



Frequent Quizzing Accelerates Classroom Learning

Chunliang Yang

Beijing Normal University

David R. Shanks

University College London

Wenbo Zhao, Tian Fan, and Liang Luo

Beijing Normal University

Overview

Although frequent testing (i.e., practice retrieval) seems to be unwelcome to many students, a wealth of studies has established that taking a test is one of the most effective strategies to facilitate learning. Specifically, over the last century, hundreds of laboratory and classroom studies on practice retrieval have documented that retrieving information from memory can more effectively facilitate learning and consolidate long-term retention than restudying the same material, a phenomenon that goes by a range of names, including: *the testing effect*, *the retrieval practice effect*, *test-enhanced learning*, *test-potentiated learning*, and *retrieval-enhanced learning* (for reviews, see Rawson & Dunlosky, 2011; Roediger & Butler, 2011; Yang et al., 2021). In addition, testing tends to be more beneficial than many other elaborative strategies, such as note-taking (Heitmann et al., 2018), concept-mapping (Karpicke & Blunt, 2011), and self-explaining (Larsen et al., 2013).

It seems compelling that learners should seek to utilize testing for optimizing their learning outcomes. However, both learners and instructors tend to underappreciate its merits despite its broad benefits, leading to its underemployment in educational settings. Besides lack of meta-awareness, other reasons for its practical underemployment will be discussed below. The main point here is that, given its efficacy, testing has not been as widely implemented in the classroom as it deserves to be. This is a concrete example of the well-known “know-do gap” between what we learn from cognitive research and what we implement in real situations (cf. Rowe & Hattie, this volume). The current chapter aims to help bridge this gap by providing practitioners with a summary of empirical findings on the classroom testing effect, providing some responses to practitioners’ potential concerns about classroom quizzing, and offering some suggestions about how to augment the enhancing effect of classroom quizzing.

Empirical Findings on Test-Enhanced Learning

Although testing has been typically considered as an assessment *of* learning, it is also an effective assessment *for* learning. Ample laboratory and classroom studies have demonstrated that testing not only consolidates long-term retention of studied information, a phenomenon termed the *backward testing effect* (for reviews, see Adesope et al., 2017; Roediger & Karpicke, 2006a; Roediger et al., 2011; Rowland, 2014), but also effectively potentiates subsequent learning of new information, a phenomenon referred to as the *forward testing effect* (for reviews, see Chan et al., 2018; Pastötter & Bäuml, 2014; Yang et al., 2017).

Test-Enhanced Consolidation of Studied Information

Roediger and Karpicke (2006b) conducted a classic study demonstrating the backward testing effect. Undergraduates studied two texts, with one text read twice and the other read once and tested once. In

a final test administered one week later, the tested passage (56%) was substantially better recalled than the restudied one (42%), reflecting that testing more effectively consolidates long-term retention than restudying.

The backward testing effect has also been documented in classroom studies. For instance, in a Brain and Behavior course, McDaniel et al. (2007) had students from the University of New Mexico read about 40 pages of textbook materials each week. After reading the weekly course assignments, students either reread ten facts selected from the reading materials or took a short answer test on these facts with corrective feedback (i.e., the correct answers) provided after they answered each question. This course lasted for 6 weeks, and students took two exams across the term, with one administered at Week 3 and the other at Week 6. In these exams, the 60 reread and 60 tested facts were assessed in a multiple-choice format. Exam performance for the tested facts (57%) was substantially better than that for reread facts (47%).

In short, testing of studied information can be employed as a practical strategy to consolidate long-term retention.

Test-Potentiated Learning of New Information

Besides the benefit that testing confers on long-term retention of studied/tested information, an emerging body of recent studies has found that testing on previously studied information can also promote subsequent learning of new information. The forward (prospective) benefits of testing on new learning reliably generalize to different types of study materials, such as lecture videos, text passages, artists' painting styles, and foreign word translations, and the effect extends to different populations, such as children, college students, older adults, and patients with traumatic brain injury (for reviews, see Chan et al., 2018; Pastötter & Bäuml, 2014; Yang & Shanks, 2018).

