent overlearning could be used to learn different material. However, many teachers still employ overlearning in the classroom. With regard to my spelling drills described above, the sheer number of repetitions ensured that I overlearned the words. That is, after a few times, I more than likely knew how to spell the word. Hence, the rest of the practice time was spent overlearning. This time could have been spent learning more words and expanding my vocabulary. The same is true for math lessons. Writing out  $2 \times 2 = 4$  twenty times is redundant, as the multiplication fact is most likely mastered long before the assignment is completed. Again, would it not be more beneficial for the student to learn several multiplication facts rather than overlearn one fact?

Another example of overlearning in mathematics involves homework assignments. For example, in learning the quadratic equation, a homework set may consist of 30 or so problems requiring a student to solve a quadratic equation. After a

few problems, the student has likely mastered the task. It may be more beneficial for the student to solve 5 quadratic equations followed by problems of other types.

It is important to note that the use of overlearning in the classroom may be directly linked to the lack of spaced practice. The typical math textbook includes many topics. A homework set (as discussed above) typically consists of 30 or so problems all of the same type. This is considered overlearning, because the student usually masters the material long before completion of the given set. If a teacher tried to implement spaced practice along with this overlearning, homework sets would become impossibly long. A better approach would be to limit the amount of overlearning, and instead give about a half dozen problems on several topics.

Future research with spacing should concentrate on its use with abstract tasks. It is difficult to find abstract tasks to utilize in this research because many of these tasks are taught in school. Perhaps the best way to conduct this research would be in a school setting, with students who have not been exposed to a task such as matrix multiplication. Although Gay (1973) and Grote (1995) both found evidence of a spacing effect with abstract tasks, both studies confounded retention interval with the manipulation of spacing versus massing. That is, as described in the introduction, the retention interval was shorter for the spacers than the massers, thereby giving the spacers an obvious advantage.

Another focus of future research should be the use of meaningfully long retention intervals to evaluate the long-term benefits of both spacing and overlearning. Although Bahrnick and Phelps (1987), and Bahrnick et al. (1993) both used extremely long retention

intervals and found promising results, these studies utilized a learning-to-criterion procedure that ensured the spacers received more practice than the massers.

Overall, one cannot overemphasize the importance of long-term retention to education. Educators should strive to teach students by methods that optimize long-term retention. Perhaps after more research is published, teachers will have the proof to push them into implementing the use of spaced practice in the classroom.

## References

- Anderson, J. R. & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, 2, 396-408.
- Austin, S. D. M. (1921). A study in logical memory. *American Journal of Psychology*, 32, 370-403.
- Baddeley, A. D. (1978). *Human Memory: Theory and Practice*, Revised Edition. Needham Heights, MA: Allyn & Bacon.
- Bahrick, H. P., Bahrick, L. E., Bahrick, A. S., & Bahrick, P. E. (1993). Maintenance of foreign-language vocabulary and the spacing effect. *Psychological Science*, 4, 316-321.
- Bahrick, H. P., & Phelps, E. (1987). Retention of Spanish vocabulary over eight years. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 344-349.
- Bloom, K. C, & Shuell, T. J. (1981). Effects of massed and distributed practice on the learning and retention of second-language vocabulary. *Journal of Educational Research*, 74, 245-248.
- Dempster, F. N. (1988). The spacing effect: A case study in the failure to apply the results of psychological research. *American Psychologist*, 43, 627-634.
- Dempster, F. N. (1989). Spacing effects and their implications for theory and practice. *Educational Psychology Review*, 1, 309-330.
- Driskell, J. E., Willis, R. P., & Cooper, C. (1992). Effect of overlearning on retention. *Journal of Applied Psychology*, 77, 615-622.
- Ebbinghaus, H. (1964). *Memory* (H. A. Ruger & C. E. Bussenius, Trans.). New York: Dover. (Original work published 1885, trans. 1913).
- Fishman, E. J., Keller, L., & Atkinson, R. C. (1968). Massed versus distributed practice in computerized spelling. *Journal of Educational Psychology*, 59, 290-296.
- Gay, L. R. (1973). Temporal position of reviews and its effect on the retention of mathematical rules. *Journal of Educational Psychology*, 64, 171-182.

