University of South Florida
Scholar Commons

Graduate Theses and Dissertations

Graduate School

4-8-2004

The Effects of Overlearning on Long-Term Retention

Kelli M. Taylor
University of South Florida

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

Part of the American Studies Commons

Scholar Commons Citation

Taylor, Kelli M., "The Effects of Overlearning on Long-Term Retention" (2004). *Graduate Theses and Dissertations.*
https://scholarcommons.usf.edu/etd/1266

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.

The Effects of Overlearning on Long-Term Retention

by

Kelli M. Taylor

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

Major Professor: Doug Rohrer, Ph.D.
Tammy Allen, Ph.D.
Kevin Thompson, Ph. D

Date of Approval:
April 8, 2004

Keywords: repetition, learning, total study time, degree of learning, study list length

© Copyright 2004, Kelli M. Taylor

Table of Contents

List of Figures

The Effects of Overlearning on Long-Term Retention

Kelli M. Taylor

ABSTRACT

Overlearning is defined as the continued study of an item immediately after it has been correctly recalled once. Based on past studies, the effectiveness of overlearning is uncertain. In the present study, the effectiveness of overlearning was tested in 3 experiments. In Experiment 1, participants studied 10 city-country pairs (e.g. doba, chad) to either a low or high degree of learning and were tested 1, 3, or 9 weeks later. In Experiments 2 and 3, participants studied varying numbers of word-definition pairs (e.g. vizard, mask) for a constant total study time. They were tested between 1 and 4 weeks later. While overlearning improved test performance, this difference in recall diminished at longer retention intervals. Furthermore, overlearning was found to be inefficient because the increase in recall was not proportional to the increase in study time. Finally, for a given amount of study time, the underlearning of more words led to higher absolute recall totals than did the overlearning of fewer words.

Introduction

Educators and trainers often use repetition as a teaching device. For instance, homework assignments often include many of the same type of math problem, repeated practice of spelling or vocabulary words, or repeated conjugations of a foreign language verb. Many of these tasks are an act of overlearning, which is defined as the continued study of material, in a single learning session, after it has been learned to a criterion of one correct recall. For example, suppose a student uses flashcards to study vocabulary words and removes a card once that item has been correctly recalled once. This is an example of *adequate learning* (AL), where the criterion is one correct recall. Any further study of the words, in the same study session, is known as *overlearning* (OL). Although overlearning helps students on a test the next day and trainees on a skills assessment shortly after training, it is not clear whether it is beneficial over longer retention intervals. Moreover, it is uncertain if the boost in recall due to underlearning is proportional to the amount of extra time required to overlearn. For example, twice the effort may not lead to twice the increase in recall.

Information on the usefulness of overlearning over long periods of retention could be very helpful when applied to classrooms and training courses. For example, if overlearning is a good learning strategy, then students should repeatedly study vocabulary words and solve many similar math problems in a single session. However, if overlearning is not particularly useful, then certain commonly used exercises could be replaced by more effective ones. For example, teachers should not assign vocabulary exercises that include the same 20 words every night for a week; instead, they should assign 20 different words each night for a total of 100 words that week. Although the assigning of more words will lead to the recall of a smaller proportion of the words, this "underlearning" of 100 words might lead to a greater absolute number of recalled words. Hence, studies on overlearning can evaluate the usefulness of such activities by examining the long-term retention of overlearned material. Furthermore, the application of the results of overlearning studies can lead to more efficient classrooms and training programs.

Benefits of Overlearning

There is an overwhelming body of research suggesting that overlearning is a valuable method of learning. That is, and perhaps not surprisingly, conditions that require overlearning produce greater recall than conditions that require lesser degrees of mastery. For example, Krueger (1929) found overlearning trumped adequate learning when participants learned words and returned for a test up to 28 days later. Postman (1961) found that participants required less time to relearn lists of words when the lists were originally overlearned. Schendel and Hagman (1982) found that overlearning helped to decrease the amount of time needed to retrain soldiers to assemble and disassemble a machine gun. Driskell, Willis, and Copper (1992) concluded from their meta-analysis that overlearning is an effective learning tool "for both physical and cognitive tasks" (p. 618).

After findings such as these, it is no wonder that overlearning is often described as a great learning technique. In fact, Fitts claimed that "the importance of continuing practice beyond the point in time where some…criterion is reached cannot be overemphasized" (1965, p.195) Likewise, Hall (1989) called overlearning practical because "continued practice on material already learned to a point of mastery can take place with a minimum effort, and yet will prevent significant losses in retention" (p. 328). Similarly, Foriska (1993) endorsed overlearning by presenting it as the mechanism children use to move information from short-term memory to long-term memory. With conclusions such as these, it is not surprising that researchers and educators advocate overlearning.

Limitations of Overlearning

Although research on overlearning does show that overlearning leads to better recall than lesser degrees of learning, these results are not surprising, as greater effort almost always produces greater performance. However, the apparent benefits of overlearning are less impressive when one considers the factors of retention interval and efficiency.

Retention Interval. While overlearning has often been shown to lead to better retention after short retention intervals, studies that employ a longer retention interval have revealed a lesser benefit. In fact, most of the studies that reveal benefits of overlearning have very short retention intervals. For instance, none of the studies in the meta-analysis by Driskell et al. (1992) used a retention interval that exceeded 28 days.

Similarly, Postman (1961) used a retention interval of only seven days, and Ausebel, Stager, and Gaite (1968) used a retention interval of only two days.