Szpunar et al. (2013) provided an illustration. In their experiment, two (Restudy vs. Test) groups of undergraduates watched a 4-segment Introductory Statistics lecture video. After watching each of Segments 1-3, the Restudy group restudied the six key concepts delivered in the just-studied segment. By contrast, the Test group was tested on these six key concepts without corrective feedback (e.g., *What is the relation between a population and a sample?*). Then, both groups were tested on Segment 4 after watching it. In brief, both groups watched the same lecture segments, and the only difference between groups was that the Restudy group restudied, whereas the Test group was tested on Segments 1-3.

As a result, the Test group correctly answered more questions (89%) than the Restudy group (65%) in the test on Segment 4, reflecting a forward benefit of testing (that is, interim testing on Segments 1-3, by comparison with restudying, boosted subsequent learning of Segment 4). Szpunar and colleagues also found that the Test group (19%) reported fewer instances of mind wandering than the Restudy group (39%) when they were occasionally prompted to report whether they were mind wandering or focusing on the lecture during the study phase. Furthermore, the Test group took more notes than the Restudy group while watching the lecture video.

Underemployment of Test-Enhanced Learning and Potential Concerns About Quizzing

Although, as discussed above, the benefits of testing are substantial, researchers have frequently expressed dismay that learners, instructors, and policy-makers tend not to appreciate these benefits and that test-enhanced learning has not been fully applied in the classroom to boost students' attainment (Roediger & Karpicke, 2006b; cf. Rowe & Hattie, this volume). For example, Tullis and Maddox (2020) asked 84 middle and 228 high school students to report their most frequently used strategies. For these

students, 43% reported frequently rereading (restudying), whereas only 16% reported using retrieval practice. Similar to students, some teachers may also lack full appreciation of the benefits of test-enhanced learning. For instance, Morehead et al. (2016) asked 146 teachers at Colorado State University to explain why they thought students should test themselves. Sixty-eight percent replied that students should test themselves to figure out how well they have mastered the material, and only 19% reported that students will learn more through testing than through rereading. These findings reflect that a majority of these teachers regard testing primarily as an assessment tool to identify knowledge gaps rather than an effective strategy to boost learning.

Besides metacognitive unawareness, underemployment of test-enhanced learning might derive from many other sources. For instance, teachers may worry that frequent quizzing consumes in-class time and/or provokes test anxiety, which in turn impairs students' academic performance. Below, we list several concerns that teachers may have about the potential drawbacks of quizzing and provide responses to these concerns.

Quizzing Borrows Time from Other Teaching Activities

One obstacle for implementing classroom quizzes is that administering them is time-consuming and scoring can sometimes be excessively demanding. Because quizzing draws time away from didactic teaching and curricula are typically too full to spare time for quizzing, some instructors may prefer to minimize the uses of classroom quizzing. However, a recent meta-analysis by Yang et al. (2021), combining results from nearly 50,000 students and over 200 research reports, demonstrated that frequent classroom quizzing enhances students' academic achievement compared to practice-as-usual. Such findings suggest that even though quizzing borrows time from didactic teaching, the tradeoff is worthwhile.

To avoid the tradeoff between teaching and quizzing, teachers can consider administering quizzes outside the classroom. Some studies have investigated the impacts of out-of-classroom quizzes on students' attainment (e.g., Grimstad & Grabe, 2004; Griswold, this volume; Marden et al., 2013). In these studies, quizzes were assigned as homework or made available online, and students were allowed to answer them at any time and place they chose. Yang et al.'s (2021) meta-analysis found that quizzes administered out of the classroom also significantly boost students' attainment, although not quite as effectively as ones administered in-class.

