

RESEARCH ARTICLE

WILEY

Regaining access to marginal knowledge in a classroom setting

Andrew C. Butler¹  | Allison C. Black-Maier² | Kathryn Campbell³ | Elizabeth J. Marsh⁴ | Adam M. Persky³

¹Department of Education and Department of Psychological & Brain Sciences, Washington University in St. Louis, St. Louis, Missouri

²Friday Institute for Educational Innovation, North Carolina State University, Raleigh, North Carolina

³Division of Pharmacotherapy and Experimental Therapeutics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

⁴Department of Psychology and Neuroscience, Duke University, Durham, North Carolina

Correspondence

Andrew C. Butler, Department of Education, Washington University in St. Louis, Box 1183, St. Louis, MO 63130-4899.
Email: andrew.butler@wustl.edu

Funding information

While working on this project, ACB was supported by the James S. McDonnell Foundation 21st Century Science Initiative in Understanding Human Cognition – Collaborative Grant No. 220020483.

Summary

Students learn large amounts of information, but not all of it is remembered after courses end – meaning that valuable class time is often spent reviewing background material. Crucially, laboratory research suggests different strategies will be effective when reactivating previously learned information (i.e. *marginal knowledge*), as opposed to learning new information. In two experiments, we evaluated whether these laboratory results translated to the classroom. Topics from prior courses were tested to document which information students could no longer retrieve. Half were assigned to a not-tested control and half to the intervention; for these topics, students answered multiple-choice questions (without feedback) that gave them the chance to recognize the information they had failed to retrieve. Weeks later, students completed a final assessment on all topics. Crucially, multiple-choice testing increased the retrieval of previously forgotten information, providing the first classroom demonstration of the reactivation of marginal knowledge.

KEYWORDS

long-term retention, marginal knowledge, multiple-choice testing, remediation

1 | INTRODUCTION

Students must acquire a vast amount of knowledge, especially in professional degree programs. However, successful learning in a course does not guarantee that this information will be remembered over the long term; students rapidly forget content if they do not continue to use it (e.g., summer learning loss; Cooper, Nye, Charlton, Lindsay, & Greathouse, 1996; Custers, 2010; Custers & Ten Cate, 2011; Ling, Swanson, Holtzman, & Bucak, 2008; Sullivan, Gregg, Adams, Rodgers, & Hull, 2013). This problem occurs even if material is well-learned initially (Emke, Butler, & Larsen, 2016; Larsen, Butler, & Roediger III, 2009) and is particularly problematic in courses where new learning builds on concepts and skills learned in prior courses.

Although this problem is often characterized as student “forgetting,” basic memory theory highlights that forgetting is not all-or-none. That is, information may be *available* (stored in memory) but not *accessible* (retrievable), with the consequence that new retrieval cues (different locations, moods, physiological states, etc.) may cue (reactivate) a seemingly forgotten memory (Tulving &

Pearlstone, 1966). Such findings provided the empirical base for theoretical ideas such as encoding specificity (Tulving & Thomson, 1973) and transfer appropriate processing (Morris, Bransford, & Franks, 1977). This distinction is not limited to episodic memories; for example, tip-of-the-tongue states are frustrating precisely because one struggles to retrieve a word or phrase that one is sure one knows (Brown & McNeil, 1966). Bahrck and colleagues (Berger, Hall, & Bahrck, 1999) coined the term *marginal knowledge* to describe cases where knowledge is stored (available) but not retrievable (accessible).

Our work follows directly from the experimental literature on marginal knowledge, so we briefly review those studies here. The laboratory work is compelling because it utilizes a control that is not possible to use in classroom studies, namely, fictional questions (e.g., *Who was Captain Child's enemy?*) for which there cannot be any marginal knowledge to reactivate. Each fictional question is paired with a real question (e.g., *Who was Robin Hood's greatest enemy?*), with the logic that an intervention that successfully reactivates marginal knowledge should improve learning more for (forgotten) real information than fictitious items (as only real items are associated with marginal

knowledge to reactivate). In the seminal study, participants took an initial short-answer test consisting of real and fictitious questions, and then briefly saw the answers (one alternative was randomly designated as “correct” for each fictitious item). Later, after varying delays, participants took the short-answer test again. The focus is on questions that participants failed to answer on the initial test. Critically, participants benefitted much more from exposure to answers of real items than fictitious ones (see also Cantor, Eslick, Marsh, Bjork, & Bjork, 2015), and that benefit lasted up to nine days for real items (whereas answers to fictitious questions were quickly forgotten). The fictional items were learnable, but they were also quickly forgotten as there was no prior knowledge to reactivate, highlighting the difference between new learning and the reactivation of previously learned information.

