Spaced Retrieval Practice in Two Fundamental Engineering Courses: Calculus and Physics

Ryan J. Patrick
Department of Psychological and Brain Sciences
College of Arts & Sciences
University of Louisville
ryan.patrick@louisville.edu

Campbell R. Bego
Department of Engineering Fundamentals
J. B. Speed School of Engineering
University of Louisville
campbell.bego@louisville.edu

Raymond J. Chastain
Department of Physics & Astronomy
College of Arts & Sciences
University of Louisville
raymond.chastain@louisville.edu

Patrica A. S. Ralston
Department of Engineering Fundamentals
J. B. Speed School of Engineering
University of Louisville
patrica.ralston@louisville.edu

Jason C. Immekus
Department of Educational Leadership, Evaluation and Organizational Development
College of Education
University of Louisville
jason.immekus@louisville.edu

Keith B. Lyle
Research and Development
Transfr
New York, NY
klyle@transfrvr.com

***Abstract*—This full-length, research-to-practice paper discusses an ongoing NSF project aimed at implementing spaced retrieval practice in Science, Technology, Engineering, and Mathematics (STEM) classes. By implementing evidence-based learning practices in the classroom, educational performance, and thereby retention and graduation rates, are thought to be improved. Retrieval practice refers to the repeated recall of information from memory to increase the information's future accessibility. Retrieval practice can be massed or spaced. In massed practice, all acts of recall occur during a single short temporal window. In spaced practice, acts of recall are temporally distributed. This project implements spaced retrieval practice in STEM barrier courses in the fields of biology, chemistry, engineering, mathematics, physics, and psychology. The current dissemination focuses on spaced retrieval practice in one section of Calculus for Engineers and two sections of Fundamentals of Physics I, with similarly delivered implementations. Spaced retrieval practice significantly improved student performance in calculus but not in the two sections of physics. Potential explanations for these results are discussed, as well as implications.**

Keywords—spaced retrieval practice, engineering education, calculus, physics, feedback.

I. Introduction

Introductory Science, Technology, Engineering, and Mathematics (STEM) courses are intended for novices to obtain fundamental knowledge in a variety of fields. Unfortunately, these courses pose a challenge for many learners and are often referred to as "barriers" to STEM degrees [1]. Anticipated growth in the U.S. STEM job market means that empirical classroom research is needed to identify effective instructional techniques that can increase student success [2]. This paper describes part of an NSF-funded study, as well as a follow-up study outside the original scope of the grant.

Initially, spaced retrieval practice was implemented in two STEM courses, calculus, and physics. These courses utilized different online learning platforms, and thus presented slightly different forms of feedback for students following retrievals. Specifically, students in calculus had access to more comprehensive feedback than students in physics. As presented in detail below, results showed that students benefitted from spaced retrieval in calculus but not in physics. Based on those preliminary results, a second study was conducted in the same physics course the following year, with a modification to the feedback settings. This paper describes the results of these studies and discusses a variety of implications.

II. Literature Review

A breadth of research has shown that long-term retention of information is improved by repeatedly recalling the information from memory—a behavior known as retrieval practice [3]–[5]. As a mnemonic technique, retrieval practice is based on the testing effect, which is the finding that long-term retention of information benefits more from retrieving the information from memory than from restudying it [5], [6]. While the testing effect has been empirically established, students do not often engage in self-directed retrieval [7]. Classroom studies have shown to be effective in improving students' ability to retain information and achieve higher grades when instructors require a greater degree of retrieval [8]–[10]. Accordingly, it is crucial that instructors provide opportunities for retrieval practice (e.g., homework, in-class work, quizzes, etc.) in order for students to benefit from the testing effect.

Additional literature in cognitive science has indicated that retention can be further improved by distributing instances of retrieval over time (spacing) versus performing retrievals in one

This project was funded by the National Science Foundation, Award #1912253.

short temporal window (massing) [3], [11], [12]. This is known as the spacing effect. Spacing retrieval attempts requires learners to recall information from long-term memory as opposed to short-term memory. Repeated retrieval from long-term memory can strengthen retrieval-recall neural connections and create new connections, increasing the ease of retrieval [13]. Spaced retrieval practice therefore builds on both the testing effect and the spacing effect.