On the bright side, with the development of new technologies, more and more modes are available to make it easy to administer and score quizzes, such as clicker response systems, course management systems (e.g., *Blackboard*), collaborative learning platforms (e.g., *PeerWise*), and smartphones. Apart from these technologies, Pyc et al. (2014) suggested alternative ways to easily incorporate quizzes into classroom instruction, such as orally asking questions and waiting for students to generate responses, or providing students with colored index cards and asking them to select a colored flashcard and hold it up to respond to true/false or multiple-choice questions. All these techniques are virtually cost-free and allow teachers to easily implement immediate quizzes. Furthermore, Yang et al.'s (2021) meta-analysis found that quiz administration mode (e.g., paper-and-pen, web-based, oral, or clicker response system) has minimal influence on the magnitude of test-enhanced learning, implying that the enhancing effect of classroom quizzing is independent of administration modality.

In short, although quizzing borrows time from other teaching activities, the tradeoff is worthwhile.

Does Frequent Quizzing Provoke Test Anxiety?

Another legitimate concern that some teachers may have, which makes them reluctant to incorporate frequent quizzing into their teaching, is that quizzing may provoke test anxiety, which causes a variety of

learning detriments, such as difficulty concentrating and poor test performance. However, two sets of research findings should go some way towards allaying this concern.

The first set of evidence is that low-stake (or no-stake) tests, which arouse minimal test anxiety, can effectively promote learning. In their meta-analysis, Yang et al. (2021) assessed the moderating role of stake-level on test-enhanced learning. Specifically, Yang and colleagues divided 295 classroom experiments into two sub-categories according to the stake-level (i.e., whether quiz performance was incorporated into course grades) of quizzes: high (Yes) vs. low (No). The results showed that both high- and low-stake quizzes significantly enhanced students' exam performance, and there was no significant difference between high- and low-stake quizzes, revealing that even low- or no-stake quizzes can facilitate learning (for related findings, see Davis, 2013). Indeed, some studies have even reported that low-stake quizzes are more beneficial than high-stake ones (e.g., Schrank, 2016).

The second set of findings is that frequent testing has little impact on or even reduces (rather than increases) test anxiety. For instance, in a large sample study (over 1,000 college participants), Yang et al. (2020) observed that interpolating tests across a study phase has minimal influence on participants' test anxiety. Szpunar et al. (2013) found that frequent tests significantly relieve test anxiety (for related findings, see Khanna, 2015). Furthermore, in a large-scale survey conducted by Agarwal et al. (2014), 72% of 1,306 middle and 102 high school students reported that frequent quizzes made them less anxious about exams, with only 8% reporting the opposite. In another survey conducted by Sullivan (2017), about 91% of 353 college students agreed with the statement *The option to retake a quiz reduced my test anxiety*, and only 3.1% disagreed.

In education, testing is frequently criticized in the media and by parents, largely due to standardized tests with high stake levels which can undeniably induce high test-anxiety (Pyc et al., 2014). However, what we recommend here is frequent low-stake quizzing or testing, which is very different from high-stake standardized tests. Furthermore, as summarized above, low-stake quizzes facilitate classroom learning without causing test anxiety.

Does Test-Enhanced Learning Generalize to Different Test Formats, Subjects, and Populations?

Teachers may wonder “Does test-enhanced learning work in my class, for the subject I’m teaching, for my particular students? Are all test formats valid?” The generalizability of test-enhanced learning to different test formats, subjects, and populations (i.e., students at different education levels) has been repeatedly investigated in previous studies. Through integrating research results from previous studies, Yang et al.’s (2021) meta-analysis found that all kinds of test formats (e.g., matching, fill-in-the-blank, short answer, multiple-choice, cued recall, free recall, and so on) reliably enhance learning (for related discussion, see Pyc et al., 2014); the classroom testing effect generalizes across all subject categories for which sufficient data are available, including Accounting/Business/Finance, Biology, Chemistry, Geography, Education, History, Medical/Nursing/ Physiology, Language/Reading/Vocabulary, Psychology, and so on. Test-enhanced learning works across different education levels, including elementary school, middle school, high school, and university/college. However, the evidence for benefits in adult education is less clear, due to a very small number of published studies. These findings jointly establish the stability and generalizability of test-enhanced learning to different test formats, subjects, and populations (Yang et al., 2021).