Past research suggests that multiple-choice tests are sometimes sufficient to reactivate marginal knowledge, given that such questions reexpose the learner to the correct answers and provide retrieval practice. In one study, following an initial short answer test, real and fake test items (from the original Berger et al., 1999 materials) were assigned to one of three conditions: multiple-choice testing without feedback, multiple-choice testing with feedback, or a no-test control condition (Cantor et al., 2015). Two findings are important for present purposes. First, multiple-choice testing improved performance on the final test 2 days later – but only for real items, for which participants had information stored in memory to reactivate. Second, the benefits of receiving feedback were similar for real and fake items, suggesting that that participants learned new information from the feedback rather than needing it to reactivate stored knowledge. Additional studies compared multiple-choice testing to exposure to the answers; less clear was whether multiple-choice testing provided a benefit over simply reading an answer (Cantor et al., 2015). This result is surprising given the overwhelming evidence that retrieval practice promotes learning (Roediger III & Butler, 2011) – but also highlights that reactivating marginal knowledge may differ from new learning.

Examinations of long-term learning suggest that students do have relevant information stored in memory that is not accessible. Such studies typically use a cross-sectional design, with students of different ages taking tests on high school content (e.g., Bahrick & Phelps, 1987; Bahrick & Hall, 1991; see Conway, Cohen, & Stanhope, 1991, for a similar demonstration with university class content). While the emphasis is normally on the stability of performance over long periods of time (following an initial period of steep forgetting), what is crucial for present purposes is that students do better on recognition tests than on production measures.

In the present research, we aimed to bridge the laboratory and the classroom to evaluate whether student knowledge from prior courses can be reactivated quickly and easily. In two studies, we obtained measures of what students remembered from their prior courses before reactivating a subset of the target material with multiple-choice quizzes. The intervention was embedded in an out-of-class assignment in Experiment 1 and in a course pretest in Experiment 2. To preview, students showed long-term benefits from the intervention, as measured at the end of their college course.

1.1 | Experiment 1

The study was conducted as part of standard educational activities within the first year “Pharmacy Bridging Course” (PHCY 500) within the doctoral pharmacy program at a large state university. This four-week course reviews core material from prerequisite coursework in physiology, organic chemistry, biochemistry, applied math, and statistics. Each section of the course administers a course pretest (a baseline assessment to measure proficiency based on prior coursework and experiences) and a post-course test to measure mastery following completion of the course. We added questions to the course pretest to identify unretrievable information, embedded the intervention in classroom assignments, and used the course posttest to assess whether our intervention was successful.

2 | METHOD

2.1 | Students

One hundred and forty-five first-year pharmacy students participated in the study (representing 95% of the total class; 5% of students did not consent to having their data released for analysis). All students were taking a bridging course designed to help them transition into the Doctor of Pharmacy program at the flagship campus of a large state university. The average age of the students was 22 years old (range 19–32), and 81% of them had obtained a college degree. All students were required to have completed at least one semester of human physiology and anatomy prior to entering pharmacy school. As a group, their mean college grade point average upon admission was 3.5 (out of 4.0), and their mean Pharmacy College Admission Test score was 88%.

2.2 | Materials

Twenty-eight **key concepts** were identified as being relevant to the current course; sample items appear in Table 1 and the entire set is available from the senior author (Persky) upon request. Because the course is successful at preparing students for the Pharmacy program (and typically leads to ceiling performance on the post-course assessment), we chose items that were not slated to be explicitly reviewed in the course but were still highly relevant to the pharmacy curriculum. There were four items in each of seven areas of physiology: cardiology, renal, hematology, central nervous system, gastrointestinal, respiratory, and endocrine.

ExamSoft software (ExamSoft, ExamSoft Worldwide, Inc. ©, 2012) was used to administer a 93-item **course pretest**; the 28 critical concepts were tested along with 65 other items that were slated for review in the course. The noncritical items were tested in multiple-choice format but the critical items were tested in short-answer format. Only 25% of critical questions were answered correctly; 75% of this material was either forgotten or inaccessible. Based on the results of this pretest, the seven topics were ranked in scores from highest to lowest and the easiest topic was excluded from the

TABLE 1 Example questions for each of the six critical topics, including the fill-in-the-blank versions used on the pre- and post-course tests, as well as the multiple-choice question used in the intervention