Likewise, studies that employ varying retention intervals show that the boost due to overlearning diminishes with time. That is, as the time between the study session and the test session grows longer, the benefits of overlearning decline, as illustrated by the hypothetical data in Figure 1A. For example, Craig, Sternthal, and Olshan (1972) found that, at several levels of overlearning, the boost in recall due to overlearning decreased as retention interval increased. Similarly, Reynolds and Glaser (1964) found that the overlearners' and adequate learners' curves converged, as shown in Figure 1A, as retention interval was lengthened. Specifically, they found that the differences between a 200% overlearning group and a 50% overlearning group existed at the 2-day interval on a recall task, yet these differences disappeared at the 21-day interval. Likewise, the meta-analysis done by Driskell et al. (1992) found the benefits of overlearning to weaken as retention interval lengthened. These findings indicate that while overlearning is beneficial over short retention intervals, this learning technique does not produce long-term retention.

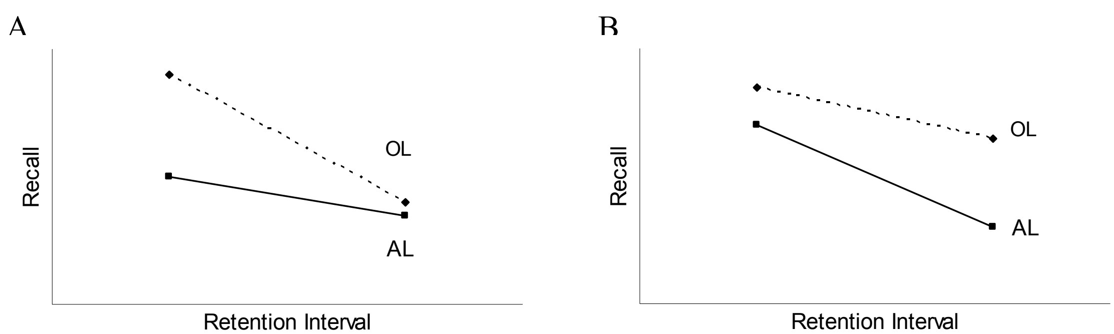


Figure 1. Hypothetical Results of an Overlearning Experiment . (A) Convergence of learning curves. (B) Divergence of Learning Curves.

One exception to this pattern of declining benefits was reported by Krueger (1929). His results showed a divergence of recall at longer retention intervals (Figure 1B). Specifically, he found that an increase in degree of learning from adequate learning to overlearning led to an increase in retention at both the 1- and 28-day retention intervals. Furthermore, at the 28-day interval, the ratio of words recalled by the overlearners to the ratio of words recalled by the adequate learners was actually larger than the same ratio at the 1-day retention interval. These conflicting results in the literature leave an unanswered question about the usefulness of overlearning at longer retention intervals.

Efficiency. Regardless of whether overlearning increases retention at certain retention intervals, it may not be the most efficient study method. That is, the increase in study time needed for overlearning may produce diminishing returns in recall benefits. To be precise, if a list of paired-associates is presented 10 (low) times, 15 (moderate) times, or 20 (high) times, recall total will naturally increase as the number of presentations increases, but the increase between the low and moderate groups will be greater than the increase between the moderate and high groups. One way of examining this inequity is by comparing varying degrees of overlearning. Using this method, Krueger (1929) found that, as degree of overlearning increased, retention also increased, but not proportionately to the increase in overlearning. In other words, when the number of trials was increased by 50%, retention increased by 48%. Yet, when the number of trials was increased by another 33⅓%, retention increased by only an additional 27%. This suggests that overlearning becomes less efficient as degree of overlearning increases. That is, the increase in study time needed to reach higher degrees of overlearning may produce a disproportionately small return in recall performance.

Similarly, Bromage and Mayer (1986) found that as the number of trials increased, recall total also increased at a negatively accelerated pace. That is, as degree of overlearning increased, the rate of increase in retention slowed. Similarly, Driskell et al. (1992) found that participants did retained more as the degree of overlearning increased, but once again, this increase in study time led to diminishing returns. Likewise, Kratochwill, Demuth, and Conzemius (1977) found that increasing the number of study trials by 20 led to an increase in retention of only one word, or only 25%. Once again, the proportional increase in retention was much less than the increase in study time.

Findings such as these suggest that the benefits of overlearning do not increase at the same rate that degree of overlearning increases, thus making it an inefficient learning strategy. That is, the proportional increase in retention due to overlearning is usually not as large as the proportional increase in study time.

Therefore, although much of the overlearning literature describes overlearning as a beneficial learning technique, a closer examination of the data reveals that the benefits of overlearning decline at longer retention intervals. In addition, there is much evidence suggesting that the benefits due to overlearning are not commensurate with the increase in total study time.

The Total Study Time Method

Another source of literature related to overlearning that is not often cited in overlearning studies are studies employing what is known as the total time paradigm. Unlike the overlearning paradigm, where total study time is varied and study list length is held constant, in the total time paradigm, total study time is held constant and study list length is varied. For example, in a total- time study, participants might be given five minutes to study a list of 5, 10, or 20 items. Of course studying 5 items will lead to the recall of a higher proportion of items, but it is unclear which condition will lead to a higher *absolute recall total* (i.e. the total number of words recalled). The implications are described in greater detail in the general discussion.

Murdock (1960) used this design and hypothesized that there is no effect of varying list length on absolute recall total. Hence, this hypothesis suggests that, when given a total time of 10 minutes to study 10 word pairs, the same number of words will be recalled as when given 10 minutes to study 20 word pairs. Yet Roberts (1972) reported results contrary to this hypothesis. He instead found the absolute recall total increased as list length increased when total time was held constant. For example, for a given 80-s study period, participants who studied 40 words recalled more words absolutely than participants who studied 20 words. Here studying more words in an allotted time led to an increase in absolute recall even though the proportion of words recalled declined. That is, the overlearning of 20 items produced a smaller absolute recall total than the underlearning (UL) of 40 items. From these results, it seems that studying more words in an allotted time can lead to a higher number of words recalled, and that overlearning is not the best learning strategy when absolute recall rather than proportion of recall is important. For example, when preparing for the verbal section of the GRE, there are thousands of vocabulary words that may appear on the test. Thus, it is best to know as many words as possible to increase the chance that you will know some of the words that do appear. In this instance, then, absolute recall is more important than proportional recall.

