Research to Practice to Research:

Intrinsic requirements of implementing and studying spaced retrieval practice in STEM courses

Campbell R. Bego
Engineering Fundamentals & Psychological and Brain Sciences
J.B. Speed School of Engineering
University of Louisville
Louisville, KY
campbell.bego@louisville.edu

Patricia A. S. Ralston
Engineering Fundamentals
J.B. Speed School of Engineering
University of Louisville
Louisville, KY
patricia.ralston@louisville.edu

Keith B. Lyle
Psychological and Brain Sciences
College of Arts & Sciences
University of Louisville
Louisville, KY
keith.lyle@louisville.edu

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

***Abstract*— This Full-Length, Research-to-Practice paper discusses intrinsic requirements that may challenge instructors if they attempt to implement spaced retrieval practice in their courses. During the first year of National Science Foundation (NSF) grant #1912253, project leaders led nine Science, Technology, Engineering, and Mathematics (STEM) instructors through a series of five interactive workshops to develop learning objectives and quiz questions. Most of the STEM instructors had to redefine their existing, multifaceted learning objectives into more specific objectives with an appropriately fine grain-size for the practice. They also worked to develop multiple questions that test the same objective with (1) comparable difficulty and (2) similar cognitive processes. Project leaders noticed that a within-subjects, counterbalanced study design presents additional challenges to implementation. Instructors were able to work around difficulties during break-out sessions in workshops and in one-on-one sessions, especially when given examples from their own discipline in one-on-one feedback.**

In this paper, we first describe the current state of spaced retrieval practice research and the purpose and plan of our active NSF grant. We then detail the implementation requirements we have discovered. Lastly, we summarize our findings with bullet-point, STEM-practitioner-centered statements about implementing spaced retrieval practice in the classroom. Identifying potential challenges of implementation and solutions to these challenges is an important step in getting the powerful memory tool of spaced retrieval practice into the STEM classroom.

Keywords — spaced retrieval practice, undergraduate STEM instruction, implementation of evidence-based practices, research to practice.

I. Spaced Retrieval Practice

Derived from memory research conducted by cognitive psychologists, spaced retrieval practice is an instructional technique in which instructors ask students multiple questions about a topic over time with intervening delays [1]. This practice is fundamentally different than many traditional Science, Technology, Engineering, and Mathematics (STEM) course structures in which content is separated into discrete units or chapters and only revisited on a cumulative final exam. Spaced retrieval practice is an effective instructional technique for two reasons. First, the act of retrieving information from memory, as opposed to restudying it, increases long-term retention of the information (the retrieval practice effect, e.g., [2]–[5]). Second, the memory-enhancing power of multiple retrievals is greater when they are spaced out over time versus when they all occur in a relatively short temporal window (the spacing effect, e.g., [1], [6]). Retention of information from introductory STEM courses is critical for students to be successful in STEM majors; thus, implementing spaced retrieval practice has the potential to improve persistence in STEM degrees and careers.

To implement spaced retrieval practice in the classroom, instructors must pose questions that require students to recall prior topics in the course. It is possible to implement spaced retrieval practice in the classroom by starting or ending class with a few questions about prior topics. It is also possible to implement the practice outside of class through short quizzes, or through homework and project-based assignments, that cover prior topics. A systematic approach to applying this practice requires:

1. Identifying the content to be spaced,
2. Writing multiple questions about the content, and
3. Determining when and how to ask the questions.

While these steps may appear easy and logical, each requires careful decision-making by instructors. Which content is the most important for students to learn and practice? What content is already spaced in the course? How many questions are available in textbooks or online systems, and which ones are the best for this practice? Should class time be used, or should students practice retrieval outside of class? After a topic is introduced, at what delay should the questions be asked? These questions and others must be considered in the implementation of spaced retrieval practice.

If instructors would like to study the impact of their implementation, yet more questions arise. The study design needs to be planned, the materials designed with tighter control, and the implementation requires more careful thought.