Topic	Fill-in-the-blank questions/ MC stems	Multiple choice alternatives
Renal	This cell is within the Bowman's capsule that wrap around capillaries of the glomerulus and the slits between the foot-like projects allow blood to be filtered. _____	A. Podocytes B. Granulate cells C. Basal lamina D. Pedithelial cells
Hematology	The antibodies produce against a foreign blood type may cause hemolysis (rupturing) or _____ (clumping) of donated erythrocytes	A. Agglutination B. Extravasation C. Degranulation D. Coagulation
Central nervous system	_____ are nonneuronal cells that maintain homeostasis, form myelin, and provide support and protection for neurons in the central and peripheral nervous systems (e.g., astrocytes, Schwann cells)	A. Glial cells B. Dendritic cells C. Microcytes D. Ganglia
Gastrointestinal	There are 3 major phases of gastric secretion. Around 20–30% of the gastric secretory response to a meal occurs before food enters the stomach due to thoughts of food, sensory stimuli, and chewing. This phase is called the _____	A. Cephalic B. Gastric C. Intestinal D. Preparatory
Respiratory	This cell is interspersed among epithelial cells of the bronchioles and are mucus secreting. _____	A. Goblet cells B. Chief cells C. Mast cells D. D-cells
Endocrine	Thecal cells produce androgens in ovaries and _____ cells produce testosterone in testes	A. Leydig B. Sertoli C. Granulosa D. Graafian

experiment. The six remaining topics were divided into two sets matched on performance on the pretest. The **course posttest** was identical to the pretest.

The **course practice assignment** included items about current course content (for which feedback was provided, given it was not the focus of the experiment) as well as the critical intervention items (for which no feedback was provided). That is, questions about one of the two sets of critical topics (counterbalanced across subjects) were interspersed throughout the assignment. For each critical topic, the same four concepts were targeted as in the pretest. Following Park (2005), students responded in two stages: first, they answered a fill-in-the-blank version of each question, to confirm current (in)accessibility. They then received the same question in multiple-choice format and selected their response from four alternatives. This hybrid

format also has the advantage of being good for learning (Butler, Marsh, Slavinsky, & Baraniuk, 2014; Park, 2005, 2010; Park & Choi, 2008; Smith & Karpicke, 2014), combining the superior mnemonic benefits that often accrue from taking an open-ended test (Kang, McDermott, & Roediger III, 2007) with the objectiveness of a multiple-choice test that facilitates grading.

2.3 | Procedure

An overview of the procedure can be found in Figure 1. Prior to the course, students completed the pretest. During the first week of the course, all participants completed the course practice assignment, which involved questions on the intervention concepts as well as other questions related to the course. The personalized learning system OpenStax Tutor (OpenStax Tutor, Rice University ©, 2015) was used to administer the assignment. Grading was based solely on completing the assignment by the assigned deadline (i.e., not performance). In addition to enabling the experiment, distributing practice on the topics had the practical benefit of helping students with time management and reducing the overall amount of work each week during a time compressed course.

During the fourth week of the course, students completed the course posttest, which was identical to the course pretest. After taking the final assessment, they received feedback on all of the questions (i.e., the material relevant to the current and future courses) to ensure that they had the correct information moving forward in their coursework.

3 | RESULTS

All analyses for both experiments were performed using SPSS 23 software (IBM Corp, 2016). Cohen's *d* is the measure of effect size reported. The results reported are for the 24 critical items (from six topics) which were not covered in the review class, half of which were targeted for reactivation using multiple-choice questions in the course assignment completed in the first week of the class.

3.1 | Course pretest: Knowledge assessment

Participants correctly answered 24% of the critical short-answer questions. Seventy-five percent were either forgotten or inaccessible, even though all material should have been learned in prerequisite courses.

3.2 | Course assignment: Performance on intervention MC items

The critical intervention questions were embedded within the course assignment completed during the first week of the class. As expected,