However, Roberts (1972) found that this increase in absolute recall due to increased study list length, when total study time was held constant, asymptotically approached a ceiling. Specifically, he found there was no significant difference between the absolute number of words recalled after studying 20 or 40 words in a 40-s study period. Thus, for a given total study time, extreme underlearning (EUL) may provide either no further boost in absolute recall total or even a decline. That is, as shown in Figure 2, does the level of absolute recall continue to increase as study list length increases, or will absolute recall begin to decline as study list length increases to extreme levels?

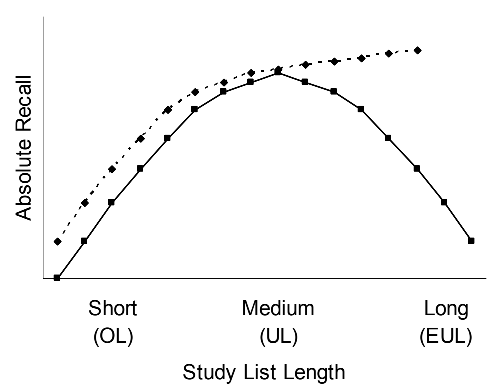


Figure 2. Hypothetical recall performance as a function of list length (constant total study time). As the degree of underlearning increases to extreme levels, the increase in absolute recall total could either reach asymptote, or begin to decline.

Although Roberts (1972) found that underlearning many words can produce greater absolute recall than overlearning fewer words, there are two characteristics of this study that hinder generalizability and ecological validity. The first is his use of a free recall paradigm, where participants are asked to recall a list of studied words. Situations in which free recall are needed do not normally present themselves in the real world. On the other hand, tasks that involve cued recall occur more often in real world settings. For example, if you are visiting a Spanish speaking country, you might need to know that the Spanish word for food is “comida”. The second limitation of Roberts’ study is his use of an immediate retention interval. Because it is usually important to retain information for more than a minute, of course it is important to test this effect at longer, more meaningful retention intervals.

Overview of the Present Studies

In the present studies, we pitted overlearning against underlearning in three different experiments to explore the benefits and limitations of overlearning. In Experiment 1, the utility of overlearning was assessed with a design normally used in the overlearning literature as described further above. Specifically, every participant studied 10 items, and total study time was varied. In Experiment 2, we assessed the utility of overlearning by varying list length rather than total study time, as in the total time studies. That is, participants studied lists of either 10 or 20 words for a period of 12 minutes. In Experiment 3, we once again held total study time constant and varied list length to assess whether any increase in absolute recall achieved by an increase in number of study words ultimately declines at longer list lengths. In all three experiments, participants returned for a test between one and four weeks later.

Experiment 1

This experiment tested the benefits of overlearning over long retention intervals using a paper/pencil task. College students either overlearned or underlearned 10 city-country pairs for either 5 or 20 minutes and were tested one, three, or nine weeks later.

Method

Participants. A sample of 130 undergraduate students participated in return for extra credit. Eight other students began the study but failed to complete it.

Design. The two variables, learning level (Lo or Hi) and retention interval (1, 3, or 9 weeks) were manipulated between subjects. Participants were randomly assigned to conditions.

Procedure. The experiment consisted of both a study session and a test session. Booklets were used during the study session. The first page of the booklets included a list of 10 city-country pairs, as listed in Appendix A. At the beginning of the study session, participants were given 60 s to study these pairs. Following this initial study trial were 5 (Lo-Learners) or 20 (Hi-Learners) feedback trials, each of a 1-min duration. Each trial corresponded to one page in the booklet. On these pages, the cities were listed on the left, and participants were asked to recall the corresponding country in the blank space provided to the right. Handwriting time was reduced by including countries with only five or fewer letters. After 50 s, participants were asked to unfold the page, which revealed the correct answers immediately to the right of the participants' answers.

Participants were then instructed to study the city-country pairs for the remaining 10 s of the 1-min trial. After the participants had completed five of these trials, the Lo-Learners left the room. After a brief rest, the Hi-Learners completed their remaining 15 trials. The words were randomly ordered each time they appeared so that participants would not merely memorize each definition's position within the list.

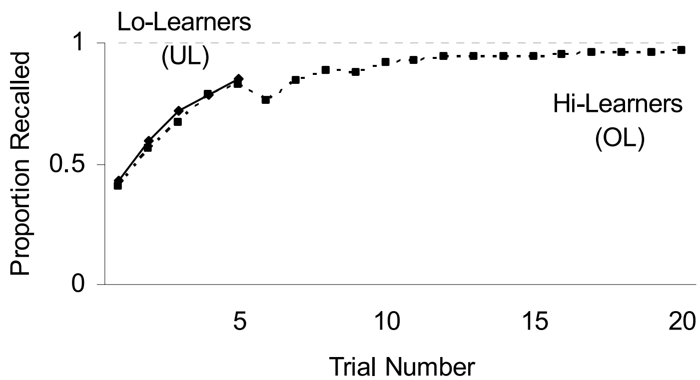
The participants were tested one, three, or nine weeks later. During the test, participants were given the name of the city and were asked to recall the country. They had three minutes to complete this task.

