# UC Merced

## Proceedings of the Annual Meeting of the Cognitive Science Society

### Title
Expanding Retrieval Promotes Long Term Retention by Preventing Rapid Rates of Forgetting

### Permalink
https://escholarship.org/uc/item/3nw4r1gn

### Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 32(32)

### ISSN
1069-7977

### Author
Callender, Aimee

### Publication Date
2010

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

# Expanding Retrieval Promotes Long Term Retention by Preventing Rapid Rates of Forgetting

**Aimee A. Callender (aac0005@auburn.edu)**
Department of Psychology, 226 Thach Hall
Auburn, AL 36849 USA

**Abstract**

Expanding retrieval, increasing the delay between retrieval attempts of recently studied material, should lead to better memory than equally spaced retrieval, however, recent results have been mixed. Ninety-six participants studied word pairs with the following expansion schedules: 1-2-3; 1-5-9; 3-5-8; 5-8-13. An evenly spaced (5-5-5) condition was also used. A final test was given after a 10 minute or 48 hour retention interval. Performance after the 48 hour retention interval was best in the 5-8-13 condition. The higher level of performance in this condition was due to no forgetting between the final learning trial and the immediate final test.

**Keywords:** Memory; Retrieval; Expanding Retrieval;

## Introduction

Spacing out study trials and retrieval attempts is a simple and effective way to improve memory for verbal material(see Cepeda, Pashler, Vul, Wixted & Rohrer, 2006 for a review). Repeated testing has also produced robust memory effects in the laboratory (Roediger & Karpicke, 2006) and classroom studies of repeated testing. One aspect of repeated testing that has produced mixed results is the effects of timing of the repeated tests or retrieval practice on long term retention. Two basic ways that repeated testing can be implemented are to evenly space out the retrieval attempts or to gradually increase the interval between each retrieval attempt, called expanding retrieval. Intuitively, expanding retrieval is thought to improve long term retention, yet the literature is mixed. This experiment investigated two different aspects of the retrieval schedule, the number of intervening items between the study period and the first retrieval attempt and the number of intervening items in the remainder of the schedule, to further delineate the conditions under which expanding retrieval may or may not improve long term retention.

## Expanding Retrieval

Expanding retrieval is a method in which an initial study period is followed by retrieval attempts that are spaced out with increasingly longer intervals between each attempt. For example, a typical expansion sequence used in the literature is a 1-5-9 sequence (see Cull, Shaughnessy, & Zechmeister, 1996; Karpicke & Roediger, 2007), meaning that after the initial study period, there is one intervening item between study and retrieval, then five intervening items between the first and second retrieval attempts, and finally, nine intervening items between the second and third retrieval attempts.

Intuitively, and anecdotally, expanding retrieval is thought to produce better long term retention than evenly spaced retrieval. Theoretically, one would expect expanding retrieval to improve long term memory based on the principle of desirable difficulty (Bjork, 1994). According to Bjork (1994), conditions that introduce difficulty into the learning process increase the likelihood of retrieving the information on a long term retention test. Expanding retrieval is built on the idea of introducing difficulty into the learning process. The initial conceptualization of expanding retrieval posited that the first retrieval attempt should occur soon after the learning trial to ensure successful retrieval. Difficulty is introduced on subsequent retrieval attempts by increasing the number of intervening items between each attempt. Gradually increasing the delay between each retrieval attempt makes each subsequent attempt more difficult than the previous attempt. Surprisingly, recent research investigating the benefits of expanding retrieval has produced mixed results. Some studies have found that expanding retrieval does improve memory when compared to an evenly spaced control condition (Cull et al., 1996) whereas others found that expanding retrieval is no better than evenly spaced retrieval (Logan & Balota, 2008).

Karpicke and Roediger (2007) explored possible reasons for the mixed results and found that the benefits of expanding retrieval depend on two factors: the time of the final test (immediate or delayed) and the number of items between the study trial and the first retrieval attempt. With respect to the time of the final test, expanding retrieval produced better performance than evenly spaced retrieval on a test that occurred 10 minutes after learning. When the final test occurred after a 48 hour retention interval, the effects reversed, and equally spaced retrieval produced superior performance compared to the expanding retrieval conditions. The results of the second factor that was investigated, the delay between the study period and the first retrieval attempt, indicated that expanding retrieval improved retention the most when the first retrieval attempt was delayed regardless of the rest of the sequence (whether it was evenly spaced or expanding).