Like retrieval practice, spacing is not frequently self-initiated by learners. Similarly, mathematics educational materials predominately employ end-of-chapter review sections in a "chunked" fashion [14]. This layout is discrete and does not space these retrieval practice opportunities.

The beneficial mnemonic effects of spaced retrieval practice have been well established in laboratory studies [15], [16], but only recently has spaced retrieval practice been studied in actual classrooms [4], [17]–[20]. Unlike laboratory studies that can be tightly controlled to change only one variable at a time, classrooms studies can be highly volatile and subject to interference from numerous extraneous variables. Differences in instructor, course structure, and course content are just a selection of variables that may impact the efficacy of spaced retrieval practice in the classroom. Additionally, in classroom studies, students have an extensive amount of control over their learning which can make it difficult to assess the efficacy of instructional techniques.

A. Spaced Retrieval Practice in Mathematics

There have been few empirical studies of spaced retrieval practice within the classroom, but most were in the field of mathematics. These studies indicated that spaced retrieval practice benefits learners' knowledge of course materials as well as reasoning skills [19]. Implementations in undergraduate precalculus courses showed improved long-term performance [4], [18]. In these studies, the researchers manipulated whether retrieval practice was massed or spaced. End-of-semester retention of practiced material was significantly improved by spacing.

Although spacing retrieval practice in mathematics appears to have long-term benefits for retention, the studies in precalculus suggest that spacing hinders immediate performance on practice activities [4], [18]. Techniques that pose challenges in the short-term but generate gains in the long-term are thought to present "desirable difficulty" [21]. The recent publication about spaced retrieval practice in calculus also found evidence for the phenomenon of desirable difficulty posed by spaced retrieval practice in mathematics [22], [23].

B. Spaced Retrieval Practice in Physics

Fewer studies have investigated the effectiveness of spaced retrieval practice in physics courses, although results have been promising. One study indicated performance in a physics course on thermodynamics benefited greatly from engagement with their app-based spaced repetition, citing nearly a full letter grade difference across three cohorts [24]. Other studies have shown that even short retrieval practice interventions can dramatically increase student performance in physics and lower performing students benefited greatly from attending lectures with distributed retrieval practice [25].

C. Feedback

The importance of feedback in spaced retrieval practice has been widely debated [26]–[33]. Reviews have indicated that spaced retrieval practice can be effective even when feedback is minimal [33]. However, other sources indicate that increased feedback is superior to lesser amounts of feedback [28]–[30], [32], [33]. Feedback allows learners to correct misconceptions and eliminate knowledge gaps. It is of importance to note that in many of the evidence-based learning practices, such as the testing effect and spaced retrieval practice, the underlying mechanism in learning from testing and retrieval is the successful retrieval of correct information [31]. If learners do not expend the effort or do not have the means to correct misconceptions, evidence-based learning practices may have limited or no benefit. Another consideration for the importance of feedback is that providing delayed feedback, as opposed to immediate feedback, may also serve as an additional retrieval interval in which students can further strengthen connections in the retrieval process [33].

D. The Current Study

The current study aims to address two major gaps in spaced retrieval practice literature. The first of which being the lack of classroom studies that test the efficacy of spaced retrieval practice. Secondly, we aim to address the sensitivity of our results to different feedback settings.

Spaced retrieval practice was implemented in introductory STEM courses: one section of *Calculus for Engineering* (fall 2020) and two sections of *Fundamentals of Physics I* (2020 and 2021). Each implementation varied slightly due to course delivery methods and assessment platforms. In the current paper, we examine the effectiveness of spaced retrieval practice in the courses. Our research questions were:

Does spaced retrieval practice benefit student learning in Calculus 2020 and Physics 2020 and 2021?

Does a change in feedback structure in Physics change the effectiveness of spaced retrieval practice?

We hypothesized that course performance on the final assessment would be significantly improved when retrieval practice was spaced as opposed to massed.

III. METHODOLOGY

A. Participants

Participants were 359 University of Louisville undergraduate students who were enrolled in one of the STEM barrier courses and who completed all practice quizzes. Calculus and Physics 2020 studies were conducted under NSF grant #1912253. The Physics 2021 study was a follow-up study. The number of participants and the demographic statistics for each course are shown in Table 1.