II. Classroom Research & Current Project

Recent collaborations between learning scientists and STEM faculty have revealed significant short- and long-term benefits of spaced retrieval practice in STEM classrooms (e.g., [7]–[9]). However, the number of classroom studies is limited and more are needed. It remains unknown whether spaced retrieval practice is effective across all STEM domains. It is also unknown whether the learning benefits are stronger in some domains than others. To assess the utility of an educational intervention, one must consider the intervention's effectiveness across domains and populations [10].

Our active NSF grant (#1912253) addresses this gap in the literature by implementing and comparing the effectiveness of spaced retrieval practice in multiple STEM domains. The project includes large courses at the University of Louisville in biology, chemistry, physics, mathematics, psychology, and engineering that are keys to various types of STEM degrees. The faculty members who are participating in this study are seasoned instructors with both discipline-specific technical expertise and interest in educational research and evidence-based educational strategies. Our proposed plan of work consists of three, year-long project phases: (1) development, (2) implementation and analysis, and (3) documentation and dissemination.

In the first year of the grant, project leaders conducted five workshops to educate STEM faculty about spaced retrieval practice and support their development of materials for classroom implementation. Rich discussions arose between faculty participants, PIs, and research staff that revealed unexpected requirements for implementing spaced retrieval practice in the STEM classroom. In the following sections, we discuss the requirements grouped by the implementation steps. We also describe innovative solutions to these challenges.

III. Content Selection

The first step in developing a spaced-retrieval-practice implementation is selecting the content for spacing. Spacing requires time between content presentation and retrieval. The earlier content is delivered, the more time is available to practice spaced retrieval. Contrastingly, the later content is delivered, the less time is available to practice. Therefore, an intrinsic limitation is that spacing can only be applied to a portion of the content in a course. **It is not possible to apply spacing to content that is introduced at the end of the semester.** In addition, as the semester advances, more content is available for spaced retrieval practice. Answering questions about prior topics takes time, which reduces time for other course activities, such as posing questions on current topics or other in-class activities. If instructors were to practice spaced retrieval practice for all content, the number of questions would increase dramatically as the semester progresses. **Instructors must be selective about the topics they want to space.** For example, in general chemistry, instructors often require that students know the names and symbols for all 118 (as of April 2020) elements. This is often taught and expected to be learned in week 1 of an undergraduate chemistry course. Using spaced retrieval practice is a good way to ensure that students remember the periodic table for a long time, but time spent on recalling all the elements would reduce the time available for practicing more advanced content.

Perhaps the greatest challenge of content selection in STEM courses is defining learning objectives that are of sufficiently fine grain-size to guide classroom instruction and assessment. Undergraduate STEM courses are often organized around broad learning objectives, using verbs such as "solve" and "analyze." To demonstrate mastery within this kind of objective, multiple pieces of information and procedures are needed. However, to gain the benefits of spaced retrieval, students need to retrieve *the same information* for each question within a given learning objective. Information can include a fact, a procedure, a categorization, a diagrammatic representation, or other types of knowledge. Because many STEM learning objectives require students to recall and combine multiple processes and pieces of information, it is important to articulate precisely what information is important. To help build materials for implementation, **learning objectives should be narrowly defined such that they target a single concept or process**. There are many good references about writing specific and measurable learning objectives [11]. Even with best practices for writing learning objectives, STEM instructors may find it difficult to break down their content into such a fine grain size.

For example, in introductory physics, a typical learning objective is: "Solve a Newton's Second Law problem." Multiple cognitive processes are required in this learning objective, including memory, reasoning, mental imagery, categorization, modeling, and problem solving. This objective can be broken down into several pieces that are more appropriate for spaced retrieval practice, as follows:

- Identify that Newton's Second Law applies in a given word problem.
- Recall Newton's Second Law equation and meaning.
- Draw a Free-Body Diagram representation of the problem.
- Write mathematical formulas to solve the problem.