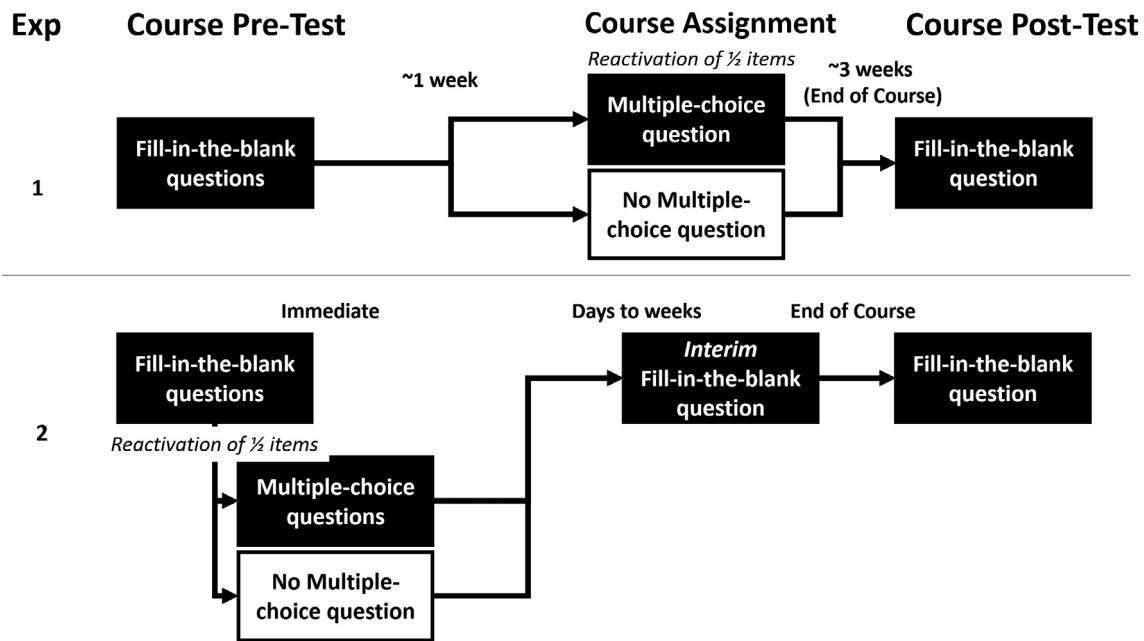


FIGURE 1 Both experiments had a course pretest and a final course posttest. The intervention was embedded in a course assignment in Experiment 1 and in the pretest in Experiment 2

students did much better on MC questions ($M = .65$) than when required to produce information from memory ($M = .32$). Given that they failed to answer the short-answer question correctly, they were above-chance at selecting the correct MC answer ($M = .50$).

3.3 | Course post-test

Figure 2 shows the main results for the final assessment in Experiment 1, which occurred 3 weeks after the intervention assignment. The data are plotted as a function of whether or not students were able to retrieve the requisite knowledge on the course pretest. Following successful retrieval, performance on the final test was high (84%), although some forgetting did occur over 3 weeks. The intervention slightly retarded this forgetting, as shown by a small but significant benefit for intervention items, relative to the control items (.88 vs. .81; $t[137] = 3.41$, $SEM = .02$, $p = .001$, $d = .32$). The key data involve the items which participants failed to retrieve initially; on the final test, participants were much more likely to answer them correctly following multiple-choice testing than for nontested control items, ($M = .42$ vs. .17; $t[144] = 10.64$, $SEM = .02$, $p < .001$, $d = 1.20$).

Within the intervention condition, performance on the final test was significantly higher if students had recognized the correct answer on the multiple-choice test, as opposed to answering it incorrectly ($M = .50$ vs. .33; $t[141] = 6.79$, $SEM = .03$, $p < .001$, $d = .56$). However, even when students answered the multiple-choice question incorrectly, the intervention was helpful: attempting to answer the question conferred an advantage over control items ($M = .33$ vs. .17; $t[142] = 5.39$, $SEM = .03$, $p < .001$, $d = .64$).

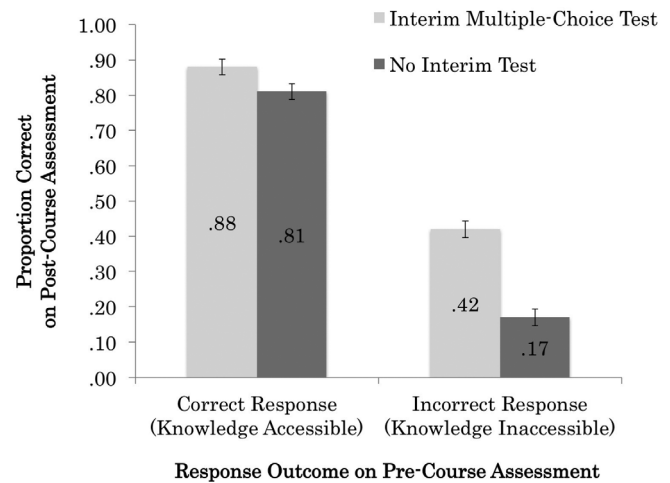


FIGURE 2 Proportion of correct responses on the post-course assessment in Experiment 1 as a function of response outcome on the course pretest (correct vs. incorrect) and experimental condition (interim multiple-choice test vs. no interim test). Error bars indicate standard error of the mean