Figure 1. Demographic Data

Course	Course Name	Male	Female	Asian	Black	Hispanic	Non-Resident Alien	Two or More Races	White	Total
Calculus 2020	*Calculus for Engineering*	131	49	9	11	10	2	10	138	180
Physics 2020	*Fundamentals of Physics*	44	62	10	6	10	1	9	70	106
Physics 2021	*Fundamentals of Physics*									73

B. Procedures

Prior to the start of the course, professors, with assistance from the researchers, identified 24 learning objectives (LOs) to be addressed on five bi-weekly quizzes. LOs were course concepts that were held to be fundamental to a sound understanding of course material. Objectives from the first half of the course were used so that quiz questions could be distributed according to the spacing schedule. The target learning objectives comprised eight objectives from weeks 1-3, eight from weeks 4-5, and eight from weeks 6-7. Once each LO was identified, instructors created or selected four questions to assess student performance, three of which were used on bi-weekly quizzes and the fourth on a criterial end-of-semester quiz.

Prior to the start of the semester, all students enrolled in the course were randomly assigned to one of two groups for the purpose of counterbalancing the assignment of objectives to condition. To account for any potential difference in difficulty in the objective pools, two sets of quizzes were created, each with half of the learning objectives spaced. Half of the students are assigned to receive each set. The schedule for the spacing manipulation is shown in Table 2.

C. Course Format: Calculus

Calculus for Engineering was conducted using a hybrid delivery format. Lectures were recorded and delivered via an online learning platform called MyMathLab®. To support and reinforce learning, the course instructors led in-person activities on Mondays, Wednesdays, and Fridays. Due to distancing constraints during the semester of implementation, students were invited to attend only one day of activities per week. For all students, there were weekly homework assignments and biweekly quizzes, which students completed using MyMathLab®. All quizzes were online, with the criterial quiz proctored via ProctorU®. In addition, 13 unit exams were administered on Tuesdays and taken either in-person or online.

As previously described, the critical course materials were 24 target learning objectives, five quizzes, and the criterial quiz. The 24 target learning objectives were selected by the lead course instructor from a larger pool of 97 objectives that were taught in the first seven weeks of the 15-week course. All were deemed fundamental to a sound understanding of calculus. Quiz questions were drawn from the MyMathLab® question library, primarily from the Thomas' Calculus textbook [34]. One question was selected for each objective. Questions required students to use mathematical procedures to calculate an answer

Table 2. Spacing Manipulation

Condition	Content	Time administered					
		Quiz 1	Quiz 2	Quiz 3	Quiz 4	Quiz 5	Criterial Quiz
Massed	Half of the LOs from weeks 1-3	Question 1 Question 2 Question 3					Question 4
	Half of the LOs from weeks 4-5		Question 1 Question 2 Question 3				Question 4
	Half of the LOs from weeks 6-7			Question 1 Question 2 Question 3			Question 4
Spaced	Half of the LOs from weeks 1-3	Question 1	Question 2	Question 3			Question 4
	Half of the LOs from weeks 4-5		Question 1	Question 2	Question 3		Question 4
	Half of the LOs from weeks 6-7			Question 1	Question 2	Question 3	Question 4

in numerical or variable form. Only a very few questions were multiple choice. MyMathLab® generated three random algorithmic variants of the question for each student to produce a total of three quiz questions per objective. Variants differed only in their superficial features (e.g., coefficients, exponents, variable names) and correct answers. To manipulate the massing or spacing of quiz questions, students either received all three algorithmic variants of a question on a single quiz or the three variants were distributed across three consecutive quizzes.

The feedback structure for the quizzes was quite explicit. Upon completion of the assessment, students were able to view correct/incorrect item feedback. Additionally, students were able to view the correct answer had their response been incorrect. If students sought further clarification for their correct/incorrect responses, they were able to click "get answer feedback" which would then open a window with a paragraph with the rationale for the correct response.