Glenberg, A. M., & Lehmann, T. S. (1980). Spacing repetitions over 1 week. *Memory & Cognition*, 8, 528-538.

Grote, M. G. (1995). Distributed versus massed practice in high school physics. *School Science and Mathematics*, 95, 97-101.

Perkins, N. L. (1914). The value of distributed repetitions in rote learning. *British Journal of Psychology*, 7, 253-261.

Rea, C. P., & Modigliani, V. (1985). The effect of expanded versus massed practice on the retention of multiplication facts and spelling lists. *Human Learning: Journal of Practical Research & Applications*, 4, 11-18.

Saxon, J. (1997a). *Algebra I (3<sup>rd</sup> Edition)*. Norman, OK: Saxon Publishers.

Saxon, J. (1997b). *Algebra II (2<sup>nd</sup> Edition)*. Norman, OK: Saxon Publishers.

Saxon Publishers. (2001). *Saxon Mathematical Results*. Norman, OK: Saxon Publishers.

## Appendices



### Appendix A: Sample set of problems

A sample set of 5 matrix multiplication problems. One of each type of problem is represented, i.e., row x column, row x square, square x column, square x square, and column x row. All answers will consist of matrices with single digit entries only.

$$1. \begin{pmatrix} 1 & 2 \end{pmatrix} \bullet \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} \_ \end{pmatrix}$$

$$2. \begin{pmatrix} 2 & 3 \end{pmatrix} \bullet \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \_ & \_ \end{pmatrix}$$

$$3. \begin{pmatrix} 5 & 2 \\ 1 & 4 \end{pmatrix} \bullet \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} \_ \\ \_ \end{pmatrix}$$

$$4. \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} \bullet \begin{pmatrix} 1 & 2 \\ 4 & 1 \end{pmatrix} = \begin{pmatrix} \_ & \_ \\ \_ & \_ \end{pmatrix}$$

$$5. \begin{pmatrix} 3 \\ 2 \end{pmatrix} \bullet \begin{pmatrix} 1 & 3 \end{pmatrix} = \begin{pmatrix} \_ & \_ \\ \_ & \_ \end{pmatrix}$$

## Appendix B: Sample set of transparencies

A sample set of transparencies that will be used to instruct students during the study sessions. Each solution is clearly expanded into a step-by-step format, with each step highlighted to show where the answers were coming from. All answers consist of matrices with single digit entries only.

1.

$$(1 \ 2) \bullet \begin{pmatrix} 2 \\ 3 \end{pmatrix} = (1 \bullet 2 + 2 \bullet 3)$$

$$(1 \ 2) \bullet \begin{pmatrix} 2 \\ 3 \end{pmatrix} = (8)$$

2.

$$(2 \ 3) \bullet \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix} = (2 \bullet 2 + 3 \bullet 1 \quad \underline{\quad})$$

$$(2 \ 3) \bullet \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix} = (7 \quad 2 \bullet 3 + 3 \bullet 1)$$

$$(2 \ 3) \bullet \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix} = (7 \quad 9)$$

Appendix B: (Continued)

3.

$$\begin{pmatrix} 5 & 2 \\ 1 & 4 \end{pmatrix} \bullet \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \bullet 1 + 2 \bullet 2 \\ \_ \end{pmatrix}$$

$$\begin{pmatrix} 5 & 2 \\ 1 & 4 \end{pmatrix} \bullet \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 9 \\ 1 \bullet 1 + 4 \bullet 2 \end{pmatrix}$$

$$\begin{pmatrix} 5 & 2 \\ 1 & 4 \end{pmatrix} \bullet \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 9 \\ 9 \end{pmatrix}$$

4.