4 | DISCUSSION

Performance on the post-course assessment showed that answering a single multiple-choice question (without feedback) reactivated marginal knowledge that was inaccessible before the course. This finding conceptually replicates prior research (Cantor et al., 2015) and extends it to a classroom setting and to more complex material than the simple declarative statements used by Berger et al. (1999). Students benefited the most when they correctly answered the multiple-choice question on the interim

test. However, it is interesting to note that students benefited even when they got the multiple-choice question wrong despite the fact that no feedback was provided. One possible explanation is that students may have later realized that one of the other alternatives for the multiple-choice question was correct, and thereby reactivated marginal knowledge. Of course, another possibility is that they could have looked up the answer at some point after taking the multiple-choice test. Although it is impossible to know for certain the mechanism(s) that produced this finding, it corresponds well with recent research showing that making errors on tests can be beneficial to learning for a variety of reasons (Arnold & McDermott, 2013; Huelser & Metcalfe, 2012; Kornell, Klein, & Rawson, 2015; Little, Bjork, Bjork, & Angello, 2012; Potts & Shanks, 2014).

4.1 | Experiment 2

A second experiment was conducted during the following year to conceptually replicate and extend the findings of Experiment 1. In the time between the two studies, the course was redesigned to be more structured (with a class period for each topic and less self-paced instruction), providing the opportunity to further evaluate the generalizability of our results.

In Experiment 2, we streamlined the intervention by combining the pretest and intervention test into a single event. That is, the pretest was changed to the hybrid test format (with each open-ended question followed by multiple-choice options), to enable it to serve as both the assessment of knowledge accessibility (via short answer responses) and the intervention used to reactivate access to marginal knowledge (via multiple-choice selections for ½ of critical items).

5 | METHOD

5.1 | Students

One hundred and forty-one first-year pharmacy students participated in the study (96% consented to having their data released for analysis). The sample of students in Experiment 2 was highly similar to Experiment 1 in age ($M = 22$, range 18–51), educational background (86% with a college degree), college GPA ($M = 3.5$), and Pharmacy College Admission Test scores ($M = 88\%$). Six students did not complete all of the measures so they were excluded from the analyses.

5.2 | Materials

The materials were the same as Experiment 1 except for the following changes: (a) some of the critical questions were slightly modified based on students' performance on those items in Experiment 1 and (b) additional questions were created to allow five questions per topic, yielding a total of 30 critical questions. For counterbalancing purposes, we used the same two sets of topics as in Experiment 1, which were match for difficulty.

5.3 | Procedure

An overview of the procedure can be found in Figure 1.

ExamSoft was again used to administer the course pretest, with the main change being that the hybrid question format was used for intervention items. That is, questions in the experimental condition required completion of a fill-in-the-blank prompt before the multiple-choice alternatives appeared. All other questions were in fill-in-the-blank format (20 questions: 5 filler questions for the 7th topic as well as 15 questions assigned to the control condition).

Over the course of the term, students took seven mid-course assessments, one for each topic. Each mid-course assessment was administered at the beginning of the class session about the topic to which it was related and consisted of the same five questions from the precourse assessment. All questions on the mid-course assessments were presented in fill-in-the-blank format regardless of whether the topic was assigned to the intervention or control condition for a particular student. Additional questions about current course materials were also included on the mid-course assessments. No feedback was given on these assessments.

Finally, students took the final test, following the procedures of Experiment 1.

6 | RESULTS

6.1 | Course pretest: Knowledge assessment

Performance on the pretest short answer questions was poor, indicating that students struggled to retrieve much of their background knowledge. Performance was slightly better for topics assigned to the control condition (12% correct) than for topics assigned to the experimental condition (10%). While this small difference was statistically significant ($M = .12$ vs. $.10$; $t[129] = 1.99$, $SEM = .01$, $p = .049$, $d = .15$), it is not problematic given that it works against our hypothesis that final performance will be higher in the experimental condition than the control.

6.2 | Course pretest: Intervention multiple-choice questions

For half of the critical questions, participants selected one of four multiple alternatives after submitting their short answer responses. On average, students selected the correct answer half of the time ($M = .50$). As in Experiment 1, students almost always selected the correct answer on the multiple-choice part of the question if they answered the question correctly on the open-ended part.

6.3 | Mid-course assessments

The top panel of Figure 3 shows the main results for the mid-course assessments. Given a retrieval failure on the pretest, the interim

multiple-choice test improved performance on the mid-course fill-in-the-blank assessment relative to questions assigned to the control condition ($M = .27$ vs. $.16$; $t[129] = 5.67$, $SEM = .02$, $p < .001$, $d = .69$). This benefit of the multiple-choice testing intervention was dependent upon students' success on that test: students performed much better on the mid-course assessment after answering the multiple-choice question correctly rather than incorrectly ($M = .42$ vs. $.13$; $t[125] = 12.90$, $SEM = .02$, $p < .001$, $d = 1.48$). Unlike Experiment 1, there was no benefit of taking the interim test if students selected an incorrect response option relative to questions assigned to the control condition ($M = .14$ vs. $.16$; $t[129] = 1.24$, $SEM = .02$, $p = .217$, $d = .17$). In addition, there was also no significant benefit of taking the interim test for questions answered correctly on the prior open-ended part ($M = .83$ vs. $.80$; $t[74] < 1$); however, it should be noted that many students did not correctly answer the open-ended part for any of the questions.