Results and Discussion

The results of the study phase are shown in Figure 3A. As shown, virtually all the Hi-Learners (overlearners) correctly recalled all 10 countries on each of the last 10 trials. Because they completed twice as many trials as that needed to reach the criterion of one correct recall, these Hi-Learners clearly relied on overlearning, as defined in the introduction. In contrast, on average, the Lo-Learners (underlearners) had not reached a criterion of even one correct recall of all 10 countries by the end of their study session. Thus, many of the Lo-Learners failed to reach even adequate learning.

The test results are shown in Figure 3B. As expected, the Hi-Learners recalled more than the Lo-Learners, as indicated by a significant main effect of degree of learning on recall, $F(1,124) = 33.29$, $p < .01$. However, the difference between the Hi-Learners and the Lo-Learners decreased dramatically with retention interval, as evidenced by the significant interaction between retention interval and degree of learning, $F(2,124) = 8.65$, $p < .01$. More specifically, the results of a Tukey HSD test revealed that the difference in recall between the Hi-Learners and Lo-Learners was significant at the 1- and 3- week intervals but failed to reach significance at the 9-week interval.

A



B

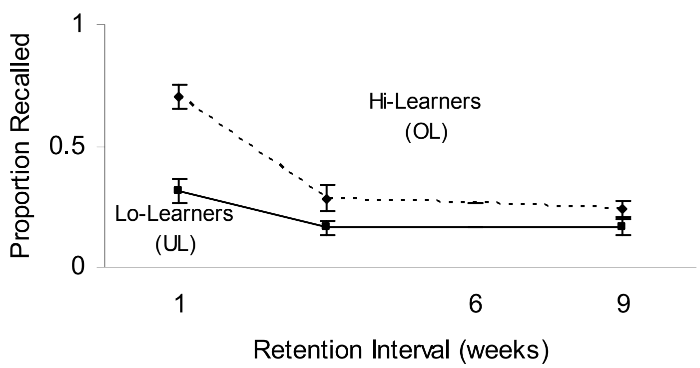


Figure 3. Results of Experiment 1. (A) Study Phase. The Hi-Learners likely overlearned the material, while the Lo-Learners failed to reach even adequate learning. (B) Test Results. Overlearning boosted recall, but the difference between the Hi- and Lo-Learners declined at longer retention intervals. Error bars indicate ±1 standard error.

These results suggest that while overlearning does boost recall, the advantages are limited. One of these limitations is the reduction of the benefits of overlearning as retention interval increases. For example, although overlearning led to a doubling of recall at the 1-week retention interval, this increase in recall decreased dramatically as retention interval increased. Furthermore, this doubling of recall becomes less impressive when you consider that total study time quadrupled. Thus, even at the point of the largest increase in recall, the benefits of overlearning did not come close to equaling the increase in study time. Consequently, overlearning is not an efficient study method.

When interpreting these results, however, one caveat should be considered. In this experiment, participants did not learn to criterion. That is, the degree of overlearning was not strictly controlled for. Although we did not specifically require participants to reach a criterion of adequate learning and overlearning, the difference between the Lo- and Hi-Learners was arguably much greater than this. Consequently, any difference found between the Hi and Lo-Learners exaggerates the benefits of overlearning because the Lo-Learners failed to reach adequate learning.

Experiment 2

This experiment examined whether it is better to overlearn a small number of items or to underlearn a larger number of items. That is, when absolute recall is more important than proportion of recall, is it beneficial to spend your time studying more or less information? For example, as described in the introduction, should a teacher assign activities geared toward learning vocabulary words for the same 20 words every night for a week, or should she assign activities for 20 different words each night, resulting in the studying of 100 words a week?

To answer this question, college students devoted 12 minutes to learning a list of either 10 or 20 word-definition pairs. Thus, those participants who studied only 10 words in the 12-min study period were likely to overlearn the words. By contrast, those participants who studied 20 words in the 12-min study period were likely to underlearn the words. One or four weeks later, all participants returned for a test.

Method

Participants. The sample contained 88 University of South Florida undergraduates who participated in return for extra credit. One hundred seven participants began the study, but 13 failed to return for the test, and four did not complete the task correctly. None participated in Experiment 1.

Design. Study list length (10 or 20) and retention interval (one or four weeks) were manipulated between subjects. Participants were randomly assigned to one of the four conditions.

Procedure. The procedure used was similar to that used in Experiment 1, except that we varied the study list length rather than the total study time. Once again, participants attended both a study session and a test session. In the study session, each participant was given a booklet and asked to study a list of 10 or 20 word-definition pairs, as listed in Appendix B. Each definition was a single word with four letters or fewer (e.g., vizard-mask) in order to minimize writing time. Participants were given 120 s to study the word pairs. Following this 2-min study period, participants completed twenty 30-s feedback trials in which they tried to recall the definitions of the words. During each feedback trial, only five words were presented, with all words appearing equally often. Thus, the 10-word Learners saw each word 10 times, whereas the 20-word Learners saw each word only five times. After 20 s of the recall phase of the feedback trials, participants unfolded the page, and the correct definitions appeared immediately to the right of their responses. Participants then studied the pairs for the remaining 10 s before turning the page to begin the next feedback trial. The words were randomly ordered each time so that participants would not merely memorize each definition's position within the list.

The participants were tested either one or four weeks later. They were given the words and asked to recall the corresponding one-word definitions. They were given four minutes to complete the test.

Results and Discussion

The study data are shown in Figure 4A. As shown, most of the 10-word Learners (overlearners) were correctly recalling all ten definitions on each of the last five study trials. Because they underwent twice as many trials as that needed to reach a criterion of one correct recall, it is assumed they underwent 100% overlearning. In contrast, by the last study trial, the 20-word Learners (underlearners) were recalling only 70% of the definitions correctly. Thus, these 20-word Learners failed to reach even adequate learning. Consequently, any difference found between the 10- and 20-word Learners exaggerates the benefits of overlearning because the 20-word learners failed to reach adequate learning.

