2021 ASEE ANNUAL CONFERENCE
Virtual Meeting | July 26–29, 2021 | Pacific Daylight Time
Paper ID #32474

Introducing Desirable Difficulty in Engineering Mathematics with Spaced Retrieval Practice

Dr. Campbell R. Bego, University of Louisville

An Assistant Professor with research in engineering education, Campbell R. Bego, PhD, PE, is interested in improving STEM student learning and gaining understanding of STEM-specific learning mechanisms through controlled implementations of evidence-based practices in the classroom. Dr. Bego has an undergraduate Mechanical Engineering degree from Columbia University, a Professional Engineering license in the state of NY, and a doctorate in Cognitive Science.

Dr. Patricia A. Ralston, University of Louisville

Dr. Patricia A. S. Ralston is Professor and Chair of the Department of Engineering Fundamentals at the University of Louisville. She received her B.S., MEng, and PhD degrees in chemical engineering from the University of Louisville. Dr. Ralston teaches undergraduate engineering mathematics and is currently involved in educational research on the effective use of technology in engineering education, the incorporation of critical thinking in undergraduate engineering education, and retention of engineering students. She leads a research group whose goal is to foster active interdisciplinary research which investigates learning and motivation and whose findings will inform the development of evidence-based interventions to promote retention and student success in engineering. Her fields of technical expertise include process modeling, simulation, and process control.

Dr. Keith B. Lyle, University of Louisville
Dr. Jason Immekus, University of Louisville

Dr. Immekus is associate professor in the Department of Educational Leadership, Evaluation, and Organizational Development.

Introducing desirable difficulty in engineering mathematics with spaced retrieval practice

NSF Award #1912253

Abstract

This study examines the difficulty introduced by spaced retrieval practice in Calculus I for undergraduate engineering students. Spaced retrieval practice is an instructional technique in which students engage in multiple recall exercises on the same topic with intermittent temporal delays in between. Spacing out retrieval practice increases the difficulty of the exercises, reducing student performance on them. However, empirical research indicates that spaced retrieval practice is associated with improvements in students' long-term memory for the retrieved information. The short-term costs and long-term benefits of spaced retrieval practice is an example of *desirable difficulty*, when more difficult exercises during the early stages of learning result in longer-lasting memory [1].

With support from the National Science Foundation (NSF), we sought to address: *Does spacing decrease performance on retrieval practice exercises in an engineering mathematics course?* Results showed that student performance was significantly lower for questions in the spaced condition than questions in the massed condition, indicating that we successfully increased the difficulty of the questions by spacing them out over time. Future work will assess final quiz performance to determine whether spacing improved long-term course performance, i.e., whether the difficulty imposed by spacing was desirable.

Introduction

First-year mathematics courses often pose barriers to success in engineering [2]–[5]. Despite encouragement from the Mathematical Association of America in 1988 to make calculus a "pump" instead of a "filter" [6], recent studies are still finding that engineering students who struggle in their mathematics courses are much less likely to persist to graduation [7]–[9]. Increasing student success in mathematics is therefore paramount to fulfilling current and future demands for engineers, which has been established as a national goal [10].

An evidence-based educational technique that has recently proven to be effective in preparatory engineering mathematics (precalculus) is spaced retrieval practice [11], [12]. In spaced retrieval practice, instructors ask multiple questions about a topic repeatedly over time, with intermittent delays[13]. Although there is copious research support for spacing, practice problems in mathematics textbooks tend to be clustered by topic such that students rarely have to revisit a question type following a delay [14]. Therefore, it is up to instructors to both discover and implement spaced retrieval practice themselves. Spaced retrieval can be implemented by asking questions about prior topics either in class or outside of class on homework assignments or quizzes (see [15] for additional details regarding implementation). Delaying questions to create spacing, also known as distributed practice, has been found to improve memory in experimental lab settings (e.g., [13], [16]), as well as in the classroom (e.g., [17]).

Spaced retrieval practice has been found to benefit both short- and long-term memory (see [18]). In two recent studies on spaced retrieval in a precalculus course for engineering students, results indicated that spaced retrieval on weekly quizzes improved student performance on a final exam as well as on a preparedness assessment at the beginning of Calculus I in the following semester [11], [12]. Maintaining learning and memory is key to cumulative mathematics course sequences, which establish a foundation for discipline-specific courses.