Although Karpicke and Roediger (2007) addressed and reduced the confusion surrounding expanding retrieval, the extant research on expanding retrieval is restricted by two limitations. One limitation (see Karpicke and Roediger, 2007, Exp. 3) was that the expanding sequences that were used to investigate the number of intervening items between

study and the first retrieval attempt were not true expanding sequences. In order to investigate the interval between the study trial and the first retrieval attempt and to control for the number of intervening items, the sequences were constructed by simply adding an initial retrieval attempt onto the standard 5-5-5- and 1-5-9 sequences. The evenly spaced sequence became a 5-5-5-5 sequence and the expanding sequence became a 5-1-5-9 sequence. The expanding sequence was no longer a true expanding sequence as it contracted before expanding. This experiment addressed this issue by comparing the 5-5-5 sequence against an expanding sequence that started with a first retrieval attempt after 5 intervening items (5-8-13).

The other limitation is that few other studies investigating expanding retrieval have distinguished between the effectiveness of different expansion sequences (but see Logan & Balota, 2008), and most studies have used a 1-5-9 expanding sequence (Cull et al, 1996; Karpicke & Roediger, 2007; Morris, Fritz, Jackson, Nichol, & Roberts, 2005). Other schedules have been used, for example Landauer and Bjork (1978) used 1-4-10 and 0-1-3-8 schedules, and in an experiment with preschool age children, Fritz, Morris, Nolan and Singleton (2007) used a 1 minute-1day-2 day expansion sequence. However, the different schedules have not been compared against one another. Other than Karpicke and Roediger's (2007) investigation of the initial retrieval attempt and Logan and Balota's (2008) study of various schedules of spacing, the remainder of the expansion schedule has not been a variable of interest. Thus, the present experiment investigated different expansion sequences to determine if the benefits of expanding retrieval depend on the particular sequence that is used.

According to Bjork's theory of Desirable Difficulty it is critical to create a task that is challenging and difficult for the individual, but is not so difficult such that the individual cannot complete the task. The difficulty of the expansion sequence can be manipulated by increasing the number of intervening items between the study trial and the first retrieval attempt or by increasing the average number of intervening items across the entire expansion schedule. This experiment manipulated both types of difficulty, but focused on the number of intervening items between the study trial and the first retrieval attempt.

## The Current Experiment

The current experiment investigated 5 different sequences: an evenly spaced control (5-5-5) and four expansion sequences (1-2-3; 1-5-9; 3-5-8; 5-8-13). The 1-5-9 sequence was chosen because it has been used in the majority of previous research. Although the expansion sequences may appear to be chosen at random, the three novel sequences used in this study (1-2-3; 3-5-8; 5-8-13) were chosen because they come from the same number sequence that occurs in nature, the Fibonacci sequence. The Fibonacci sequence is a naturally expanding sequence in which each number is determined by adding together the two previous numbers in the sequence. By using portions of the Fibonacci sequence for each of the novel sequences used in this experiment, each expanding sequence expands in the same way.

The expansion sequences used in this study were chosen from the Fibonacci sequence by considering two factors. First, the critical issue in this experiment was to investigate the number of intervening items between the study trial and the first retrieval attempt. Thus, sequences were selected by choosing different starting points in the Fibonacci sequence to correspond to the initial retrieval attempts that have been used in previous literature (1 or 5). This resulted in the 1-2-3 and 5-8-13 sequences. The other factor that was considered was the average number of intervening items across the entire sequence. The average number of intervening items in the 5-5-5 and 1-5-9 sequences that are typically used in the literature is 5. Thus, the 3-5-8 condition was selected.

The comparison of most interest was to investigate the difference in performance between the 5-5-5 condition and the expanding condition that also had five intervening items between the study trial and the first retrieval attempt, 5-8-13. This comparison was vital because Karpicke and Roediger's (2007) investigation of this variable found that delaying the first retrieval attempt improved performance regardless of the rest of the expansion sequence. However, the expansion sequences were different from the expansion sequences used in this experiment as their sequences started over from 1 after the first retrieval attempt (5-1-5-9) whereas the expansion sequence in the present experiment continued to expand from that first retrieval attempt (5-8-13).

## Method

### Participants

Ninety-six undergraduate students at Auburn University participated in the study in exchange for extra credit for a psychology course. Participant were between the ages of 17 and 25, 74% were female, 26% were male, 88% were white, 9% were African American, and 3% identified themselves as another race/ethnicity.