The proportional test results are illustrated in Figure 4B. Not surprisingly, the 10-word Learners, who studied each word twice as long as the 20-word Learners, recalled a greater proportion of words than the 20-word Learners, $F(1,84) = 12.21$, $p < .01$. However, as in Experiment 1, the benefits of overlearning decreased dramatically with retention interval, as indicated by the significant interaction between retention interval and list length, $F(1,84) = 6.15$, $p < .05$.

Interestingly, as shown in Figure 4C, when the absolute recall data were analyzed, we found that the 20-word Learners (underlearners) actually recalled more words than the 10-word Learners (overlearners), $F(1,84) = 4.14$, $p < .05$. Furthermore, as retention interval increased, the benefits of underlearning on absolute recall total stayed constant, as demonstrated by the statistical nonsignificance of the list length x retention interval interaction ($F<1$).

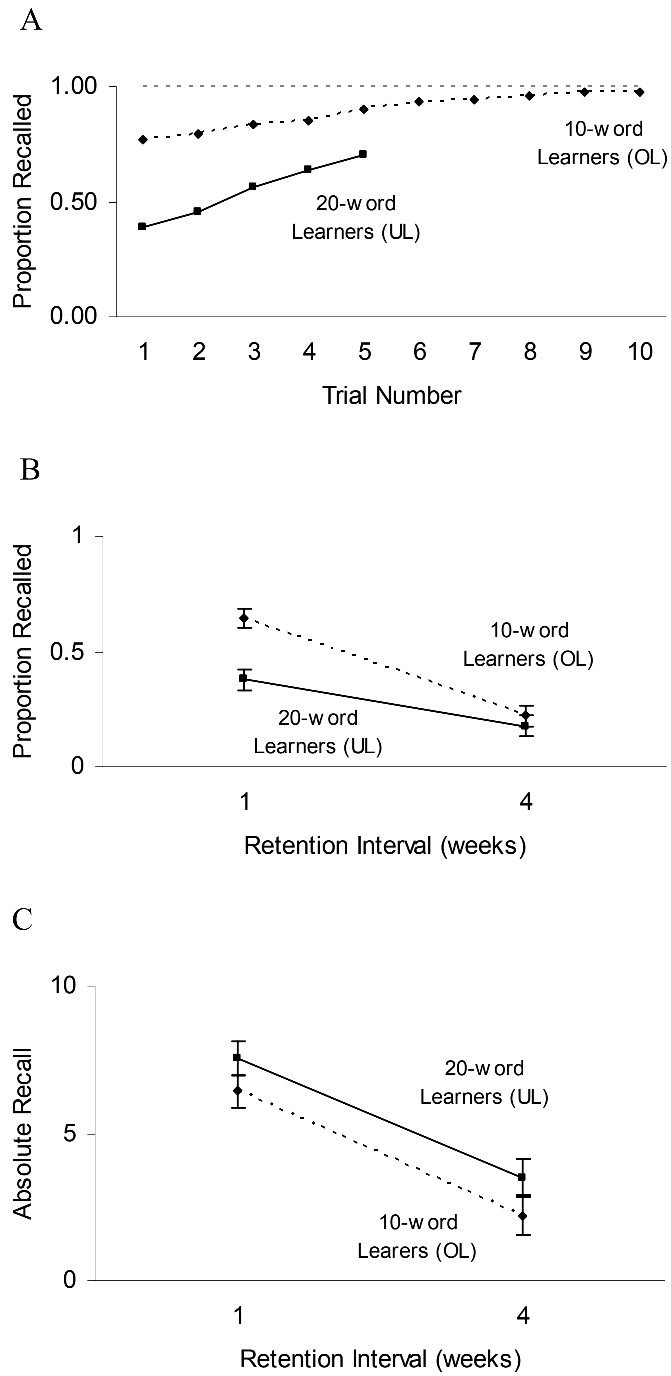


Figure 4. Results of Experiment 2. (A) Study Phase. (B) Proportional Test Results. (C) Absolute Test Results. Error bars indicate ±1 standard error.

As in Experiment 1, overlearning led to a higher proportion of words recalled than underlearning, but this difference decreased as retention interval increased. Once again, it seems that overlearning is not useful when material must be maintained over long retention intervals. Furthermore, overlearning appears to be inefficient. As shown in Figure 4C, the overlearners spent twice as much time studying each word, and still recalled fewer words absolutely than the underlearners at both one and four weeks. This finding suggests that, when given a set amount of time to study, instead of studying a small amount of material, you should study a larger amount of material. In return, you will be able to recall more information absolutely.

Experiment 3

The results of Experiment 2 suggest that underlearning more information produces greater absolute recall totals than overlearning less information, but does this finding extend to extreme underlearning? The purpose of Experiment 3 was to find such a limit, if one exists. As in Experiment 2, total study time was held constant. Study list length was manipulated, with participants studying list lengths of 5 (EUL), 10 (UL), and 20 (OL).

Because the extreme underlearning condition required very rapid presentation rates, we were required to use a computer. Using a computer based task also allowed us to precisely control for total time spent on each word.

Method

Participants. Participants were 32 undergraduates from the University of South Florida who participated in return for extra credit. Twelve additional students began the study but failed to complete the test phase. None participated in Experiments 1 or 2.

Design. List length (5,10, and 20 words) was a within-subjects variable.