D. *Course Format: Physics 2020*

Physics 2020 was conducted in an online-synchronous format using Blackboard©. The course was based on a system of modules designed and built by the instructor. The modules provided students with a short reading about the upcoming material, a pre-test, then a series of short (typically 8 to 12 minute) videos presenting the basic material of the course. Students could follow up each video with self-check questions to test their understanding before taking a quiz over all the material presented in the module. The course was delivered using a flipped course format, in which students completed an assigned module prior to attending the synchronous meetings. The synchronous meetings were intended to focus on the application of the material presented in the modules. In the synchronous meetings, students engaged in problem solving, often in groups, in which students were to integrate the information they obtained from the asynchronous portion of the class. The synchronous meetings were held within Collaborate Ultra, the video conferencing platform embedded within Blackboard©. Students were given weekly homework assignments as practice for the problem-solving strategies introduced in the synchronous meetings. All quizzes were online, with the criterial quiz proctored via Respondus LockDown Browser®.

The 24 target learning objectives selected for the study covered the key concepts of mechanics and were presented with a particle model framing. The questions were written by the instructor and typically asked students to recall a critical equation for a particular physical concept. The majority of the questions asked students to either use the equation to make a numerical calculation or use proportional reasoning to determine the answer. There were also a small set of questions that tested students' ability to represent appropriate relations diagrammatically by selecting the correct image from a series of options. All questions in the physics quizzes were multiple choice.

Students were able to review correct/incorrect item feedback once the assessment window closed. The feedback was limited to whether the student had answered correctly; it did not include the correct answer or the answer that the student had selected.

E. *Course Format: Physics 2021*

Physics 2021 followed the same flipped structure as Physics 2020, where students were expected to complete the assigned modules prior to each formal class meeting. However, the group sessions in Physics 2021 were delivered in-person, with students meeting in a traditional large lecture hall. The bi-weekly quizzes were still delivered through Blackboard©, and the criterial quiz was administered in-person with live proctoring. Aside from the students enrolled and the class meetings and criterial quiz being in-person rather than online, everything else was the same. The course materials were identical to Physics 2020, and the spacing schedule was identical to the Calculus and Physics 2020 implementations.

An effort was made in Physics 2021 to increase the amount of feedback presented to students. Where Physics 2020 provided limited feedback in the form of correct/incorrect item feedback, students in the Physics 2021 course were able to see all the possible answer choices with the answer they had selected and the correct answer identified.

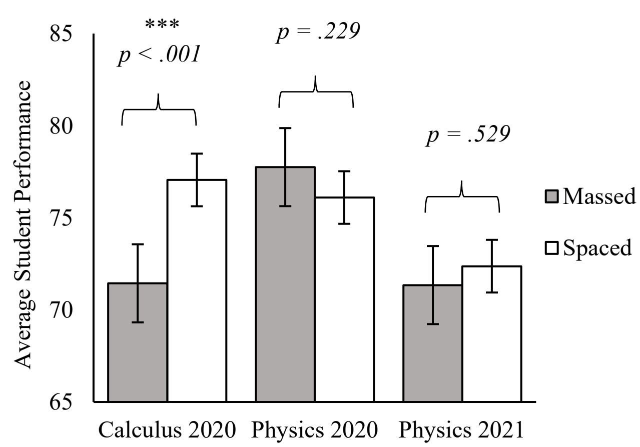
All materials from all courses are available upon request.

IV. Results

According to the Shapiro-Wilk test, the Calculus 2020 ($W = .97$, $p = .001$), Physics 2020 ($W = .92$, $p < .001$), and Physics 2021 data ($W = .95$, $p = .009$) was positively skewed, as is common in a classroom study. Although this violates the normality assumption, we report the results of *t* tests for practical interpretation. We also conducted all analyses using the nonparametric Wilcoxon Signed-Ranks test and all results were consistent with the *t* test results.

In Calculus, there was a statistically significant difference in student performance, $t(179) = 4.64$, $p < .001$. Performance was better when retrieval was spaced ($M = 77.06\%$, $SD = 17.45\%$) as opposed to massed ($M = 71.44\%$, $SD = 18.10\%$). The mean difference was 5.61% with a 95% confidence interval ranging from 3.22% to 8.00%. The Cohen's *d* statistic (.35) indicated a small-to-medium effect size.