### Design and Materials

The experiment used a 2 (retention interval) x 7 (testing schedule) mixed design. The final retention interval (10 minutes or 48 hours) was manipulated between subjects and the various testing schedules were manipulated within subjects. The testing schedules included an evenly spaced control (5-5-5), 3 expanding schedules based on the Fibonacci sequence (1-2-3; 3-5-9; 5-8-13), a standard expanding schedule (1-5-9), and two single test conditions in which participants took a single test immediately (Single0) or after 2 intervening items (Single2).

The experiment was based on Karpicke and Roediger (2007), using word pairs in which the first word was an unfamiliar (low frequency) word such as *Tumbrel,* and the

second word was a one word synonym or definition, *Cart*. During the study phase, the word pairs were presented together, *Tumbrel-Cart*, and during the testing phases the first word was presented and participants were required to type in the appropriate word pair. Fifty-six word pairs were constructed. Forty of the word pairs were critical word pairs. For each of the 7 testing conditions, there were 5 word pairs to study. The remaining 5 items served as unstudied control items. Eight counterbalancing conditions were constructed to allow for each set of 5 word pairs to be rotated through each of the testing and unstudied conditions. This resulted in 40 critical word pairs (5 for each of the 7 conditions, counterbalanced so remaining 5 were unstudied or control items). Additionally, 16 filler items were included for a total of 56 word pairs. Two filler items were used as buffers at the beginning of the task, and 3 filler items were used as buffers at the end of the task. The remaining fillers were interspersed throughout the list to allow for the appropriate spacing of all of the study and test trials. This resulted in a total of 142 trials in the experiment.

## Procedure

Data collection occurred in groups of up to 15 participants in a computer lab. Participants were instructed that they would study word pairs that included one familiar word and one unfamiliar word. They were instructed to study the word pairs and type in the appropriate response during the test trials. Each of the trials (study or test) was 8 s. with a 500 ms. intertrial interval. Participants had to spend the entire 8 s. viewing the study screen, but during the test trials they were allowed to press Enter to move on to the next trial once they entered their response. If no response had been entered within 8 s., the computer program automatically advanced to the next trial. This task (142 trials) generally took between 15 and 20 minutes for participants to complete.

The final retention test tested participants on the 40 critical word pairs. Thirty-two of the participants took the final test after 10 minutes, and 32 took the test after a 48 hour delay. Participants were instructed that they were going to be tested on the words that they had learned previously, and to type in the appropriate word pair. Each trial was 14 s. (participants could press Enter to advance to the next trial once they entered a response) and the interstimulus interval was 500 ms. If no response had been entered in, the computer program automatically advanced after 14 seconds. This task generally took about 10 minutes to complete.

## Results

Three separate analyses were conducted. The first analysis investigated performance on the learning trials, that is, each retrieval attempt during the expanding or evenly spaced sequence. A second analysis was conducted on final cued-recall performance. The third analysis investigated forgetting between the last learning trial and the final cued-recall test.

### Learning Trials

A 5(spacing condition) x 3 (learning trial) within subjects ANOVA was conducted (the single trial conditions were not included in this analysis). The effect of learning trial was significant, $F$ (2, 464) = 5.82, $p$ = .003, *partial* $?^2$ = .02, which was qualified by a significant interaction of learning trial and spacing condition $F$ (8, 930) = 4.52, $p$ < .001, *partial* $?^2$ = .04. Table 1 shows that performance in the two conditions in which there was only one intervening item between the study trial and the initial retrieval attempt (1-2-3 and 1-5-9) was much higher on the learning trials than the other three spacing conditions (3-5-8; 5-8-13; 5-5-5), indicating that the trials that had more intervening items before the first retrieval attempt were more difficult to learn (See Table 1).

Table 1: Performance on learning trials by spacing condition.

| Spacing | Trial 1 | Trial 2 | Trial 3 |
|---|---|---|---|
| 1-2-3 | .68 (.02) | .56 (.02) | .58 (.02) |
| 1-5-9 | .66 (.03) | .52 (.03) | .56 (.06) |
| 3-5-8 | .43 (.03) | .48 (.07) | .42 (.03) |
| 5-8-13 | .46 (.03) | .44 (.03) | .43 (.04) |
| 5-5-5 | .43 (.03) | .44 (.03) | .45 (.03) |

*Note. Standard errors are in parentheses.*

### Final Cued Recall

A 2 (delay) x 7 (spacing condition: 1-2-3; 1-5-9; 3-5-8; 5-8-13; 5-5-5; Single0; Single2) repeated measures ANOVA was conducted on the final recall performance. Both the within-subjects effect of spacing and the effect of delay were significant; $F$ (6, 89) = 25.45, $p$ < .001, *partial* $?^2$ = .64, and $F$ (1, 94) = 26.14, $p$ < .001, *partial* $?^2$ =.22, respectively. The spacing by delay interaction was marginally significant, $F$ (6, 89) = 2.15, $p$ = .06, *partial* $?^2$ = .07. As Table 2 shows, performance in the 5-8-13 condition was the highest in both the immediate condition (although not significantly) and was significantly higher than the other spacing conditions (except for the 1-5-9 condition) after the 48 hour delay. The 5-5-5 spacing condition was not significantly better than the other spacing conditions on either the immediate or delayed test.