These specific objectives about singular cognitive actions can be practiced, and thus they lend themselves to spaced retrieval practice.

IV. Question Development

Spaced retrieval practice depends on asking multiple questions on a given topic over time. The additional practical challenge is that **each question for a given topic must require students to recall the <u>same</u> information**. Writing narrow learning objectives helps instructors target the same information with different questions. Questions are most frequently used in summative assessment to measure whether students know a piece of information. In this way, it seems redundant to ask about the same information multiple times. However, asking the same question or very similar questions is the very essence of retrieval practice. Developing similar questions is more or less straightforward depending on the course content, as well as the resources available. In mathematics, testing students' ability to apply solution methods can be repeated easily by replacing

This project was funded by the National Science Foundation (NSF) through the program Improving Undergraduate STEM Education: Education and Human Resources (IUSE: EHR) program.

coefficients, variable names, and other values in a problem. There are many online textbooks for mathematics that have large test banks of questions already categorized by learning objective, topic, and chapter. Utilizing such a test bank with a large number of questions per objective can make the development of spaced retrieval materials easier. Not all test banks are alike, however. Biology and chemistry test bank questions are likely to vary in the information that is retrieved. For example, questions might ask for retrieval of only a few element names and symbols, or different processes within a cycle. This would not be appropriate for spaced retrieval practice.

Writing questions that require similar information to be recalled can be particularly difficult with learning objectives about category recognition and application. Based on our work with the nine STEM faculty members, these types of objectives occur frequently throughout STEM domains. In chemical engineering thermodynamics, for example, it is important for students to understand many properties of materials within different locations on a P-T diagram (Fig. 1). Many material properties coincide with locations on the diagram, either within a region, along a line, or at an intersection point. If questions each ask about only one property, phase, or location in the diagram, then the question set does not target the memory of the diagram. However, if questions directly ask about locations on the diagram, then students are practicing retrieval of the information in the diagram. **Question sets are appropriate if they are designed such that knowing the answers to early questions leads towards better performance on the final question.** Appropriate and inappropriate sets of questions for the chemical engineering example are illustrated in Table 1. As illustrated in these examples, one method of getting the categorization questions to be similar is to have the same multiple choice answers across the categories.

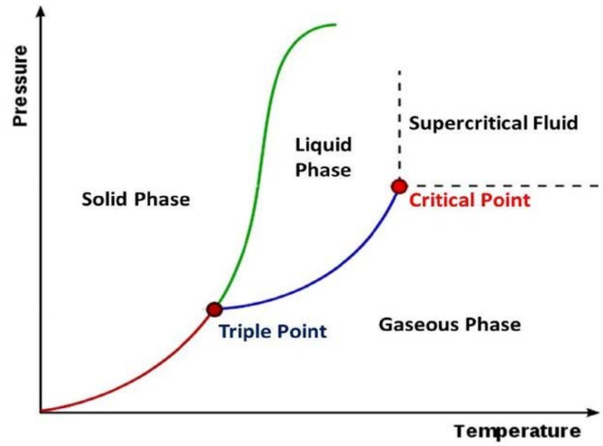
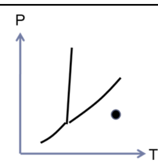
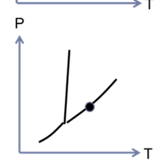
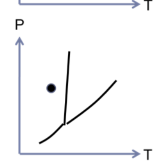
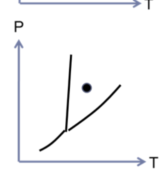


Fig. 1. An example P-T diagram taught and used in Chemical Engineering Thermodynamics. Important learning objectives include understanding the properties of materials in each phase, at each point, and along each line in the diagram. Image from *[12]*.