$$\begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} \bullet \begin{pmatrix} 1 & 2 \\ 4 & 1 \end{pmatrix} = \begin{pmatrix} 3 \bullet 1 + 1 \bullet 4 & \_ \\ \_ & \_ \end{pmatrix}$$

$$\begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} \bullet \begin{pmatrix} 1 & 2 \\ 4 & 1 \end{pmatrix} = \begin{pmatrix} 7 & 3 \bullet 2 + 1 \bullet 1 \\ \_ & \_ \end{pmatrix}$$

$$\begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} \bullet \begin{pmatrix} 1 & 2 \\ 4 & 1 \end{pmatrix} = \begin{pmatrix} 7 & 7 \\ 1 \bullet 1 + 2 \bullet 4 & \_ \end{pmatrix}$$

$$\begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} \bullet \begin{pmatrix} 1 & 2 \\ 4 & 1 \end{pmatrix} = \begin{pmatrix} 7 & 7 \\ 9 & 1 \bullet 2 + 2 \bullet 1 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} \bullet \begin{pmatrix} 1 & 2 \\ 4 & 1 \end{pmatrix} = \begin{pmatrix} 7 & 7 \\ 9 & 4 \end{pmatrix}$$

Appendix B: (Continued)

5.

$$\begin{pmatrix} 3 \\ 2 \end{pmatrix} \bullet (1 \ 3) = \begin{pmatrix} 3 \bullet 1 & \_ \\ \_ & \_ \end{pmatrix}$$

$$\begin{pmatrix} 3 \\ 2 \end{pmatrix} \bullet (1 \ 3) = \begin{pmatrix} 3 & 3 \bullet 3 \\ \_ & \_ \end{pmatrix}$$

$$\begin{pmatrix} 3 \\ 2 \end{pmatrix} \bullet (1 \ 3) = \begin{pmatrix} 3 & 9 \\ 2 \bullet 1 & \_ \end{pmatrix}$$

$$\begin{pmatrix} 3 \\ 2 \end{pmatrix} \bullet (1 \ 3) = \begin{pmatrix} 3 & 9 \\ 2 & 2 \bullet 3 \end{pmatrix}$$

$$\begin{pmatrix} 3 \\ 2 \end{pmatrix} \bullet (1 \ 3) = \begin{pmatrix} 3 & 9 \\ 2 & 6 \end{pmatrix}$$

## Figure Captions

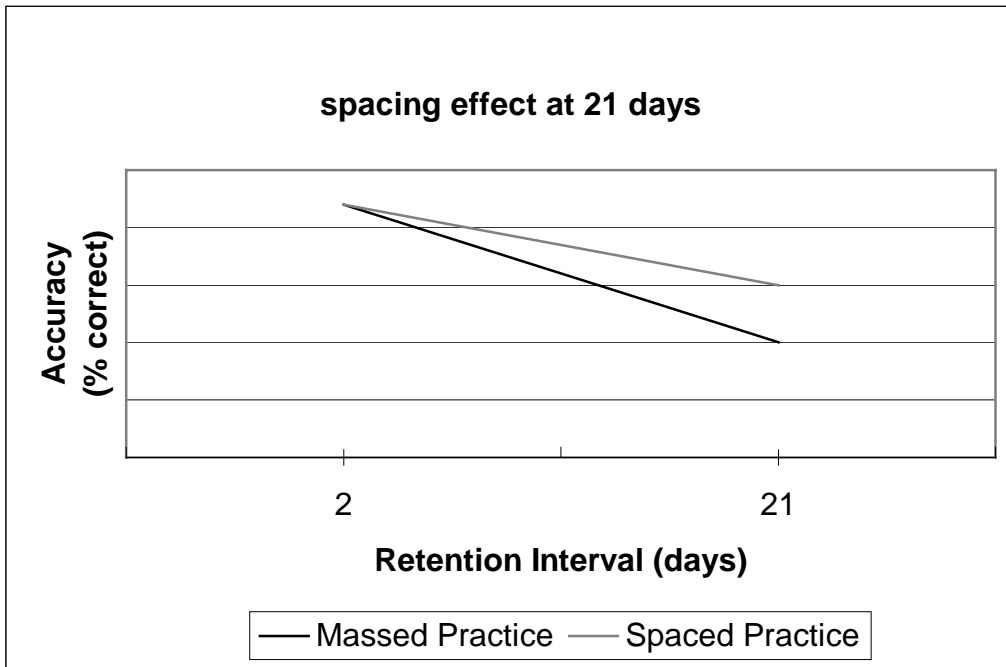
*Figure 1.* Illustrative Data. (A) Hypothetical results showing benefits of spacing at only the longer retention interval. (B) Hypothetical results showing a benefit of massing at the 2-day interval, and a benefit of spacing at the 21-day interval. The figure shows evidence of a crossover interaction of retention interval and type of practice. Error bars indicate  $\pm 1$  standard error.