From a learner's perspective, spaced retrieval practice adds an additional layer of difficulty during the learning process. Specifically, instead of students applying information that they may have recently reviewed, they are required to recall previously learning information to categorize problems and identify appropriate solution methods. Thus, performance on spaced retrieval questions is often lower than if the same questions were asked without spacing (e.g., [12]), referred to as massed.

Although low immediate performance may seem to indicate that learning is hindered instead of aided, the concept of desirable difficulty proposes that difficulty during initial learning can result in lasting learning gains [19], [20]. To understand desirable difficulty, a distinction must be made between performance and learning: performance is highly context-based and responsive to environmental cues and affect, whereas learning is a permanent change in knowledge that results in long-term retrieval and transfer of information.

Learning scientists define desirable difficulties as those that activate cognitive processes that support deep learning, such as encoding or retrieval [19]. According to the leading researchers on desirable difficulty, Bjork and Bjork, good practices that impose desirable difficulty include: varying the learning context, i.e., not studying in the same place every time [21]; interleaving [22]; spacing out study sessions over time [16]; and self-testing instead of rereading during study sessions [18], [23]. Not all difficulties are desirable. Methods that distract attention, increase working memory load with extraneous information, or confuse the learner would hinder long-term learning.

The difficulty imposed by spaced retrieval practice, if observed, is likely desirable, since the difficulty arises from students' activation of retrieval processes. Because students are forced to recall previously-learned information without any contextual cues, their memory is strengthened.

Current Work

As part of NSF Award #1912253, we implemented spaced retrieval practice in Calculus I for engineering students at the University of Louisville's J. B. Speed School of Engineering. The three-year grant is currently in its second year, the study having been implemented in Fall 2020.

As a preliminary analysis on the data available for the ASEE timeframe, we asked the following research question:

> *RQ: Does spacing decrease performance on retrieval practice exercises in an engineering mathematics course?*

Methods

Participants

Participants (*N* = 183) were students enrolled in Calculus I for Engineers in Fall 2020 at the University of Louisville's J. B. Speed School of Engineering who completed all retrieval practice assignments. Students who completed only a portion of assignments (*N* = 55) were excluded.

Procedures

Course instructors selected 8 target learning objectives (LOs) and assessment questions from an online learning textbook for three units of time (weeks 1-3, 4-5, and 6-7), for a total of 24 LOs and questions. The online textbook offered several algorithmic variants for each assessment question, so that the same question could be asked multiple times but appear different to students.

Researchers used the selected LOs and questions to build 5 quizzes (administered after weeks 3, 5, 7, 9, and 11) and a final assessment which was administered during class time on the last day of class. The experimental manipulation was within-subjects, as follows:

- Half of the LOs were assigned to a *massed* condition, in which the assessment question was administered three times on a single quiz.
- The other half of the LOs were assigned to a *spaced* condition, in which the question was administered over three consecutive quizzes.
- The final assessment included all 24 questions.

The distribution of questions over the quizzes is illustrated in Table 1 for LOs in the massed condition and in the spaced condition.

The columns of Table 1 illustrate the content covered in each quiz. Quiz 2, for example, included questions from all LOs in Unit 2 and half of the LOs for Unit 1. Quiz 3 had the most questions, and included questions from Units 1, 2, and 3. Quiz 5 only included questions from half of the LOs of Unit 3.

At the beginning of the semester, students were randomly divided into groups A and B and assigned a set of five quizzes to each group. Group A was assigned odd-numbered LOs in a massed condition and even-numbered LOs in a spaced condition. Group B was assigned even-numbered LOs in a massed condition and odd-numbered LOs in a spaced condition. This provided counterbalancing such that each objective was spaced for half of the students and massed for the other, removing the potential confound of objective difficulty.

Table 1. Distribution of questions as a function of unit, condition, and quiz number.