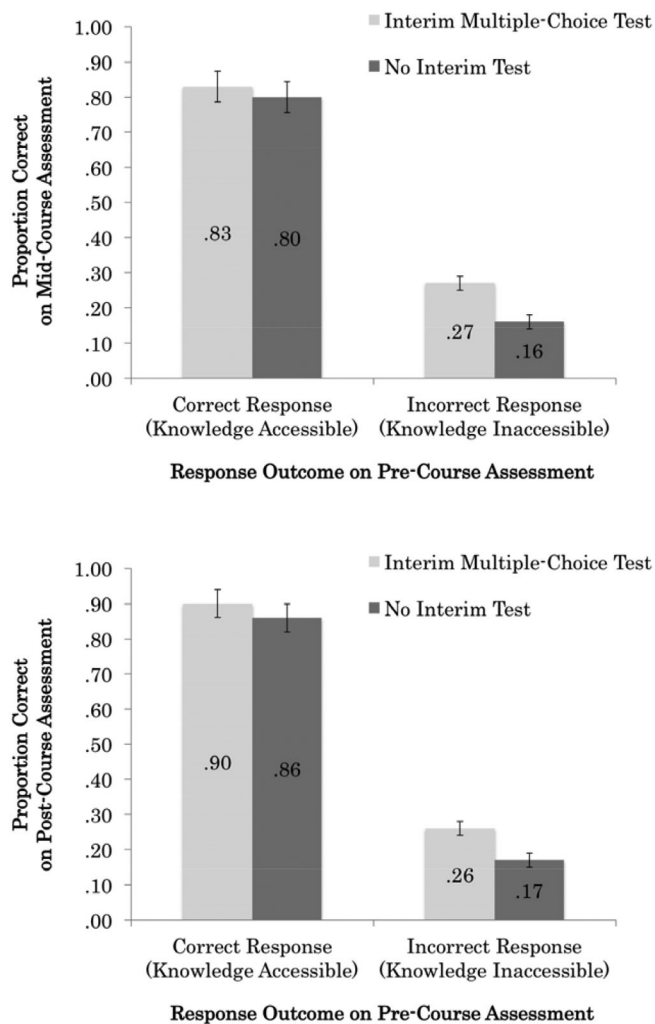


FIGURE 3 Proportion of correct responses on the mid-course assessments (top panel) and post-course assessment (bottom panel) in Experiment 2 as a function of response outcome on the open-ended part of the pre-course test (correct vs. incorrect) and experimental condition (interim multiple-choice test vs. no interim test). Error bars indicate standard error of the mean

6.4 | Course posttest

The bottom panel of Figure 3 shows the main results for the post-course assessment in Experiment 2. Overall, the results were consistent with the mid-course assessment. First, the benefit of taking the multiple-choice part of the test extended until the end of the course. That is, for information students could not retrieve on the pretest, multiple-choice testing led to a benefit on the final test, compared to control items ($M = .26$ vs. $.17$; $t[129] = 5.57$, $SEM = .02$, $p < .001$, $d = .56$). Again, this benefit was much stronger given the selection of the correct answer to the multiple-choice question ($M = .40$ vs. $.12$; $t[125] = 12.95$, $SEM = .02$, $p < .001$, $d = 1.43$).

There were two differences as compared to Experiment 1. First, there was no benefit when students answered multiple-choice questions incorrectly; they actually performed significantly worse on the final test after their errors, relative to control questions ($M = .13$ vs. $.17$; $t[129] = 2.55$, $SEM = .02$, $p = .012$, $d = .29$). Second, the multiple-choice test did not retard forgetting. That is, after answering the question correctly on the pretest, there was no further benefit of multiple-choice testing (on the final test), ($M = .90$ vs. $.86$; $t[74] < 1$).

7 | DISCUSSION

Experiment 2 replicated the main finding from Experiment 1 – multiple-choice testing reactivated marginal knowledge, even without feedback after answering. When students were unable to generate the answer on the open-ended part of the question on the pre-assessment (i.e., knowledge was inaccessible), answering the multiple-choice version immediately afterwards led to significantly better performance later in the course when asked the same open-ended question again. In contrast with Experiment 1, this finding was driven exclusively by items for which students correctly answered the interim multiple-choice question. Although there was a benefit to students of answering the interim multiple-choice question incorrectly in Experiment 1, this finding was not replicated in Experiment 2. It should also be noted that the size of the multiple-choice effect (on the final test) was smaller in Experiment 2 ($d = .56$) than in Experiment 1 ($d = 1.20$).