Figure 1. Results



In Physics 2020, the difference was not significant, $t(105) = -1.21$, $p = 229$, with relatively equal performance when retrieval was spaced ($M = 76.10\%$, $SD = 18.90\%$) versus massed ($M = 77.75\%$, $SD = 17.01\%$). The mean difference was -1.65% with a confidence interval ranging from -4.36% to 1.06%. The Cohen's d statistic (.12) indicated a negligible effect size.

The spacing manipulation was also not significant in Physics 2021, $t(72) = .54$, $p = .592$. Performance in the spaced condition ($M = 72.37\%$, $SD = 17.12\%$) was relatively equal to that in the massed condition ($M = 71.35\%$, $SD = 15.77\%$), The mean difference was 1.03% with a confidence interval ranging from -2.78% to 4.83%. The Cohen's d statistic (.06) indicated a negligible effect size.

V. Discussion

The effect of spaced retrieval practice was different in Calculus versus Physics (2020 and 2021). In Calculus, spaced retrieval practice significantly benefited student performance on the criterial quiz. In Physics, there were no significant effects.

The provision of additional feedback in the second physics course (2021) was intended to determine whether a small change in the *amount of feedback* moderated the effectiveness of spaced retrieval practice. There was no effect of spaced retrieval in either Physics 2020 or 2021, indicating that the providing additional feedback was inconsequential. We therefore cannot conclude that amount of feedback is a moderating factor. Although there was a benefit of spacing in Calculus and not in Physics 2020, and the feedback structures were different, there are many other possible reasons why the effects could differ.

The first is differences in the *type of knowledge* required in these courses. Much of the spaced retrieval practice literature is based on verbal recall. Fixed declarative knowledge requires only the retrieval of a piece of information from memory whereas answering calculus and physics questions involves more complex (i.e., multi-step) problem-solving work. In calculus, solving quiz questions requires students to know when and how to apply various procedures. Students are likely building and using procedural knowledge during the practice quizzes. In physics, students must remember larger principles and also translate word problems into visuospatial diagrams and then into formulas for computations. These students are likely engaging in conceptual knowledge and spatial transformations during problem-solving. Although a type-of-knowledge boundary condition is possible (spaced retrieval could benefit memory of declarative verbal and procedural mathematical knowledge but not visuospatial knowledge), it has not been studied directly. In fact, one piece of evidence indicates the counterpoint; a benefit of spaced retrieval was observed in a physics course [24]. It is therefore more likely that other variables that also differ between the courses are better explanations for the differences in effects.

A more likely moderating factor is the *amount of intrinsic spacing* in each course. In Physics (2020 and 2021), the core concepts of mechanics are covered in the first eight weeks using a point particle model. In the following month, those same concepts were revisited using an extended object model. The physics course is designed to return to previous material, and therefore intrinsically includes spacing. There may have been no benefit from spacing in this physics course because retrieval in the course itself is already spaced. Although the calculus course did include some amount of intrinsic spacing (as most calculus courses do), it is arguably less than that in this physics course. Our study did not attempt to quantify the amount of intrinsic spacing in each course, but it could be a moderating factor of the effectiveness of spaced retrieval practice.

Another possible factor for consideration is the *integration* of quizzes into the existing courses, and the resulting emphasis on the practice quizzes and cumulative quiz. In Calculus, there were many weekly exams, and the biweekly quizzes were in the same format as the homework (in the online platform). In Physics (2020 and 2021), there were only 4 exams, and the biweekly quizzes may have been treated like assessments rather than like homework. This could have made students focus more on raising their performance on these quizzes, even though they were the same overall percentage of their grades in each course. Focusing on the quizzes might mean that students studied more, raising their performance on objectives in the massed condition, thereby masking any potential effect of spacing.

A. Limitations

A potential moderating factor that we did not control was the *student sample.* Calculus students were in the engineering school, whereas most Physics students were in Arts & Sciences with a premedical major, and there were also differences in demographics (see Table 1). It is possible that these student groups study or learn differently. We did not study this; a controlled experiment would have students from both programs learning both subjects. However, it could be that spaced retrieval practice is more effective for students who study either infrequently or less effectively because the manipulation is a study aid.