*Figure 2.* Experiment 1. (A) Results of Experiment 1 including all participants. A benefit of massing is seen at both retention intervals. (B) Results of Experiment 1 including only those who scored at least 80% on study session 1. A benefit of massing is seen at 2-days and a benefit of spacing is seen at 21-days. Error bars indicate  $\pm 1$  standard error.

*Figure 3.* Experiment 2. (A) Results of Experiment 2 including all participants. A small benefit of high-massing is shown at both retention intervals. (B) Results of Experiment 2 including those participants who scored at least 50% on the study session. A small benefit of high-massing is shown at both retention intervals. Error bars indicate  $\pm 1$  standard error.

Illustrative Data

A



B

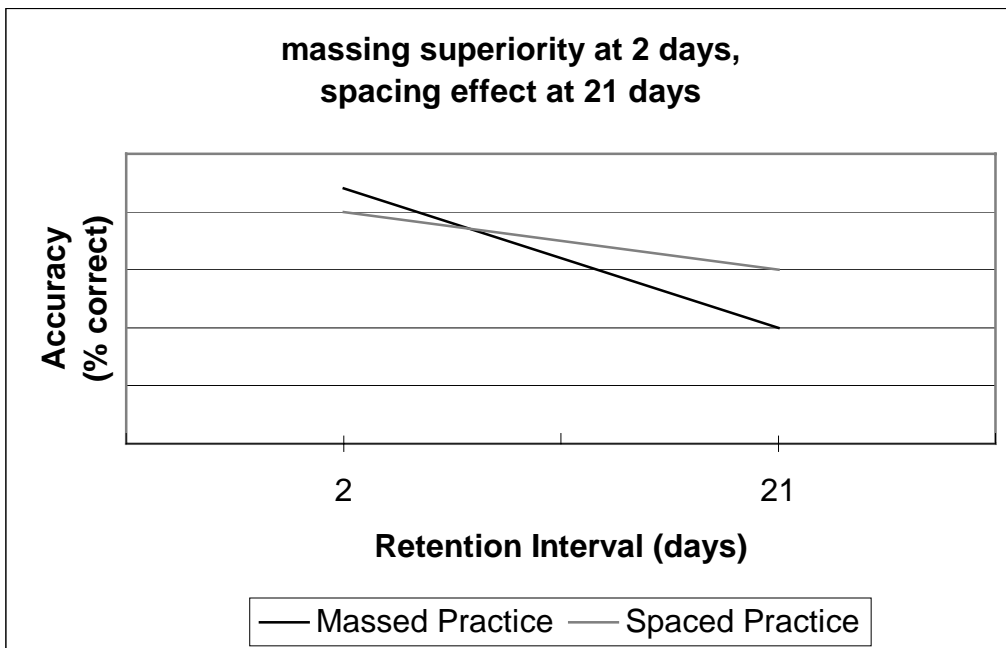
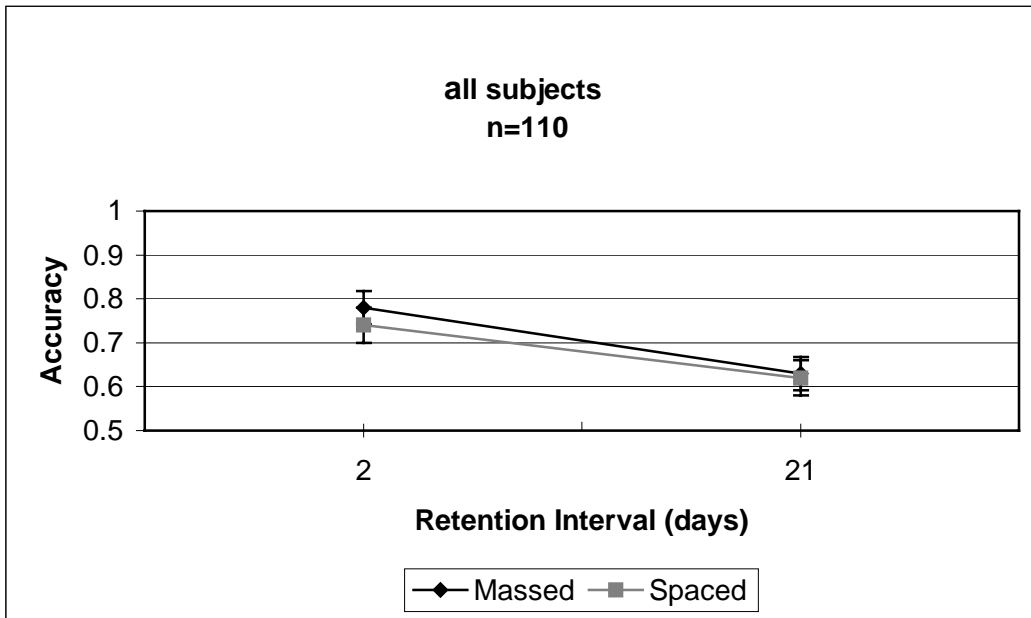