Procedure. The procedure was similar to that of Experiment 2. Participants attended both a study session and a test session, and the study session was controlled by a computer. During the study session, each participant completed a 5-word practice list before completing 5-, 10-, and 20-word experimental lists, as listed in Appendix C. For each list, participants were given a 1-s prompt to get ready to study and then shown each word-definition pair (e.g. vizard-mask) for three seconds. After this initial study period, the participant completed eight (5-word Learners), four (10-word Learners), or two (20-word Learners) feedback trials depending on the list length. Before each feedback trial began, a 1-s “Get ready to remember” prompt appeared. Then each word from the list appeared individually and the participant had five seconds to type in the correct definition. Then, if the participant answered correctly, the word “correct” appeared on the screen. If the participant answered incorrectly, the word “incorrect” appeared along with the correct answer. This feedback remained for three seconds in either instance. After every word on the list appearred, the participants rested for five seconds before beginning the next feedback trial. The words were randomly ordered each time so that participants could not merely memorize each definition’s position within the list. Furthermore, the order of the conditions was also random to prevent confounds such as fatigue and interference.

One week later the participants were tested by paper and pencil. They were given the words and asked to recall the corresponding one-word definitions. They had five minutes to complete the task.

Results and Discussion

The results of the study phase are shown in Figure 5A. It appears that by the fifth trial, in the 5-word condition, most participants were recalling all five words correctly. Thus, it is assumed that by the eighth trial, these participants had experienced 100% overlearning. In contrast, in both the 10- and 20- word conditions, participants on average failed to reach adequate learning.

As shown in Figure 5C, we were successful in finding an upper limit to the finding of Experiment 2. As in Experiment 2, as list length increased, absolute recall also increased, $F(2,62) = 8.39$, $p < .01$. However, recall did not increase monotonically as list length increased. Post hoc tests attribute the main effect of degree of learning to the differences between the overlearners and the two underlearning conditions. There was however, no significant difference between the 10- and the 20-word conditions. Thus suggesting that the increase in study list length from 10 to 20 words did not result in an increase in recall.

These results imply that while it is better to spend your time underlearning more information, there is a limit at which studying more information in an allotted time will no longer increase recall. Once again, this experiment did not use a learning to criterion design, so the results should be interpreted cautiously.

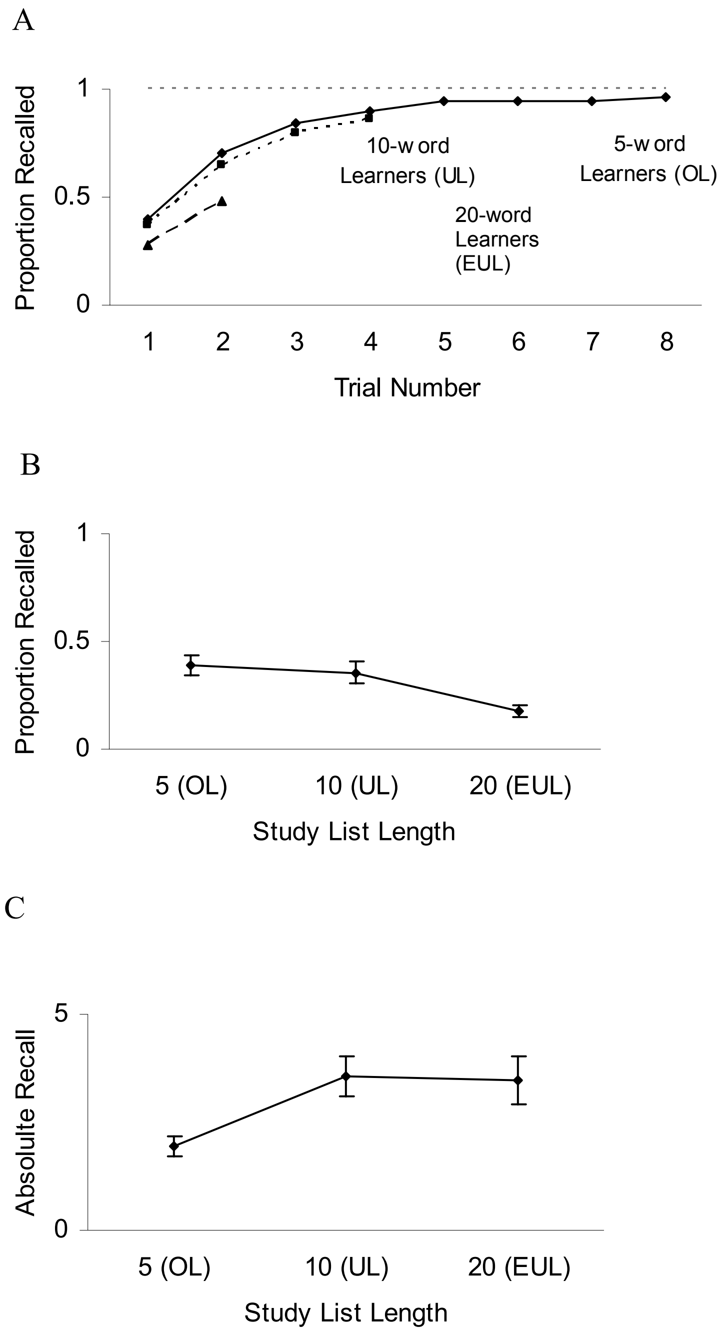


Figure 5. Results of Experiment 3. (A) Study Phase. The 5-word Learners (overlearners) likely overlearned the words, whereas the 10-word Learners (underlearners) and the 20-word Learners (extreme underlearners) failed to reach even adequate learning. (B) Proportional Test Results. The 5- and 10-word Learners both recalled a higher proportion of words than the 20-word Learners. (C) Absolute Test Results. The 10-and 20-word Learners both recalled a higher absolute total of words than the 5-word learners. Error bars indicate ±1 standard error.