Question difficulty and time-to-answer also need to be considered. If instructors decide to space 20 objectives, the 20 corresponding questions could take 20 minutes or could take several hours, depending on question difficulty and other factors. In undergraduate STEM courses, it is likely that many

TABLE I. QUESTION SET COMPARISON FROM CHEMICAL ENGINEERING THERMODYNAMICS

This table shows two sets of questions for a chemical engineering thermodynamics learning objective. The objective is to *Understand how a P-T diagram represents material phases*. The questions on the left successfully get students to practice retrieving the same information each time (the locations of material phases in the diagram with respect to the lines), whereas the questions on the right target different pieces of information (the shape of the diagram, discrete locations, material phases).

Appropriate for Spaced Retrieval Practice		Not Appropriate
1. Identify the equilibrium phase(s) at the indicated point. (gas, liquid, solid, gas and liquid, solid and liquid).		1. In a P-T phase diagram the line separating solid from liquid is nearly (horizontal, vertical).
2. Identify the equilibrium phase(s) at the indicated point. (gas, liquid, solid, gas and liquid, solid and liquid).		2. In a P-T phase diagram the region at high temperature and low pressure is (liquid, gas).
3. Identify the equilibrium phase(s) at the indicated point. (gas, liquid, solid, gas and liquid, solid and liquid).		3. In a P-T phase diagram the region at high pressure and very low temperature is (liquid, solid).
4. Identify the equilibrium phase(s) at the indicated point. (gas, liquid, solid, gas and liquid, solid and liquid).		4. In a P-T phase diagram gas and solid phases can exist in equilibrium (below, above) the triple point.

existing questions require a significant amount of time for students to answer. **The best questions are those that get students to practice retrieval of information without spending too much time.** Although some test banks may include average time spent per question (e.g., Pearson's MyMathLab), this information is rarely available for textbooks and instructor-generated questions. It is therefore something that instructors must keep in mind as they are developing materials. Multiple choice questions, which may take less time to answer, may be useful in this regard, although they can be time consuming for instructors to create.

V. Syllabus Integration

In addition to challenges in content identification and question development, it can also be difficult for instructors to figure out how to incorporate spaced retrieval practice in their classrooms, either during class time or outside of class. In implementing spaced retrieval practice, **it is likely that instructors will need to alter other assignments.** An easy way to implement spaced retrieval is by adding questions on prior topics to existing homework assignments or quizzes. There is innately more time available outside of class, but there is always a tradeoff. **It takes time to practice retrieval**, just like it takes time to lecture, lead an activity, and do homework assignments. For courses in which multiple practice opportunities already exist, it will only require rearranging the timing of the questions. For other courses that do not require students to practice the same information multiple times already, integrating spaced retrieval into the course may require more work to be done by the instructors.

If the implementation is merely asking questions at a different time than usual, instructors can decide whether or not to be explicit about spaced retrieval in the syllabus. Stating that homework or quiz assignments are "cumulative" is a good way to normalize asking questions about prior topics. Alternatively, an explicit description could also help students understand that the topics they see repeated are important and the practice they are being given will help their long-term memory. This can be motivating for students, making them more likely to participate.

Instructors also need to choose how much time they want to put between retrievals. Research indicates that delays on the order of days and weeks increase retention after durations 5-10 times as long as the initial delay [13]. For example, a spacing interval of one week will benefit retrieval 5-10 weeks following the final retrieval act. For a semester-long undergraduate course, one week would be very appropriate.

One practical example comes from our current project. In this grant, we streamlined the syllabus integration to make implementation methods similar across the STEM courses. We also worked to design an implementation method that would require minimal course redesign on the part of participating faculty. We added five ancillary quizzes to be taken online, outside of class, and a final quiz to be taken during the class. The quizzes occur approximately every two weeks throughout the semester, administered via online learning management systems. Some instructors were able to add these quizzes to their syllabi without removing any other assignments, while others found it difficult to justify adding more content at home and had to remove other assignments.