Figure 1

Experiment 1

A



B

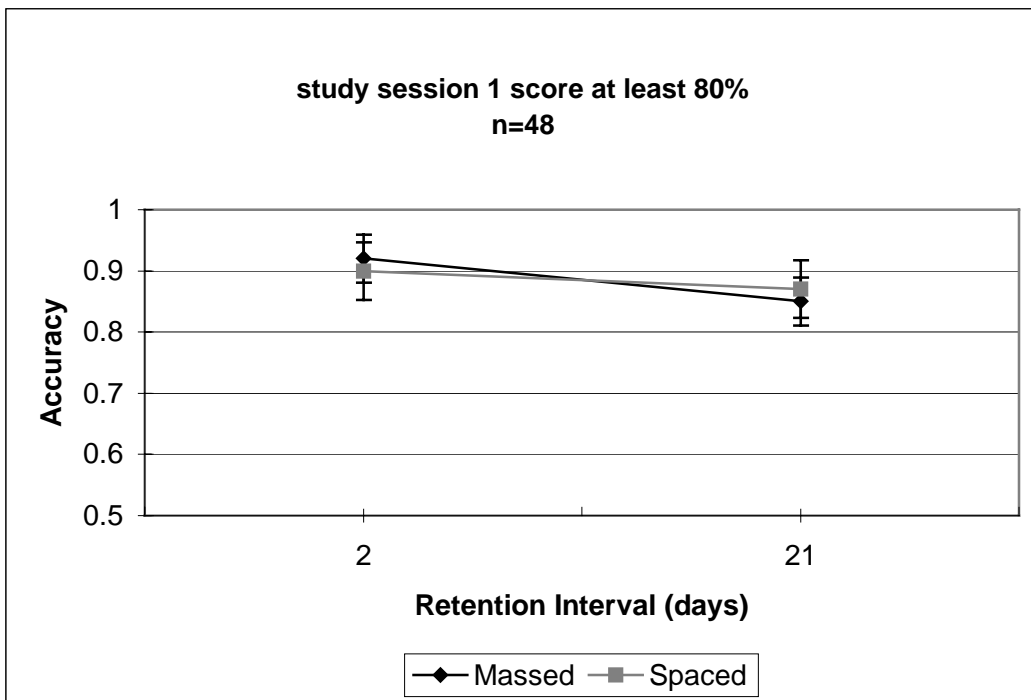
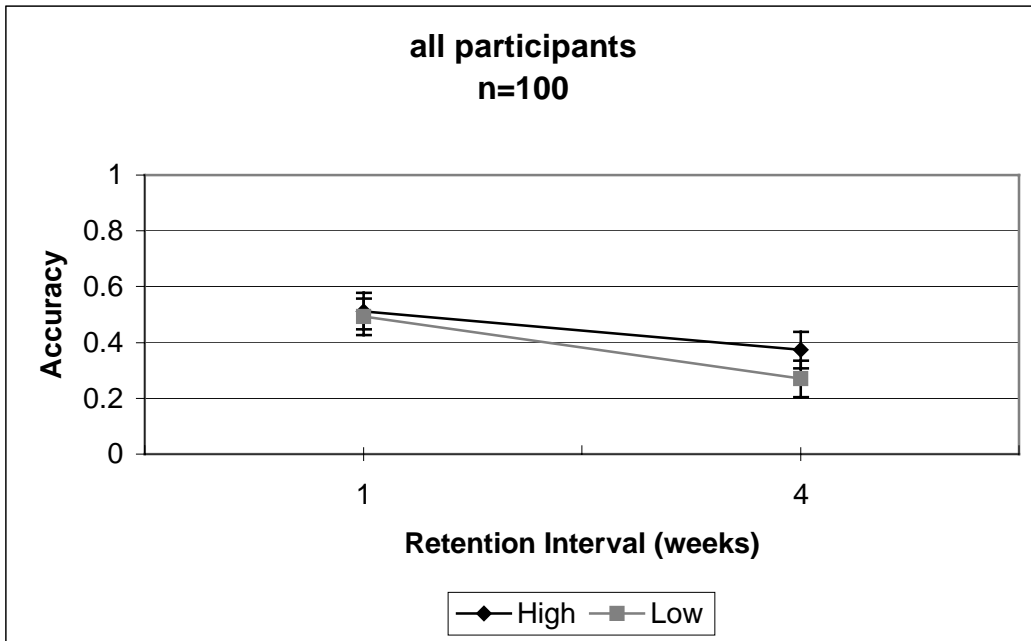