The changes in methodology across experiments (see Figure 1) may explain the slightly different results. In Experiment 2, the multiple-choice test intended to reactivate marginal knowledge was delivered immediately after knowledge was assessed, as opposed to mid-way through the course. In addition, the questions from Experiment 1 were modified for Experiment 2, and the difference in performance on the multiple-choice test in the two experiments suggests that these changes made the questions harder. Overall, the results of Experiment 2 suggest that the finding that a multiple-choice test can help reactivate marginal knowledge is quite robust; however, the magnitude of this effect seems to depend upon performance on the multiple-choice test, suggesting that factors that increase successful retrieval on the test may increase the amount of knowledge that is reactivated.

8 | GENERAL DISCUSSION

The results of these two classroom experiments show that taking a multiple-choice test, even without feedback, allows students to regain access to some of the knowledge learned in prerequisite courses. At the beginning of a Pharmacy review course, students were unable to retrieve many critical concepts from prior classes – knowledge that was important to facilitating their learning in future courses in the curriculum. Answering multiple-choice questions outside of class was sufficient to reactivate access to some of this knowledge – and this access lasted the entire course term. By demonstrating the generalizability of laboratory work on marginal knowledge to the classroom, our study opens up the possibility of using this approach to improve educational practice.

In particular, one potentially promising use of this finding would be strategic, just-in-time interventions to reactivate relevant marginal knowledge prior to new learning. To be clear, such an intervention would be unlikely to have as strong of an effect as substantive instruction – but it would be cheaper and quicker, while at the same time providing useful feedback to instructors about which material requires classroom coverage. Of course, students could always be instructed to restudy material on their own outside of class, perhaps yielding similar results. However, such an approach would not yield the same information to instructors, and its success would depend upon a student's ability to accurately judge whether information was learned (something that struggling students are particularly poor at; Kruger & Dunning, 1999).

Despite the potential for success in reactivating marginal knowledge, multiple-choice interventions should be coupled with other efforts to promote the maintenance of knowledge. Our studies are the first to target the kind of complex information required in college courses, as opposed to simple declarative facts – but we do not yet know, for example, whether the benefits would occur for all types of materials (e.g., computations). Furthermore, some types of learning do not lend themselves to multiple-choice testing, such as the skill of writing. Thus, our work is a first step toward improving the maintenance of knowledge, but only represents one tool in what is likely a complex toolbox. From the perspective of researchers, we know a lot about how to facilitate new learning, but much less about maintenance of knowledge (Butler & Raley, 2015) – highlighting the need for research in this area.

We close by noting that the loss of access to knowledge is particularly acute in health professional education. Entry-level degrees for most health professions require completion of prerequisite coursework prior to entering the degree program and utilize curricula in which courses build closely upon each other. Furthermore, every health position has some knowledge that students and practitioners must have at their fingertips. Much of this knowledge could be looked up or re-learned with time; however, if it is inaccessible in the moment, then new learning could be hindered, or even worse – a mistake could be made in a clinical situation with a patient.

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

AUTHOR CONTRIBUTIONS

All authors contributed to the idea for the research and designed the experiments. A.M.P. and K.C. collected the data with assistance from A.C.B. and A.C.B.-M. A.C.B. and A.C.B.-M. analyzed the data. A.C.B., A.C.B.-M., and A.M.P. drafted the manuscript. All authors edited the manuscript.

DATA AVAILABILITY STATEMENT

Correspondence and requests for materials and data should be addressed to ACB (andrew.butler@wustl.edu)