	Quiz 1	Quiz 2	Quiz 3	Quiz 4	Quiz 5	Final Quiz
Unit 1 Massed LOs	3 Questions					1 Question
Unit 1 Spaced LOs	1 Question	1 Question	1 Question			1 Question
Unit 2 Massed LOs		3 Questions				1 Question
Unit 2 Spaced LOs		1 Question	1 Question	1 Question		1 Question
Unit 3 Massed LOs			3 Questions			1 Question
Unit 3 Spaced LOs			1 Question	1 Question	1 Question	1 Question

Materials

Online Textbook. The questions in the experimental manipulation were drawn from Pearson's MyLabsPlus ® Calculus I content. For details on the LOs and questions, please contact the corresponding author.

Quizzes. Quizzes included only the massed and spaced questions from target LOs in units 1-3. Quizzes 1-5 were assigned after weeks 3, 5, 7, 9, and 11, respectively. Students could access quizzes from Friday at 1pm through Sunday night at midnight through the online textbook. These quizzes were un-proctored, but students were forced to complete each quiz in one sitting and were limited to a specified amount of time after opening the quiz. The amount of time was proportional to the number of questions (maximum: 24 questions, 90 minutes). The amount of time was such that there was plenty of time available (i.e., students had enough time to answer all questions), but students would not have the flexibility of leaving the assignment open and doing something else during the time frame.

Syllabus Integration. Quizzes were called "Cumulative Quizzes" to differentiate them from weekly homework assignments, weekly exams, and the cumulative final exam. The Cumulative Quizzes were assigned a value of 7% of the overall grade. If students completed all 5 quizzes and a final assessment, they received a bonus of 10% on their quiz average, up to 100% maximum.

Analysis

We compared student performance on massed questions to student performance on spaced questions. Student performance was averaged for all questions asked on the cumulative quizzes in a specific condition. We then performed a paired t test to compare student performance in the massed and spaced conditions. The criteria to reject the null hypothesis (no difference between conditions) was assumed to be $p < 0.05$.

Results

The paired *t* test revealed that students performed significantly worse on spaced questions (M = 77.91%, SD = 13.10%) than massed questions (M = 80.60%, SD = 12.74%), $t(182) = 3.00$, $p = .003$, Hedges $g_{av} = .21$. The mean difference was 2.69%, 95% CI [.92%, 4.45%]. Means are illustrated in Figure 1.

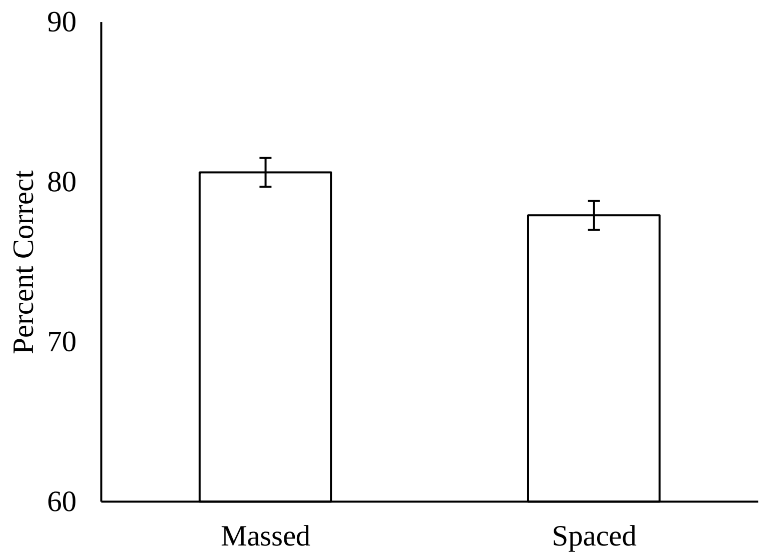

Figure 1. Quiz performance by experimental condition. Error bars = ±SE.

Discussion

In the current implementation of spaced retrieval practice in Calculus I for engineering students, we observed significantly lower performance in the spaced condition than the massed condition with a mean difference of 2.61%. This mid-award result is encouraging, as it indicates that spaced retrieval was associated with significantly greater difficulty than massed retrieval. We believe greater difficulty is desirable because it requires students to activate retrieval processes, which is known to facilitate learning and improve memory.