General Discussion

Summary

Although overlearning resulted in a boost in proportion of words recalled in all three experiments, closer inspection of the results revealed that there were limitations to these benefits. More specifically, the benefits of overlearning decreased at longer retention intervals, and overlearning also proved to be an inefficient study method.

This observed decline in the benefits of overlearning at longer retention intervals is in agreement with past studies (Craig, Sternthal, & Olshan, 1972; Reynolds & Glaser, 1964). That is, while overlearners recalled a much higher proportion of words than the underlearners one week after the study session, this difference declined with retention interval (i.e. the time elapsed between the study session and the test session), as shown in Figures 4B and 5B. This result was consistent when tested using two different designs. In Experiment 1, overlearning was shown to be less effective at longer retention intervals when study list length was held constant and total study time was varied. Likewise, Experiment 2 revealed a rapidly diminishing benefit of overlearning when study list length was varied, and total study time was held constant. Thus, we found that overlearning proved less useful at long retention intervals with two different experimental designs. This convergence suggests that overlearning is not a good learning strategy when long-term retention is important.

In addition, overlearning also seems to be a highly inefficient study method, as suggested by past studies (Driskell et al., 1992; Kratochwill, Demuth, & Conzemius, 1977). In fact, in Experiments 1, 2 and 3, the percentage increase in recall gained from overlearning the material was less than the percentage increase in total study time. That is, the boost in recall due to overlearning was not proportional to the increase in total study time. In Experiment 1, for example, the Hi-Learners studied four times longer than the Lo-Learners, and at best, outperformed them by only double (Figure 3B). Similarly, the overlearners (10-word Learners) in Experiment 2 spent 100% more time studying each word than did the underlearners (20-word Learners), as shown in Figure 4B. In return for this effort, the overlearners recalled only 70% more than the underlearners at the 1-week interval and only 26% more at the 4-week interval. Likewise, in Experiment 3, the overlearners (5-word Learners) studied each word twice as long as underlearners (10-word Learners) and still recalled only 8% more (Figure 5B). This is evidence that the time committed to overlearning is nowhere near proportional to the benefits achieved.

The inefficiency of overlearning becomes even more apparent when the absolute recall totals of Experiments 2 and 3 are examined. In Experiment 2, although twice as much time was spent learning each word in the 10-word list, participants recalled a higher number of words from the 20-word list (Figure 4C). Furthermore, this increase in absolute recall total (i.e. the total number of words recalled) remained virtually consistent across both retention intervals. Similarly, in Experiment 3, when less time was devoted to learning each word in both the underlearning (10-word) and the extreme underlearning (20- word) conditions, participants nevertheless recalled more than when the words were overlearned (5-word condition) (Figure 5C). These results suggest that underlearning 10 or 20 words will probably result in greater absolute recall total than overlearning 5 words in the same amount of time.

However, underlearning does not always produce greater absolute recall totals than overlearning. That is, underlearning can be too extreme to result in any recall benefits. For example in Experiment 3, increasing the list length from 10 to 20 while holding the study time constant did not produce an increase in absolute recall (Figure 5C). This suggests that the participants did not have enough time to study each word in the 20-word (extreme overlearning) condition. As a result, they did not benefit from the increase in list length. These results are evidence, that in extreme measures, underlearning can be equally as inefficient as overlearning.

In summary, while overlearning did boost recall on a test a week later, this boost in recall disappeared as retention interval lengthened. Therefore, it may not be wise to spend time overlearning material if it needs to be recalled in the distant future. Furthermore, while overlearning led to increased recall when compared to underlearning, this increase in recall was never proportional to the increase in total study time. Thus, the time invested to overlearn was never proportional to the increase in recall. Finally, when more words were studied in a given amount of time, a higher number of words were recalled at test. Even more striking, this benefit did not decrease with retention interval as did the benefits of overlearning. So, regardless of retention interval, underlearning more words led to greater absolute recall.

Implications

A lot of research has been done that could help improve certain facets of education. Unfortunately, this research is not often used to evaluate and update educational procedures. The results from the current studies have practical implications for education and learning in general.

One problem often described by educators is that students do not retain information past test day. Cooper et al. (1996) expressed this concern of teachers by relaying that students forget a large amount of material during summer breaks. The finding that overlearning is only advantageous at short retention intervals suggests that, if the material being taught is something that should be remembered over long periods of time, as is most information taught in schools, perhaps overlearning is not the best method to use. For example, the purpose of teaching vocabulary is to give students a large, ever growing pool of terminology. Thus, rather than spending 30 minutes to overlearn a relatively short list of words, the present data suggest it might be better to underlearn a longer list. Thus, as in Experiments 2, this should lead to the retention of a larger number of words, regardless of retention interval.

Another situation in which absolute total recall is important occurs when studying for the GRE. On the GRE, any of thousands of words could be presented to you. Thus, the total number of words recalled is important because, as the total number of words you know increases, your chances to do well also increase. So, the findings of Experiments 2 and 3, suggest that underlearning may produce higher absolute recall than overlearning.

Overlearning techniques are also often used in math courses. Teachers may assign students 20 or more of the same type of math problem for homework in one night. Once students figure out how to complete the first problem correctly, every subsequent problem constitutes overlearning. As evidenced by our results in Experiment 1, while this overlearning may boost their scores in the immediate test, it may not produce long-term learning. Furthermore, as evidenced by the results of Experiments 1 and 2, the extra time students spend to overlearn is not likely to produce a proportional gain in recall.