Does Quizzing Only Enhance “Inert Knowledge”?

Testing (quizzing) has been occasionally met with a criticism that it is a “drill-and-kill” strategy, which only promotes retention of “inert knowledge” that cannot be applied to solve new problems in new

contexts. However, this criticism has been repeatedly contradicted by studies showing that testing can effectively promote knowledge transfer and application (for a review, see Carpenter, 2012). For example, Butler (2010) had students study six texts, three of which were repeatedly restudied, and the other three repeatedly tested (e.g., *What is the wing structure of a bat like relative to that of a bird?*). On a final test administered one week later, students were instructed to apply knowledge studied from the texts to solve new inferential questions in a very different domain (e.g., *How would a military aircraft modeled after a bat wing differ from traditional aircrafts?*). The results showed greater accuracy for new inferential questions for tested passages (68%) than for restudied ones (44%), suggesting that testing promotes far transfer knowledge application, enabling learners to solve new problems in new contexts (for related findings, see Karpicke & Blunt, 2011).

In another study conducted by Kromann et al. (2009), 81 medical students studied resuscitation skills in an in-hospital resuscitation course. In the intervention (testing) condition, students took a 30-minute quiz following three and a half hours of instruction and training. By contrast, those in the control condition received four hours of instruction and training without quizzing. Two weeks later, all students undertook an assessment to measure their resuscitation skills, in which they were provided cardiac arrest scenarios and required to perform the resuscitation treatment. Assessment performance in the intervention condition (83%) was much better than in the control condition (73%), implying that testing can enhance skill learning.

In their meta-analysis, Yang et al. (2021) coded classroom studies into three categories according to their exam content type: Fact (i.e., memory of specific course contents, such as historical events), Concept (e.g., concept comprehension and inductive inferences that go beyond direct experience, such as learning rules of syntax in a foreign language course), and Problem-solving (i.e., applying learned knowledge or skills to solve problems, such as performing resuscitation treatment in cardiac arrest scenarios). Through integrating results of previous classroom studies, Yang et al. found that class quizzing not only promotes learning of facts, but also facilitates concept learning (Jacoby et al., 2010; Karpicke et al., 2014; Yang & Shanks, 2018) and enhances knowledge application in the service of problem solving (Kromann et al., 2010; Sennhenn-Kirchner et al., 2018).

In summary, the above empirical and meta-analytic findings jointly contradict the view that testing is a “drill-and-kill” strategy and only produces mnemonic benefits to inert knowledge. Instead, these findings imply that testing is also an effective tool to promote knowledge transfer and application.

Does Frequent Quizzing Widen the Achievement Gap among Students?

It is reasonable for teachers to suspect that frequent quizzing may widen the achievement gap (e.g., between novices and experts in a subject area, less-skilled and more-skilled readers, students with overall lower and higher academic performance). For instance, do some students benefit less from class quizzing or does class quizzing exert detrimental effects on some students’ academic performance because they typically perform less well on quizzes? Poor relative quiz performance may cause a sense of failure, leading to poorer academic performance. If this is the case, might frequent quizzing widen the achievement gap among students? Recent research demonstrates that this concern may be misplaced – indeed, the opposite seems to be the case. For example, Agarwal et al. (2017) and Yang et al. (2020) observed that the benefits of testing are larger for individuals with low than with high working memory capacity. Brewer and Unsworth (2012) found that individuals with lower IQs experience a larger enhancing effect of testing than those with higher IQs. Although high test anxiety is typically associated with poor learning achievement, Yang et al. (2020) and Clair et al. (2020) consistently observed little difference in the magnitude of test-enhanced learning between students with high and low levels of test anxiety.

These findings point to the conclusion that students, in general, will benefit from testing. Hence, frequent quizzing may narrow, rather than broaden, the achievement gap among students.

Optimizing the Benefits of Quizzing

The above discussion summarizes the broad benefits of classroom quizzing and provides responses to some of the concerns that teachers might have about frequent quizzing. Below, we offer several suggestions about when and how to implement quizzes to maximize the magnitude of test-enhanced learning.