Figure 2

Experiment 2

A



B

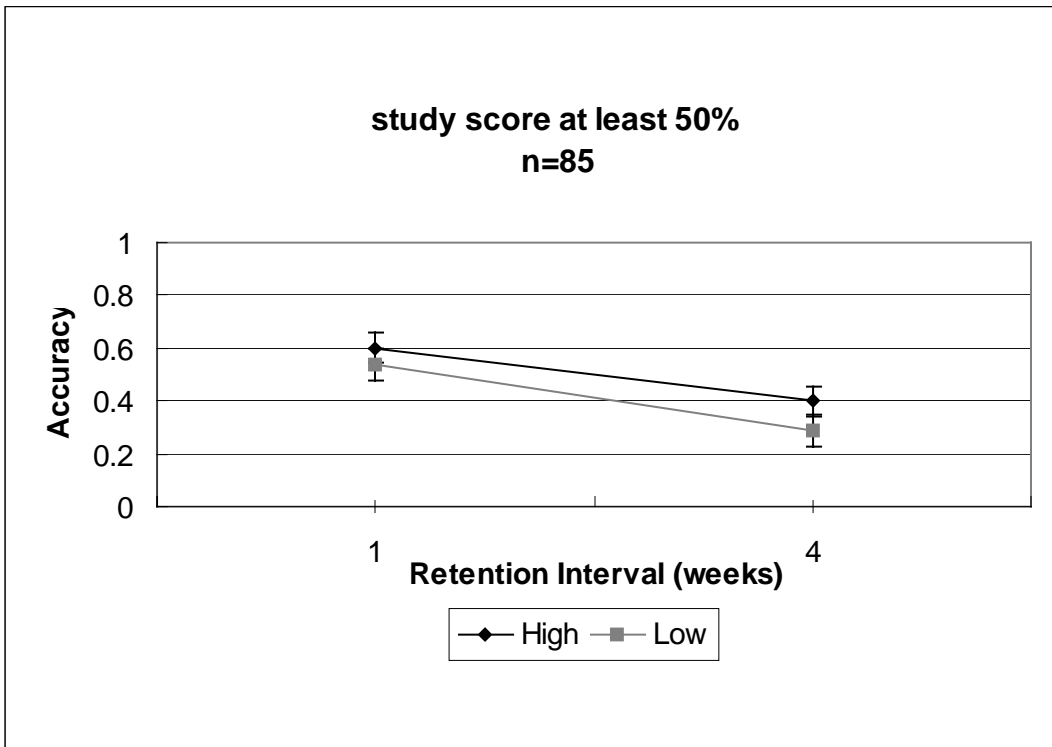


Figure 3