Secondly, we must acknowledge the limitations of validity of these results. Validity limitations include attrition (analyses of only complete datasets), instrumentation (Physics questions were newly developed for this study), and replicability (each study was done only once). Results should be interpreted accordingly. However, the within-subjects design with counterbalancing avoided other threats to validity, and the results can add meaningfully to our existing knowledge about spaced retrieval practice.

VI. Conclusions & Future directions

This paper discussed the implementation of spaced retrieval practice in STEM courses with different results. Spaced retrieval benefitted students in calculus but not in physics, regardless of feedback structure. Possible moderating factors include amount of feedback available, type of knowledge required, intrinsic spacing in the courses, integration of retrieval opportunities into the courses, and sample differences. Future work is aimed at examining these potential moderating factors in more detail.

VII. References

[1] A. Redmond-Sanogo, J. Angle, and E. Davis, "Kinks in the STEM pipeline: Tracking STEM graduation rates using science and mathematics performance," *School Science and Mathematics*, vol. 116, no. 7, pp. 378–388, Nov. 2016, doi: 10.1111/ssm.12195.

[2] R. Noonan, "STEM Jobs: 2017 Update," US Department of Commerce, 2017. [Online]. Available:

https://www.commerce.gov/sites/default/files/migrated/reports/stem-jobs-2017-update.pdf

[3] 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, Sep. 2012. doi: 10.1007/s10648-012-9205-z.

[4] K. B. Lyle, C. R. Bego, R. F. Hopkins, J. L. Hieb, and P. A. S. Ralston, "How the amount and spacing of retrieval practice affect the short- and long-term retention of mathematics knowledge," *Educational Psychology Review*, vol. 32, no. 1, pp. 277–295, Mar. 2020, doi: 10.1007/s10648-019-09489-x.

[5] 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, Nov. 2014, doi: 10.1037/a0037559.

[6] K. B. Mcdermott, "Practicing retrieval facilitates learning," *Annual Review of Psychology*, 2020, doi: 10.1146/annurev-psych-010419.

[7] J. Dunlosky and R. Ariel, "Self-regulated learning and the allocation of study time," *Psychology of Learning and Motivation - Advances in Research and Theory*, vol. 54, pp. 103–140, Jan. 2011, doi: 10.1016/B978-0-12-385527-5.00004-8.

[8] M. A. McDaniel, J. L. Anderson, M. H. Derbish, and N. Morrisette, "Testing the testing effect in the classroom," *European Journal of Cognitive Psychology*, vol. 19, no. 4–5, pp. 494–513, Jul. 2007, doi: 10.1080/09541440701326154.

[9] K. B. Lyle and N. A. Crawford, "Retrieving essential material at the end of lectures improves performance on statistics exams," *Teaching of Psychology*, vol. 38, no. 2. pp. 94–97, 2011. doi: 10.1177/0098628311401587.

[10] F. C. Leeming, "The exam-a-day procedure improves performance in psychology classes," *Teaching of Psychology*, vol. 29, no. 3, pp. 210–212, 2002, doi: 10.1207/S15328023TOP2903_06.

[11] 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, p. 354, 2006.

[12] N. Kornell, "Optimising learning using flashcards: Spacing is more effective than cramming," *Applied Cognitive Psychology*, vol. 23, no. 9, pp. 1297–1317, 2009, doi: 10.1002/acp.1537.

[13] S. K. Carpenter, "Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 35, no. 6, p. 1563, 2009.

[14] D. Rohrer, R. F. Dedrick, and M. K. Hartwig, "The scarcity of interleaved practice in mathematics textbooks," *Educational Psychology Review*, vol. 32, no. 3. Springer, pp. 873–883, Sep. 01, 2020. doi: 10.1007/s10648-020-09516-2.

[15] G. B. Maddox and D. A. Balota, "Self control of when and how much to test face-name pairs in a novel spaced retrieval paradigm: An examination of age-related differences," *Aging, Neuropsychology, and Cognition*, vol. 19, no. 5, pp. 620–643, Sep. 2012, doi: 10.1080/13825585.2011.640658.