Administer Quizzes After Teaching

In previous studies, quizzes have occasionally been presented before formal teaching, with to-be-taught content tested prior to teaching. Thus, a student might be quizzed about the causes of an historical event prior to a lesson on those causes. Under such circumstances, quiz performance is likely to be quite poor, compared to the more standard procedure in which quizzes are completed at the end of a class with taught content assessed in these post-class quizzes. Surprisingly, pre-class quizzes have been found to enhance learning. Although some recent experiments (e.g., Pan & Sana, 2021) have found pre-testing to be more effective than post-testing, Yang et al.'s (2021) meta-analysis aggregating all available evidence up to 2019 found that the enhancing effect of post-class quizzes is nearly triple that of pre-class quizzes. Hence, instructors are advised to administer quizzes after formal lecturing (cf. Carpenter, et al., this volume).

Provide Corrective Feedback

It has been well established in both laboratory and classroom studies that providing corrective feedback (i.e., the correct answer to a given test question) is a fortifier of test-enhanced learning since corrective feedback enables learners to detect and correct wrong answers. Many laboratory and classroom studies have shown that the memorial benefits of testing are stronger when followed by corrective feedback (e.g., Bangert-Drowns et al., 1991; Kang et al., 2007; Vojdanoska et al., 2010). Two meta-analyses also reached the same conclusion. Rowland (2014), who synthesized results from 159 laboratory experiments, found that the magnitude of the testing effect is doubled when feedback is provided, in comparison with not providing feedback. More recently, Yang et al. (2021), who integrated results from 222 classroom studies, observed that providing corrective feedback following quizzing increased the magnitude of quiz-induced enhancement by a factor of about 1.5. It is indeed unusual to administer a quiz without providing corrective feedback in the classroom (Yang et al., 2021).

These consistent findings lead to the strong recommendation that instructors offer corrective feedback after quizzing.

Administer Repeated Quizzing

Although it has been widely documented that a single test is sufficient to enhance memory compared to restudying, many laboratory studies have observed that repeated tests (i.e., with studied content tested repeatedly) produce a larger enhancing effect on knowledge retention and transfer than a single test (e.g., Butler, 2010; Dunlosky et al., this volume; Roediger & Karpicke, 2006b). The enhancing effect of repeated tests has been re-confirmed by many classroom studies. Moreover, Yang et al.'s (2021) meta-analysis coded the number of test repetitions (i.e., how many times the studied information was tested), and conducted analyses to quantify the relation between the magnitude of test-enhanced learning and the number of test (quiz) repetitions. The results showed a clear trend that the more occasions on which class content is quizzed, the more effectively quizzing aids exam performance.

These findings endorse a recommendation that teachers should consider incorporating repeated quizzing on key course concepts into classroom practice. For instance, teachers can test key concepts repeatedly in quizzes administered immediately after the class, one week later, and one month later. Of course, teachers should also take the tradeoff between repeated quizzing and didactic teaching into consideration.

Administer Long-Term Quizzing Interventions

Yang et al.'s (2021) meta-analysis investigated how the effectiveness of test-enhanced learning varies as a function of quiz treatment duration. Specifically, classroom studies were classified into four sub-categories according to the duration of their quizzing intervention: *Single class* (i.e., studies which only administered a quiz in a single class), *< Semester* (i.e., studies which implemented the quiz treatment for less than a semester, but longer than a single class), *= Semester* (i.e., studies which included quizzes across a whole semester), and *> Semester* (i.e., studies which included quizzes across more than a semester, such as across an academic year). Integrating results across 573 classroom experiments, Yang and colleagues found that the longer the quizzing intervention, the greater the magnitude of test-enhanced learning.

Test-enhanced engagement and learning motivation can account for the positive relationship between the classroom testing effect and treatment duration. For instance, Schrank (2016) found that about 90% of students in a daily-exam-class attended his Introductory Sociology class, and attendance rates were maintained steadily across a whole semester. By contrast, attendance rates in a control class, which did not take a daily quiz, dramatically decreased from about 90% to about 66% across the semester. Consistently, other studies have observed that frequent tests enhance note-taking (Szpunar et al., 2013), boost motivation (e.g., attention and study time, see Pastötter et al., 2011; Yang et al., 2017), and reduce task-unrelated mind wandering (Jing et al., 2016; Szpunar et al., 2013).