Another decision-making point comes when assigning a value to the completion of spaced retrieval practice. It is important to assign a high enough value such that students will participate in the retrieval practice with their best efforts. It is also recommended that the value not be a major factor in students' overall grade. We have found that **5-10% is a good value to be placed on spaced retrieval practice**. For research, it is also necessary in a study to give incentives for students to complete all questions or quizzes. A sample syllabus integration is given for a psychology statistics course in Figure 2.

VI. Research to Practice to Research

Studying the effect of spaced retrieval practice imposes its own requirements on implementation. Instructors need to consider potential differences in students, learning objectives, and overall course design. Many research designs are possible, including quasi-experimental, concurrent between-subjects, and randomized, and research designs with tighter controls increase the implementation complexity. Our current study design is within-subjects and counterbalanced. Within-subjects means that every student experiences content in two conditions: massed and spaced. At the end, each student will contribute two data points from the final quiz: performance on massed content and performance on spaced content. It will then be possible to test whether performance on spaced content is significantly better than performance on massed content. Assignment of objectives to quizzing condition (massed or spaced) is counterbalanced such that every objective appears in the massed condition for half of the students and in the spaced condition for the other students. In this way, we control for individual differences between students and variability in objective difficulty.

In a within-subjects counterbalanced research design, **multiple questions must be appropriate when asked at the**

Blackboard Quizzes and Pre-Final Exam Quiz: Five quizzes will be administered on Blackboard on the dates shown on the course schedule. They will be cumulative. Any content covered in the class up to the time of the quiz is fair game. Each quiz will become available at 1pm on a Friday and be due by 11:59pm on the following Sunday. Each quiz must be completed in a single session and within a specified amount of time, which shall not exceed 75 minutes. I strongly encourage you to study for these quizzes before you start them and attempt to complete them without reference to your notes or other supporting material. This will simulate exam conditions and reveal to you what you do and do not know. A similar quiz will be administered in class, using scantrons, on Tuesday, December 1.

Your percentage correct on the six quizzes will be averaged together and, collectively, will count for 6% of your final course grade. It is very important that you complete all six quizzes. If you complete all six, your quiz average will be increased by 10% for the purposes of calculating your final course grade. For example, if your average on the six quizzes is 85%, it will be treated as 95% when calculating your final grade in the class. But, remember, you will receive this bonus only if you complete all six quizzes at the times specified on the course schedule. **Multiply your average quiz score (including the 10% bonus, if applicable) by .06 to figure out how much of your final grade will come from these quizzes.**

Fig. 2. An example of a syllabus integration for an experimental psychology course. In this case, the manipulation is being studied and so is not mentioned explicitly.

same time (massed) and when asked at multiple different time points (spaced). Therefore, questions should not be exact replications, because seeing the same question multiple times in quick succession in the massed condition would be strange to students. Contrastingly, asking the same question more than once is actually a good way to easily implement spaced retrieval if not running this kind of study. Also, when conducting a study, **questions should not increase in difficulty and depth as the semester progresses**, because some students receive them all at once at an early point in the semester.

VII. Summary of Implementation Requirements

As described in detail above, implementation of spaced retrieval practice requires instructors to select content to space, write questions, and figure out how to incorporate it into their courses. Throughout our developmental year, we found one-on-one discussions between learning scientists and instructors to be invaluable. These discussions were most meaningful in the middle of the development of the implementation as opposed to at the beginning of the process. It is only when instructors actually begin to write questions that the limitations and requirements become clear. Hopefully, this article will assist instructors who wish to attempt this practice. Best practices are summarized as follows:

- **Selecting Content**
  - Instructors must be selective about the topics they want to space because there is not enough time to space every topic.
  - Learning objectives should be narrowly defined such that they target a single item, concept, or procedure.
  - Spacing cannot be applied to content introduced at the end of a course.
- **Writing Questions**
  - Each question for a given topic must require students to recall the same information.
  - Question sets are appropriate if they are designed such that knowing the answers to early questions leads towards better performance on the criterial question.
  - The best questions are those that get students to practice retrieval of information or processes without spending too much time.
- **Syllabus Integration**
  - It takes time to practice retrieval.
  - It is likely that instructors will need to alter other assignments.
  - 5-10% is a good value to be placed on spaced retrieval practice, and an incentive to complete all assignments is recommended.
- **In Research**
  - Questions should not increase in difficulty and depth as the semester progresses.
  - Questions must be appropriate whether asked at the same time or across multiple time points.