[16] J. M. Logan and D. A. Balota, "Expanded vs. equal interval spaced retrieval practice: Exploring different schedules of spacing and retention interval in younger and older adults," *Aging, Neuropsychology, and Cognition*, vol. 15, no. 3, pp. 257–280, May 2008, doi: 10.1080/13825580701322171.

[17] A. C. Butler, E. J. Marsh, J. P. Slavinsky, and R. G. Baraniuk, "integrating cognitive science and technology improves learning in a stem classroom," *Educational Psychology Review*, vol. 26, no. 2, pp. 331–340, 2014, doi: 10.1007/s10648-014-9256-4.

[18] 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," *Educational Psychology Review*, vol. 28, no. 4, pp. 853–873, Dec. 2016, doi: 10.1007/s10648-015-9349-8.

[19] B. M. May, "Effects of spaced, repeated retrieval practice and test-potentiated learning on mathematical knowledge and reasoning," *International Journal of Mathematical Education in Science and Technology*, 2021, doi: 10.1080/0020739X.2021.1961034.

[20] 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," *Applied Cognitive Psychology*, vol. 33, no. 5, pp. 732–743, Sep. 2019, doi: 10.1002/acp.3507.

[21] E. L. Bjork and R. A. Bjork, "Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning," *Psychology and the real world: Essays illustrating fundamental contributions to society*, vol. 2, no. 59–68, 2011.

[22] C. R. Bego, P. A. Ralston, K. B. Lyle, and J. Immekus, "Introducing desirable difficulty in engineering mathematics with spaced retrieval practice," 2021.

[23] K. B. Lyle, C. R. Bego, P. S. Ralston, and J. C. Immekus, "Spaced retrieval practice imposes desirable difficulty in calculus learning," *Educational Psychology Review*, vol. in-press, 2022.

[24] A. Voice and A. Stirton, "Spaced repetition: towards more effective learning in STEM," *New Directions in the Teaching of Physical Sciences*, no. 15, Jan. 2020, doi: 10.29311/ndtps.v0i15.3376.

[25] V. Gjerde, B. Holst, and S. D. Kolstø, "Retrieval practice of a hierarchical principle structure in university introductory physics: Making stronger students," *Physical Review Physics Education Research*, vol. 16, no. 1, Jan. 2020, doi: 10.1103/PHYSREVPHYSEDUCRES.16.013103.

[26] E. R. Fyfe and B. Rittle-Johnson, "Mathematics practice without feedback: A desirable difficulty in a classroom setting," *Instructional Science*, vol. 45, no. 2, pp. 177–194, 2017.

[27] J. Hattie and M. Gan, "Instruction based on feedback," in *Handbook of research on learning and instruction*, Routledge, 2011, pp. 263–285.

[28] R. L. Bangert-Drowns, C.-L. C. Kulik, J. A. Kulik, and M. Morgan, "The instructional effect of feedback in test-like events," *Rev Educ Res*, vol. 61, no. 2, pp. 213–238, 1991.

[29] A. C. Butler, J. D. Karpicke, and H. L. Roediger III, "Correcting a metacognitive error: feedback increases retention of low-confidence correct responses.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 34, no. 4, p. 918, 2008.

[30] H. Pashler, N. J. Cepeda, J. T. Wixted, and D. Rohrer, "When does feedback facilitate learning of words?," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 31, no. 1, p. 3, 2005.

[31] S. H. K. Kang, K. B. McDermott, and H. L. Roediger III, "Test format and corrective feedback modify the effect of testing on long-term retention," *European Journal of Cognitive Psychology*, vol. 19, no. 4–5, pp. 528–558, 2007.

[32] R. W. Kulhavy and W. A. Stock, "Feedback in written instruction: The place of response certitude," *Educ Psychol Rev*, vol. 1, no. 4, pp. 279–308, 1989.

[33] H. L. Roediger and A. C. Butler, "The critical role of retrieval practice in long-term retention," *Trends in Cognitive Sciences*, vol. 15, no. 1. pp. 20–27, Jan. 2011. doi: 10.1016/j.tics.2010.09.003.

[34] J. R. Hass, C. E. Heil, and M. C. Weir, *Thomas' calculus*, 14th ed. Pearson Education, 2017.