There are, of course, instances in which overlearning would be the best strategy. For instance, if the proportion of information recalled were important, overlearning strategies would be the most useful. For example, it is important when administering CPR that you remember and complete *all* the steps. In this case, absolute recall is less relevant, because recalling every detail is crucial. Thus, for instances such as these, overlearning would be helpful, but as demonstrated in Experiments 1 and 2, the material would have to be retrained periodically to reduce forgetting.

In conclusion, before deciding on a study method, be sure to consider the desired result. If proportion of recall is desired, and the retention interval is short, overlearning could be useful. However, if your goal is to retain as much information as possible regardless of retention interval, underlearning a lot of information would be optimal.

Future Studies

When planning future studies of the benefits of overlearning, a few considerations should be made. Unlike the present studies, many learning experiments require participants to learn words to a criterion. That is, in the study session, the feedback trials continue until the participant has correctly recalled every word the required number of times. For example, in another study we recently completed, participants were required to learn word pairs to a criterion of either 1 or 9 correct recalls before moving on. This process can be likened to the procedure often used when studying with flash cards. Thus, in the condition in which participants are learning to a criterion of one, if they correctly recall the definition of a word, that word is removed from the list. They continue this process until no words are left in the list. The use of this design will control for degree of learning. It will also allow us to test the advantages and disadvantages of overlearning at precise degrees.

In each of these experiments, the sizes of the study list lengths were limited due to time constraints. In fact, the longest study list was 20 words. Study books for tests such as the SAT or GRE contain vocabulary lists thousands of words long. Thus, in future studies the parameters of list length and total study time should be expanded. While there should not be qualitative differences in the results as compared to the results of our experiments, the results would more closely approximate the study process for such exams.

The gender composition of the sample should also be considered in future experiments. In the present studies, many more females than males were tested. Thus, although we found no gender differences, this highly unbalanced ratio of females to males resulted in very low power for these tests. Even though gender differences usually do not exist in memory tasks, future studies should use more balanced samples.

Finally, although it is not likely that school-aged children differ qualitatively in learning styles, these studies should be replicated with a sample of young children before we generalize these results to that population. Results from school-aged populations will help answer questions about how children learn and retain information. They will also be useful in increasing classroom effectiveness.

References

Ausubel, D. P., Stager, M., & Gaite, J. H. (1968). Retroactive facilitation in meaningful verbal learning. *Journal of Educational Psychology*, *59*, 250-255.

Bromage, B. K., & Mayer, R. E. (1986). Quantitative and qualitative effects of repetition on learning from technical texts. *Journal of Education Psychology, 78*, 271-278.

Cooper, H., Nye, B., Charlton, K., & Lindsay, J. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research*, *66*, 227-268.

Craig, C. S., Sternthal, B., & Olshan, K. (1972). The effect of overlearning on retention. *Journal of General Psychology, 87*, 85-94.

Driskell, J.E., Willis, R. P., & Copper, C. (1992). Effect of overlearning on retention. *Journal of Applied Psychology, 77*, 615-622.

Fitts, P. M. (1965). Factors in complex skill training. In R. Glaser (Ed.), *Training Research and Education* (pp. 177-197). New York: Wiley.

Foriska, T. J. (1993). What every educator should know about learning. *Schools in the Middle, 3*, 39-44.

Hall, F. H. (1989). Retention and forgetting: The functional approach continued. In *Learning and Memory* (2nd ed., pp. 317-345). Boston, MA: Allyn & Bacon.

Kratochwill, T. R., Demuth, D. M., & Conzemius, W. C. (1977). The effects of overlearning on preschool children's retention of sight vocabulary words. *Reading Improvement, 14*, 223-228.

Krueger, W. C. F. (1929). The effect of overlearning on retention. *Journal of Experimental Psychology, 12*, 71-78.

Murdock, B. B. (1960). The immediate retention of unrelated words. *Journal of Experimental Psychology, 60,* 222-234.

Postman, L. (1962). Retention as a function of degree of overlearning. *Science, 135*, 666-667.

Reynolds, J. H., & Glaser, R. (1964). Effects of repetition and spaced review upon retention of a complex learning task. *Journal of Educational Psychology, 55*, 297-308.

Roberts, W. A. (1972). Free recall of word lists varying in length and rate of presentation: A test of total-time hypotheses. *Journal of Experimental Psychology, 92*, 365-372.

Schendel, J. D. & Hagman, J. D. (1982). On sustaining procedural skills over a prolonged retention interval. *Journal of Applied Psychology, 67*, 605-610.

Appendices

Appendix A

List of city-country pairs for Experiment 1.

City	Country
Axim	Ghana
Bari	Italy
Chiba	Japan
Doba	Chad
Lugo	Spain
Mago	Tonga
Pune	India
Rabat	Malta
Talara	Peru
Yumen	China

Appendix B

List of vocabulary words for Experiment 2.

Word	Definition	Word	Definition
anta	pier	acrogen	fern
cerate	wax	beldam	hag
cicatrix	scar	cess	tax
excrescence	wart	elver	eel
fosse	moat	emmet	ant
peruke	wig	mentum	chin
stannum	tin	mome	fool
tippet	cape	ruga	fold
vizard	mask	salver	tray
weir	dam	talar	robe

Note: The overlearners studied the 10 words in the left column. The underlearners studied all 20 words.

Appendix C

List of additional vocabulary words for Experiment 3.

Word	Definition	Word	Definition
acarus	mite	pneuma	soul
brad	nail	portent	omen
bullock	ox	roughleg	hawk
cabal	plot	swagman	hobo
concordat	pact	victual	food
kismet	fate	vitellus	yolk
occident	west	wheal	welt
phyllome	leaf		

Note: These words were used in addition to the words used in Experiment 2.