To conclude, these findings suggest that instructors consider administering long-term quizzing interventions, such as regular quizzes across a whole semester or an academic year.

Summary

Hundreds of studies have demonstrated that testing (quizzing) is an effective strategy to boost classroom learning. However, learners, instructors, and policymakers tend to lack full appreciation of the virtues of testing, leading to its underemployment in practice (see Rowe & Hattie, this volume, for a discussion of this issue). The current chapter attempts to bridge this know-do gap and provide practitioners with scaffolding in implementing quizzing in the classroom. In short, testing consolidates long-term retention of studied information (i.e., the backward testing effect) and potentiates subsequent learning of new information (i.e., the forward testing effect). Although teachers may have concerns about potential drawbacks of quizzing, research findings run counter to many of these concerns (see Rowe & Hattie, this volume). Instructors are recommended to administer quizzing in the classroom, and quizzes should be implemented after formal lecturing. Following quizzing, corrective feedback ought to be provided. Repeated quizzing and long-term quizzing treatment are more beneficial than single/short-term quizzing treatment.

Author Note

This chapter was supported by the Natural Science Foundation of China (32000742; 32171045), the Fundamental Research Funds for the Central Universities (2019NTSS28), and the United Kingdom Economic and Social Research Council (ES/S014616/1).

References

- Adesope, O. O., Trevisan, D. A., & Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of Educational Research, 87*(3), 659-701. <https://doi.org/10.3102/0034654316689306>
- Agarwal, P. K., D'Antonio, L., Roediger, H. L., McDermott, K. B., & McDaniel, M. A. (2014). Classroom-based programs of retrieval practice reduce middle school and high school students' test anxiety. *Journal of Applied Research in Memory and Cognition, 3*(3), 131-139. <https://doi.org/10.1016/j.jarmac.2014.07.002>
- Agarwal, P. K., Finley, J. R., Rose, N. S., & Roediger, H. L. (2017). Benefits from retrieval practice are greater for students with lower working memory capacity. *Memory, 25*(6), 764-771. <https://doi.org/10.1080/09658211.2016.1220579>
- Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61*(2), 213-238. <https://doi.org/10.3102/00346543061002213>
- Brewer, G. A., & Unsworth, N. (2012). Individual differences in the effects of retrieval from long-term memory. *Journal of Memory and Language, 66*(3), 407-415. <https://doi.org/10.1016/j.jml.2011.12.009>
- Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(5), 1118-1133. <https://doi.org/10.1037/a0019902>
- Carpenter, S. K. (2012). Testing enhances the transfer of learning. *Current Directions in Psychological Science, 21*(5), 279-283. <https://doi.org/10.1177%2F0963721412452728>
- Chan, J. C., Meissner, C. A., & Davis, S. D. (2018). Retrieval potentiates new learning: A theoretical and meta-analytic review. *Psychological Bulletin, 144*(11), 1111-1146. <https://doi.org/10.1037/bul0000166>
- Clair, B. S., Putnam, P., Miller, H. L., Larsen, R., & Jensen, J. L. (2020). Test-enhanced learning and incentives in Biology education. *CBE—Life Sciences Education, 19*(3) (ar40), 1-6. <https://doi.org/10.1187/cbe.19-11-0226>
- Davis, K. A. (2013). Using low- and no-stakes quizzing for student self-evaluation of readiness for exams. *International Journal of Construction Education and Research, 9*(4), 256-271. <https://doi.org/10.1080/15578771.2013.809036>
- Grimstad, K., & Grabe, M. (2004). Are online study questions beneficial? *Teaching of Psychology, 31*(2), 143-146. https://doi.org/10.1207/s15328023top3102_8
- Heitmann, S., Grund, A., Berthold, K., Fries, S., & Roelle, J. (2018). Testing is more desirable when it is adaptive and still desirable when compared to note-taking. *Frontiers in Psychology, 9*, 2596. <https://doi.org/10.3389/fpsyg.2018.02596>
- Jacoby, L. L., Wahlheim, C. N., & Coane, J. H. (2010). Test-enhanced learning of natural concepts: Effects on recognition memory, classification, and metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(6), 1441-1451. <https://doi.org/10.1037/a0020636>