ORCID

Andrew C. Butler  <https://orcid.org/0000-0002-6367-0795>

REFERENCES

- Arnold, K. M., & McDermott, K. B. (2013). Test-potentiated learning: Distinguishing between direct and indirect effects of tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 940–945.
- Bahrack, H. P., & Hall, L. K. (1991a). Lifetime maintenance of high school mathematics content. *Journal of Experimental Psychology: General*, *120*, 20–33.
- Bahrack, H. P., & Phelps, E. (1987). Retention of Spanish vocabulary over 8 years. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 344–349.
- Berger, S. A., Hall, L. K., & Bahrack, H. P. (1999). Stabilizing access to marginal and submarginal knowledge. *Journal of Experimental Psychology: Applied*, *5*, 438–447.
- Brown, R., & McNeil, D. (1966). The “tip of the tongue” phenomenon. *Journal of Verbal Learning and Verbal Behavior*, *5*, 325–337.
- Butler, A. C., Marsh, E. J., Slavinsky, J. P., & Baraniuk, R. G. (2014). Integrating cognitive science and technology improves learning in a STEM classroom. *Educational Psychology Review*, *26*, 331–340.
- Butler, A. C., & Raley, N. D. (2015). The future of medical education: Assessing the impact of interventions on long-term retention and clinical care. *Journal of Graduate Medical Education*, *7*, 483–485.
- Cantor, A. D., Eslick, A. N., Marsh, E. J., Bjork, R. A., & Bjork, E. L. (2015). Multiple-choice tests stabilize access to marginal knowledge. *Memory & Cognition*, *43*, 193–205.
- Conway, M. A., Cohen, G., & Stanhope, N. (1991). On the very long-term retention of knowledge acquired through formal education: Twelve years of cognitive psychology. *Journal of Experimental Psychology: General*, *120*, 395–409.
- Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research*, *66*, 227–268.
- Custers, E. J. (2010). Long-term retention of basic science knowledge: A review study. *Advances in Health Sciences Education*, *15*, 109–128.
- Custers, E. J., & Ten Cate, O. T. (2011). Very long-term retention of basic science knowledge in doctors after graduation. *Medical Education*, *45*, 422–430.
- Emke, A. R., Butler, A. C., & Larsen, D. P. (2016). Effects of Team-Based Learning on short-term and long-term retention of factual knowledge. *Medical Teacher*, *38*, 306–311.
- ExamSoft Worldwide, Inc. (2012). *ExamSoft*. Dallas, TX: ExamSoft Worldwide, Inc.
- Huelsen, B. J., & Metcalfe, J. (2012). Making related errors facilitates learning, but learners do not know it. *Memory & Cognition*, *40*, 514–527.
- IBM Corp. (2016). *IBM SPSS statistics, Version 23.0*. Armonk, NY: IBM Corp.

- Kang, S. H., McDermott, K. B., & Roediger, H. L., III. (2007). Test format and corrective feedback modify the effect of testing on long-term retention. *European Journal of Cognitive Psychology, 19*, 528–558.
- Kornell, N., Klein, P. J., & Rawson, K. A. (2015). Retrieval attempts enhance learning, but retrieval success (versus failure) does not matter. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*, 283–294.
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self assessments. *Journal of Personality and Social Psychology, 77*, 1121–1134.
- Larsen, D. P., Butler, A. C., & Roediger, H. L., III. (2009). Repeated testing improves long-term retention relative to repeated study: A randomised controlled trial. *Medical Education, 43*, 1174–1181.
- Ling, Y. U., Swanson, D. B., Holtzman, K., & Bucak, S. D. (2008). Retention of basic science information by senior medical students. *Academic Medicine, 83*, S82–S85.
- Little, J. L., Bjork, E. L., Bjork, R. A., & Angello, G. (2012). Multiple-choice tests exonerated, at least of some charges: Fostering test-induced learning and avoiding test-induced forgetting. *Psychological Science, 23*, 1337–1344.
- Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16*(5), 519–533.
- OpenStax CNX. (2015). *OpenStax tutor*. Houston, TX: OpenStax CNX.
- Park, J. (2005). Learning in a new computerized testing system. *Journal of Educational Psychology, 97*, 436–443.
- Park, J. (2010). Constructive multiple-choice testing system. *British Journal of Educational Technology, 41*, 1054–1064.
- Park, J., & Choi, B. C. (2008). Higher retention after a new take-home computerised test. *British Journal of Educational Technology, 39*, 538–547.
- Potts, R., & Shanks, D. R. (2014). The benefit of generating errors during learning. *Journal of Experimental Psychology: General, 143*, 644–667.
- Roediger, H. L., III, & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences, 15*(1), 20–27.
- Smith, M. A., & Karpicke, J. D. (2014). Retrieval practice with short-answer, multiple-choice, and hybrid tests. *Memory, 22*, 784–802.
- Sullivan, P. B., Gregg, N., Adams, E., Rodgers, C., & Hull, J. (2013). How much of the paediatric core curriculum do medical students remember? *Advances in Health Sciences Education, 18*, 365–373.
- Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior, 5*, 381–391.
- Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*(5), 352–373.

How to cite this article: Butler AC, Black-Maier AC, Campbell K, Marsh EJ, Persky AM. Regaining access to marginal knowledge in a classroom setting. *Appl Cognit Psychol.* 2020;1–8. <https://doi.org/10.1002/acp.3679>