Table 2: Final recall as a function of spacing condition and delay.

| Spacing | 10 min. | 48 hours |
|---|---|---|
| 1-2-3 | .37 (.03) | .17 (.03) |
| 1-5-9 | .37 (.04) | .18 (.04) |
| 3-5-8 | .34 (.04) | .15 (.04) |
| 5-8-13 | .42 (.04) | .25 (.04) |
| 5-5-5 | .38 (.04) | .15 (.04) |
| Single 0 | .14 (.02) | .02 (.02) |
| Single 2 | .23 (.03) | .07 (.03) |

*Note. Standard errors are in parentheses.*

## Forgetting

A 2 (retention interval) X 5 (spacing condition: 1-2-3; 1-5-9; 3-5-8; 5-8-13; 5-5-5) repeated measures ANOVA was conducted on the amount of forgetting that occurred between the final learning trial and the final test. There was significantly more forgetting in the delayed condition than the immediate condition, $F$ (1, 94) =55.69, $p < .001$, *partial* $?^2$=.37. Forgetting also depended on the spacing condition, $F$ (4, 91) = 19.28, $p < .001$, *partial* $?^2$=.46, but there was no interaction. Table 3 shows the amount of forgetting in each condition. Comparing the 5-8-13 condition to the 5-5-5 condition, there was less forgetting on the immediate test (although not significantly less), and significantly less forgetting on the delayed final test, $p = .001$.

Table 3: Forgetting as a function of spacing and delay.

| Spacing | 10 min. | 48 hours |
|---|---|---|
| 1-2-3 | .18 (.03) | .45 (.03) |
| 1-5-9 | .13 (.09) | .44 (.09) |
| 3-5-8 | .08 (.03) | .27 (.03) |
| 5-8-13 | .02 (.03) | .18 (.03) |
| 5-5-5 | .07 (.03) | .30 (.03) |

*Note. Standard errors are in parentheses.*

## Discussion

Expanding retrieval can lead to better long term retention than evenly spaced retrieval when a slight modification is made to the original conceptualization of the method. Long term retention is best when the learning trials are constructed by combining a delayed initial retrieval attempt with an expanding sequence for the remainder of the learning trials. In accordance with Karpicke and Roediger (2007), increasing the number of intervening items between the study trial and the first retrieval attempt is critical to improve learning and memory of the information. Additionally, as the current experiment shows, it is also critical for the sequence to expand for optimal long-term retention.

Several findings from this experiment provide insight into the conditions under which expanding retrieval will produce better long term retention than evenly spaced retrieval. The first relevant finding is, as stated above, the first retrieval attempt after the study period must be sufficiently difficult. Both of the spacing conditions that only had a single item between study and retrieval produced high performance on the initial retrieval attempt, but those conditions also saw a large decline in performance across the learning trials. The 1-2-3 and 1-5-9 conditions, for example, both had a 10 percentage point drop in performance between the first and third retrieval attempts. The conditions that were more difficult and had at least three items between studying and the first retrieval attempt had no forgetting across the learning trials. Thus, increasing the difficultly of the first retrieval attempt produced steady performance throughout the learning trials. Maintaining a steady rate of performance throughout learning may be important in preventing rapid forgetting from occurring.

The next important finding was concerned with forgetting in each of the conditions. Forgetting, as Tables 1 and 3 show, was quite rapid in the 1-2-3 and 1-5-9 conditions both during learning (as performance decreased across learning trials) and during the final test retention intervals. Although there was no forgetting across learning trials for the sequences that began with a larger number of intervening items, performance on the third and final learning trial was still higher in the 1-2-3 and 1-5-9 conditions than the 3-5-8, 5-8-13 or 5-5-5 conditions. However, on the final test given 10 minutes later, the 5-8-13 condition produced numerically better performance than the other conditions, including the evenly spaced (5-5-5) condition. This result was due to almost no forgetting in the 5-8-13 condition. This trend continued, and the 5-8-13 condition produced significantly better performance than evenly spaced retrieval on the final test administered after a 48 hour retention interval. In this retention interval condition, the 5-8-13 sequence produced half the amount of forgetting as the 5-5-5 condition.