- Jing, H. G., Szpunar, K. K., & Schacter, D. L. (2016). Interpolated testing influences focused attention and improves integration of information during a video-recorded lecture. *Journal of Experimental Psychology: Applied*, 22(3), 305-318. <https://doi.org/10.1037/xap0000087>
- Kang, S. H. K., McDermott, K. B., & Roediger, H. L. (2007). Test format and corrective feedback modify the effect of testing on long-term retention. *European Journal of Cognitive Psychology*, 19(4-5), 528-558. <https://doi.org/10.1080/09541440601056620>
- Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, 331(6018), 772-775. <https://doi.org/10.1126/science.1199327>
- Karpicke, J. D., Blunt, J. R., Smith, M. A., & Karpicke, S. S. (2014). Retrieval-based learning: The need for guided retrieval in elementary school children. *Journal of Applied Research in Memory and Cognition*, 3(3), 198-206. <https://doi.org/10.1016/j.jarmac.2014.07.008>
- Khanna, M. M. (2015). Ungraded pop quizzes: Test-enhanced learning without all the anxiety. *Teaching of Psychology*, 42(2), 174-178. <https://doi.org/10.1177%2F0098628315573144>
- Kromann, C. B., Bohnstedt, C., Jensen, M. L., & Ringsted, C. (2010). The testing effect on skills learning might last 6 months. *Advances in Health Sciences Education*, 15, 395-401. <https://doi.org/10.1007/s10459-009-9207-x>
- Kromann, C. B., Jensen, M. L., & Ringsted, C. (2009). The effect of testing on skills learning. *Medical Education*, 43(1), 21-27. <https://doi.org/10.1111/j.1365-2923.2008.03245.x>
- Larsen, D. P., Butler, A. C., & Roediger, H. L. (2013). Comparative effects of test-enhanced learning and self-explanation on long-term retention. *Medical Education*, 47(7), 674-682. <https://doi.org/10.1111/medu.12141>
- Marden, N. Y., Ulman, L. G., Wilson, F. S., & Velan, G. M. (2013). Online feedback assessments in physiology: Effects on students' learning experiences and outcomes. *Advances in Physiology Education*, 37(2), 192-200. <https://doi.org/10.1152/advan.00092.2012>
- McDaniel, M. A., Anderson, J. L., Derbish, M. H., & Morrisette, N. (2007). Testing the testing effect in the classroom. *European Journal of Cognitive Psychology*, 19(4-5), 494-513. <https://doi.org/10.1080/09541440701326154>
- Morehead, K., Rhodes, M. G., & DeLozier, S. (2016). Instructor and student knowledge of study strategies. *Memory*, 24, 257-271. <https://doi.org/10.1080/09658211.2014.1001992>
- Pan, S. C., & Sana, F. (2021). Pretesting versus posttesting: Comparing the pedagogical benefits of errorful generation and retrieval practice. *Journal of Experimental Psychology: Applied*, 27(2), 237-257. <https://doi.org/10.1037/xap0000345>
- Pastötter, B., & Bäuml, K. H. T. (2014). Retrieval practice enhances new learning: The forward effect of testing. *Frontiers in Psychology*, 5, 286. <https://doi.org/10.3389/fpsyg.2014.00286>
- Pastötter, B., Schicker, S., Niedernhuber, J., & Bäuml, K. H. T. (2011). Retrieval during learning facilitates subsequent memory encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(2), 287-297. <https://doi.org/10.1037/a0021801>
- Pyc, M. A., Agarwal, P. K., & Roediger, H. L. (2014). Test-enhanced learning. In V. A. Benassi, C. E. Overson, & C. M. Hakala (Eds.), *Applying science of learning in education: Infusing psychological science into the curriculum* (pp. 78-90). Society for the Teaching of Psychology. <http://teachpsych.org/ebooks/asle2014/index.php>