The difference we observed is similar to the 2.0% difference value reported in a recent study of spaced retrieval practice in precalculus for engineers [12]. In that study, the implementation consisted of weekly quizzes, and the spacing manipulation was different; the first question was asked in the week the topic was introduced, the second question was asked a week later, and the third question was asked two weeks after that. In our implementation, we had equal spacing in consecutive assignments which were each 2 weeks apart. The similarity of the results could indicate that the difficulty of spaced retrieval is generalizable in some way, although more studies are necessary to determine what factors contribute to the amount of difficulty (e.g., the temporal delay length, the type of knowledge retrieved, discipline).

It is also important to consider the impact of spaced retrieval practice on students' final grades. In first-year engineering courses, it is preferable to implement educational techniques that that don't hurt performance in such a way that demotivates students or penalizes their GPAs. Because the performance reduction averaged at only 2.61%, and this reduction was applied to 7% of students' grades (with an opportunity for a small bonus for full participation), it is not likely that this strategy was detrimental to student performance overall. It is also expected that performance increased on a final assessment due to spacing. Spaced retrieval practice therefore appears to be a promising method to use in engineering mathematics.

Limitations

It is possible that this analysis underestimates the difference in performance between massed and spaced questions. In both conditions, one question was asked on the quiz immediately following the unit; spacing was only applied to questions 2 and 3 of the set of 3. It is possible that there is a bigger difference between performance on spaced and massed questions than we are observing in the mean average. We will be investigating this idea in future work.

Importantly, we did not yet assess whether the difficulty imposed by spaced retrieval practice in engineering mathematics was desirable. To do this, we need to determine whether spacing benefitted learning on the final assessment. We expect there to be a learning benefit because of the robust effect of spacing observed throughout the literature, and because spaced retrieval invokes cognitive processes that are supportive of deep learning. We will address learning directly in a future publication.

Conclusion

For NSF Award #1912253, we implemented spaced retrieval practice in an engineering mathematics course. In this mid-award paper and poster, we found that temporal spacing significantly reduced student performance on quiz questions. We believe that spaced retrieval practice imposes desirable difficulty, and our future work will investigate whether spacing led to increased learning.

References

[1] R. A. Bjork, “Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. Shimamura (Eds.),” *Metacognition: Knowing about knowing.* Cambridge, MA: MIT Press., pp. 185–205, 1994.

[2] J. Moore, “Undergraduate mathematics achievement in the emerging ethnic engineers programme,” *Int. J. Math. Educ. Sci. Technol.*, 2005, doi: 10.1080/00207390500063975.

[3] P. K. Subramaniam, M. Cates, and G. Borislava, “Improving success rates in calculus,” *MAA Focus*, vol. 28, no. 5, pp. 20–21, 2008.

[4] M. W. Ohland, A. G. Yuhasz, and B. L. Sill, “Identifying and removing a calculus prerequisite as a bottleneck in clemson’s general engineering curriculum,” *J. Eng. Educ.*, vol. 93, no. 3, pp. 253–257, 2004, doi: 10.1002/j.2168-9830.2004.tb00812.x.

[5] Y. Min, G. Zhang, R. A. Long, T. J. Anderson, and M. W. Ohland, “Nonparametric survival analysis of the loss rate of undergraduate engineering students,” *J. Eng. Educ.*, vol. 100, no. 2, pp. 349–373, 2011, doi: 10.1002/j.2168-9830.2011.tb00017.x.

[6] L. A. Steen, Ed., *Calculus for a new century: a pump, not a filter. Papers from a national colloquium (MAA Notes No. 8, October 28-29, 1987)*. Washington, DC: Mathematical Association of America, 1988.

[7] C. R. Bego, J. L. Hieb, and P. A. S. Ralston, “Barriers and bottlenecks in engineering mathematics: How performance throughout a math sequence affects retention and persistence to graduation,” in *October 2019 IEEE Frontiers in Education Conference (FIE)*, 2019.

[8] J. J. Pembridge and M. A. Verleger, “First-year math and physics courses and their role in predicting academic success in subsequent courses,” in *Proceedings of the 120th ASEE Annual Conference and Exposition*, 2013.