Because spaced retrieval practice has been found to improve undergraduate student recall in the short-term (i.e., within a semester) and the long-term (i.e., in the following semester), application to STEM barrier courses could improve student success, potentially thereby increasing the number of STEM graduates. Identifying and describing implementation challenges will help open more classrooms to this evidence-based technique.

Acknowledgments

We are grateful for the invaluable feedback and willingness to collaborate from our faculty participants Raymond Chastain, Physics, Lora Haynes, Psychology, Lenore Hoyt, Chemistry, Rachel Pigg and Shira Rabin, Biology; Matthew Scobee, Mathematics, and Thomas Starr, Chemical Engineering.

References

[1] 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.

[2] J. D. Karpicke and H. L. Roediger, “The critical importance of retrieval for learning,” *Science*, vol. 319, no. 5865, pp. 966–968, 2008, doi: 10.1126/science.1152408.

[3] H. L. Roediger and J. D. Karpicke, “The power of testing memory: Basic research and implications for educational practice,” *Perspect. Psychol. Sci.*, vol. 1, no. 3, pp. 181–210, 2006, doi: 10.1111/j.1745-6916.2006.00012.x.

[4] H. L. Roediger and M. A. Pyc, “Inexpensive techniques to improve education: Applying cognitive psychology to enhance educational practice,” *J. Appl. Res. Mem. Cogn.*, vol. 1, no. 4, pp. 242–248, 2012.

[5] H. L. Roediger and J. D. Karpicke, “Test-enhanced learning: Taking memory tests improves long-term retention,” *Psychol. Sci.*, vol. 17, no. 3, pp. 249–255, 2006, doi: 10.1111/j.1467-9280.2006.01693.x.

[6] N. Kornell, “Optimising learning using flashcards: Spacing is more effective than cramming,” *Appl. Cogn. Psychol.*, vol. 23, no. 9, pp. 1297–1317, 2009, doi: 10.1002/acp.1537.

[7] 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.*, vol. 2019, 2019, doi: 10.1007/s10648-019-09489-x.

[8] 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.

[9] A. C. Butler, E. J. Marsh, J. P. Slavinsky, and R. G. Baraniuk, “Integrating cognitive science and technology improves learning in a STEM classroom,” *Educ. Psychol. Rev.*, vol. 26, no. 2, pp. 331–340, 2014, doi: 10.1007/s10648-014-9256-4.

[10] J. Dunlosky, K. A. Rawson, E. J. Marsh, M. J. Nathan, and D. T. Willingham, “Improving students’ learning with effective learning techniques: Promising directions from cognitive and educational psychology,” *Psychol. Sci. Public Interes.*, vol. 14, no. 1, pp. 4–58, 2013, doi: 10.1177/1529100612453266.

[11] N. E. Gronlund, *How to write and use instructional objectives*. Macmillan Pub Co; 5th edition, 1995.

[12] R. Gupta, A. K. Mishra, and A. K. Pathak, “Supercritical Fluid technology a boon for pharmaceutical Particle Manufacturing,” in *Science & Technology For Human Development*, 2014, Poster.

[13] N. J. Cepeda, N. Coburn, D. Rohrer, J. T. Wixted, M. C. Mozer, and H. Pashler, “Optimizing distributed practice theoretical analysis and practical implications,” *Exp. Psychol.*, vol. 56, no. 4, pp. 236–246, 2009, doi: 10.1027/1618-3169.56.4.236.