The 5-5-5- and 5-8-13 conditions were considered the critical comparison in this study and addressed the question of whether expanding retrieval can produce superior performance on a long term retention test compared to evenly spaced retrieval. On the learning trials, these two conditions produced similar levels of performance. After a 10 minute retention interval the 5-8-13 condition produced slightly better results, and after 48 hours, the 5-8-13 condition produced significantly better memory for the word pairs than the 5-5-5 condition.

Thus far, the question of whether expanding retrieval can produce superior performance compared to evenly spaced retrieval on immediate and long term retention tests has been mixed at best. In fact, equally spaced retrieval has generally produced better performance than expanding retrieval on tests taken at least 24 hours later. Karpicke and Roediger (2007) even noted that, "we know of no existing study using a continuous paired associate learning task…that has shown that expanding retrieval produces greater long-term retention (after delays greater than 24 hr) than equally spaced practice." The present experiment is one that does show that expanding retrieval produces greater long-term retention after 48 hours. Based on the current results, it is not entirely surprising that the previous research has been so mixed. Considering that a key factor in this experiment was the difficulty of the initial retrieval attempt, the 1-5-9 condition that has been used most widely is not a sequence that would be expected to improve long term retention. This also explains why the 5-5-5 condition, which begins with a difficult initial retrieval attempt, has produced such good performance in previous experiments.

The unique 5-8-13 expansion sequence, which combined a difficult first retrieval attempt with expanding retrieval, not only resulted in virtually no forgetting during both the

learning trials and on a test administered after a 10 minute retention interval. This particular sequence prevented the rapid forgetting that normally occurs immediately after learning takes place. Further, and of most importance to the question at hand, is that the combination of a difficult initial retrieval attempt and an expanding sequence resulted in half the amount of forgetting on the long-term retention test when compared to the evenly spaced control.

In summary, increasing the difficulty of the initial retrieval attempt protects against rapid forgetting that can occur within minutes of the study trial. This is evidenced by the small amount of forgetting during learning and on the 10 minute retention interval test in both the 5-5-5 and the 5-8-13 conditions. However, expanding retrieval from a difficult initial retrieval attempt produces better performance on a final test 48 hours. Combining a difficult initial retrieval attempt with even more difficult subsequent attempts reduces the forgetting that normally occurs immediately after learning (described by Ebbinghaus, 1885/1913; Rubin, Hinton & Wenzel, 1999) and up to 48 hours later.

## Acknowledgments

This project was funded by Auburn University's College of Liberal Arts.

## References

Bjork, R.A. (1994). Memory and metamemory considerations in the training of human beings. I J.Metcalfe & a. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185-205). Cambridge, MA: MIT Press.

Cepeda N., Pashler, H., Vul, E., Wixted, J., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132,* 354-380

Cull, W., Shaughnessy, J., & Zechmeister, E. (1996). Expanding understanding of the expanding-pattern-of-retrieval mnemonic: Toward confidence in applicability. *Journal of Experimental Psychology: Applied, 2,* 365-378

Fritz, C., Morris, P., Nolan, D., & Singleton, J. (2007). Expanding retrieval practice: An effective aid to preschool children's learning. *The Quarterly Journal of Experimental Psychology, 60*(7), 991-1004.

Karpicke, J.D. & Roediger, H. L. (2007). Expanding retrieval practice promotes short-term retention, but equally spaced retrieval enhances long-term retention. *Journal of Experimental Psychology: Learning, Memory & cognition,* 33, 704-719.

Landauer, T.K. & Bjork, R.A. (1978). Optimum rehearsal patterns and name learning. In M.M. Gruneberg, P.E. Morris, & R.N. Sykes (Eds.), *Practical aspects of memory* (pp. 625-632). London: Academic Press.

Logan, J.M. & Balota, D.A. (2008). Expanded vs. equal interval spaced retrieval practice: Explorign different schedules of spacing and retention interval in younger and older adults. *Aging, Neuropsychology and Cognition, 15,* 257-280.

Morris, P., Fritz, C., Jackson, L., Nichol, E., & Roberts, E. (2005). Strategies for Learning Proper Names: Expanding Retrieval Practice, Meaning and Imagery. *Applied Cognitive Psychology, 19*(6), 779-798

Roediger, H., & Karpicke, J. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science, 1,* 181-210.

Rubin, D., Hinton, S., & Wenzel, A. (1999). The precise time course of retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 1161-1176.