[9] D. G. Beanland, “Challenges and opportunities facing the education of engineers-Address to Victoria division of engineers Australia seg meeting.” Melbourne, 2010.

[10] PCAST President’s Council on Advisors on Science and Technology and President’s Council on Advisors on Science and Technology, “Engage to Excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics,” Washington, DC: Office of the President, 2012.

[11] R. F. Hopkins, K. B. Lyle, J. L. Hieb, and P. A. S. Ralston, “Spaced retrieval practice increases college students’ short-and long-term retention of mathematics knowledge,” *Educ. Psychol. Rev.*, vol. 28, no. 4, pp. 853–873, 2016, doi: 10.1007/s10648-015-9349-8.

[12] K. B. Lyle, C. R. Bego, R. F. Hopkins, P. A. S. Ralston, and J. L. Hieb, “How the amount and spacing of retrieval practice affect the short- and long-term retention of mathematics knowledge,” *Educ. Psychol. Rev.*, 2019, doi: 10.1007/s10648-019-09489-x.

[13] S. K. Carpenter, N. J. Cepeda, D. Rohrer, S. H. K. Kang, and H. Pashler, “Using spacing to enhance diverse forms of learning: Review of recent research and implications for instruction,” *Educational Psychology Review*, vol. 24, no. 3. pp. 369–378, 2012, doi: 10.1007/s10648-012-9205-z.

[14] D. Rohrer, R. F. Dedrick, and M. K. Hartwig, “The scarcity of interleaved practice in mathematics textbooks,” *Educ. Psychol. Rev.*, vol. 2020, pp. 1–11, 2020.

[15] C. R. Bego, P. A. Ralston, K. B. Lyle, and J. C. Immekus, “Research to practice to research: Intrinsic requirements of implementing and studying spaced retrieval practice in STEM courses,” in *October 2020 IEEE Frontiers in Education Conference (FIE)*, 2020, pp. 1–5.

[16] N. J. Cepeda, H. Pashler, E. Vul, J. T. Wixted, and D. Rohrer, “Distributed practice in verbal recall tasks: A review and quantitative synthesis,” *Psychol. Bull.*, vol. 132, no. 3, pp. 354–380, 2006, doi: 10.1037/0033-2909.132.3.354.

[17] R. A. R. Gurung and K. Burns, “Putting evidence-based claims to the test: A multi-site classroom study of retrieval practice and spaced practice,” *Appl. Cogn. Psychol.*, vol. 33, no. 5, pp. 732–743, 2019, doi: 10.1002/acp.3507.

[18] C. A. Rowland, “The effect of testing versus restudy on retention: A meta-analytic review of the testing effect.,” *Psychol. Bull.*, vol. 140, no. 6, pp. 1432–1463, 2014, doi: 10.1037/a0037559.

[19] E. L. Bjork and R. A. Bjork, “Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning,” in *Psychology and the real world: Essays illustrating fundamental contributions to society*, 2nd Editio., M. A. Gernbacher and J. Pomerantz, Eds. New York, NY: Worth, 2014, pp. 59–68.

[20] N. C. Soderstrom and R. A. Bjork, “Learning versus performance: An integrative review,” *Perspect. Psychol. Sci.*, vol. 10, no. 2, pp. 176–199, 2015, doi: 10.1177/1745691615569000.

[21] S. M. Smith and E. Vela, “Environmental context-dependent memory: A review and meta-analysis,” *Psychon. Bull. Rev.*, vol. 8, no. 2, pp. 203–220, 2001, doi: 10.3758/BF03196157.

[22] M. S. Birnbaum, N. Kornell, E. L. Bjork, and R. A. Bjork, “Why interleaving enhances inductive learning: The roles of discrimination and retrieval,” *Mem. Cogn.*, vol. 41, no. 3, pp. 392–402, 2013, doi: 10.3758/s13421-012-0272-7.

[23] J. D. Karpicke and H. L. Roediger, “Repeated retrieval during learning is the key to long-term retention,” *J. Mem. Lang.*, 2007, doi: 10.1016/j.jml.2006.09.004.