- Rawson, K. A., & Dunlosky, J. (2011). Optimizing schedules of retrieval practice for durable and efficient learning: How much is enough? *Journal of Experimental Psychology: General*, *140*(3), 283-302. <https://doi.org/10.1037/a0023956>
- Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, *15*(1), 20-27. <https://doi.org/10.1016/j.tics.2010.09.003>
- Roediger, H. L., & Karpicke, J. D. (2006a). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, *1*(2), 181-210. <https://doi.org/10.1111/j.1745-6916.2006.00012.x>
- Roediger, H. L., & Karpicke, J. D. (2006b). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*(3), 249-255. <https://doi.org/10.1111/j.1467-9280.2006.01693.x>
- Roediger, H. L., Putnam, A. L., & Smith, M. A. (2011). Ten benefits of testing and their applications to educational practice. *Psychology of Learning and Motivation*, *55*, 1-36. <https://doi.org/10.1016/B978-0-12-387691-1.00001-6>
- Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin*, *140*(6), 1432-1463. <https://doi.org/10.1037/a0037559>
- Schrank, Z. (2016). An assessment of student perceptions and responses to frequent low-stakes testing in introductory sociology classes. *Teaching Sociology*, *44*(2), 118-127. <https://doi.org/10.1177%2F0092055X15624745>
- Sennhenn-Kirchner, S., Goerlich, Y., Kirchner, B., Notbohm, M., Schiekirka, S., Simmenroth, A., & Raupach, T. (2018). The effect of repeated testing vs repeated practice on skills learning in undergraduate dental education. *European Journal of Dental Education*, *22*(1), e42-e47. <https://doi.org/10.1111/eje.12254>
- Sullivan, D. (2017). Mediating test anxiety through the testing effect in asynchronous, objective, online assessments at the university level. *Journal of Education and Training*, *4*(2), 106-123. <https://doi.org/10.5296/jet.v4i2.10777>
- Szpunar, K. K., Khan, N. Y., & Schacter, D. L. (2013). Interpolated memory tests reduce mind wandering and improve learning of online lectures. *PNAS*, *110*(16), 6313-6317. <https://doi.org/10.1073/pnas.1221764110>
- Tullis, J. G., & Maddox, G. B. (2020). Self-reported use of retrieval practice varies across age and domain. *Metacognition and Learning*, *15*, 129-154. <https://doi.org/10.1007/s11409-020-09223-x>
- Vojdanoska, M., Cranney, J., & Newell, B. R. (2010). The testing effect: The role of feedback and collaboration in a tertiary classroom setting. *Applied Cognitive Psychology*, *24*(8), 1183-1195. <https://doi.org/10.1002/acp.1630>
- Yang, C., Luo, L., Vadillo, M. A., Yu, R., & Shanks, D. R. (2021). Testing (quizzing) boosts classroom learning: A systematic and meta-analytic review. *Psychological Bulletin*, *147*(4), 399-435. <https://doi.org/10.1037/bul0000309>
- Yang, C., Potts, R., & Shanks, D. R. (2017). The forward testing effect on self-regulated study time allocation and metamemory monitoring. *Journal of Experimental Psychology: Applied*, *23*(3), 263-277. <https://doi.org/10.1037/xap0000122>

- Yang, C., & Shanks, D. R. (2018). The forward testing effect: Interim testing enhances inductive learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(3), 485-492. <https://doi.org/10.1037/xlm0000449>
- Yang, C., Sun, B., Potts, R., Yu, R., Luo, L., & Shanks, D. R. (2020). Do working memory capacity and test anxiety modulate the beneficial effects of testing on new learning? *Journal of Experimental Psychology: Applied*, 26(4), 724-738. <https://doi.org/10.1037/xap0000278>