Modulators of Test-Enhanced Learning in Post-Secondary Biology

Bryn St. Clair*and Jamie L. Jensen

Department of Biology
Brigham Young University
Provo, UT 84602

*Corresponding author. Dept. of Biology, 4102 LSB, Brigham Young University, Provo, UT 84602. bes@byu.edu

Abstract

Cognitive scientists and psychology researchers have given growing attention to evidence of the testing effect, that is, the improvement of students' recall through memory-retrieval practice in the form of quizzes and exams. While laboratory experiments consistently show dramatic positive effects on learning through the testing effect, discipline-specific education researchers have sought to generalize these findings in real, instead of simulated classrooms. Our objective in this review was to survey recent findings on the testing effect in post-secondary biology education and synthesize how those findings may modulate learning in the post-secondary biology classroom. We found that: (a) Increased exam frequency increases the testing effect; (b) Corrective feedback on exams may enhance the testing effect; (c) Incentives, such as points, may decrease the positive outcome of learning through the testing effect, though little research in actual classrooms on this widely used practice is found; (d) Individual differences in student achievement and preparation may moderate the effect. We consider how further research on the testing effect may be useful for instructors' decisions regarding its use.

Introduction

In recent decades, science education reform has involved stronger collaboration between cognitive and education researchers interested in applying their findings from laboratory research to subject-specific disciplines. Discipline-based education research (DBER) is central to the effort at the post-secondary level (e.g., Carpenter et al., 2016). Several principles of learning from cognitive-science laboratory studies have hypothetical applications for instructional practice in STEM classrooms. In this article, we review the recent findings on one such principle of learning for application in post-secondary biology education —the testing effect.

The testing effect is the improvement of students' learning through classroom testing. It is also known as retrieval practice, practice quizzing, or test-enhanced learning (for a review see Roediger & Butler, 2011). For example, information that appears on a quiz will more likely be recalled later than information not tested. Terms such as test, exam, and quiz are most commonly associated with assessment in education. In simple language, assessment is the measurement of how much students have learned of what the teacher has taught. In this way, assessment informs teachers of the effectiveness of their instruction in terms of student learning (Black & Wiliam, 1998). Even though an exam on previously presented content is sometimes considered, in itself, a neutral learning event, researchers in cognitive psychology, education and neurobiology have reported that learning may be enhanced when it is retrieved through testing (Roediger & Karpicke, 2006). Specifically, retrieving information from memory may enhance the cues for future retrieval.

A recent resurgence of research on this topic has been largely conducted in cognitive-science laboratories in controlled studies (Roediger & Butler, 2011) using varied material, including word pairs (Carrier & Pashler, 1992; Karpicke & Roediger, 2008), general facts (Butler et al., 2008), trivia (McDaniel & Fisher, 1991), and textual passages (Kang et al., 2007). Researchers have also examined the differential effects of the testing effect using varied assessment formats, including multiple-choice items (Marsh et al., 2007), open- and closed-book items (Agarwal et al., 2008), and inference items (Karpicke, 2012). The results of a multitude of studies on the testing effect have been featured in confirmatory meta-analyses (Bangert-Drowns et al., 1991; Phelps, 2012; Rowland, 2014; Schwieren et al., 2017). Officials at the Institute of Education Sciences, sponsored by the US Department of Education, have recommended the adoption of testing and retrieval practice at all levels of education, including the post-secondary level (Pashler et al., 2007).

The extent to which research on the testing effect applies to learning biology at the post secondary level is critical in making decisions in its use

(Daniel, 2012). Given both the consistent findings on the testing effect in laboratory studies, and researchers' confidence in recommending the method as a means of improving student learning, there is a growing motivation to understand the mechanistic boundaries that may influence the testing effect as it applies to learning in discipline-specific classrooms. Perhaps due to the demand for STEM education reform, the body of testing effect literature in recent years has expanded to include information pertinent to the application of the testing effect in post-secondary biology classrooms. A summary of the testing effect in biology education at the post-secondary level is not found in the current literature. Here we present an overview of the testing effect specifically in post-secondary biology education and discuss the implications for further research and application of the testing effect. The key questions being addressed are:

- How does the classroom structure of assessment influence learning through the testing effect in biology education?
- How does assessment format and content influence learning through the testing effect in biology education?
- What student characteristics influence learning through the testing effect in biology education?

How does the classroom structure of assessment influence the testing effect in biology education?

Assessment frequency

Researchers have found a positive relationship between the higher frequency of classroom assessments and academic achievement (Phelps, 2012). The importance of testing frequency has also been shown more specifically in biology education research at the university-level. For example, frequent quizzing improved learning outcomes in post-secondary biology over standard unit exams (Bailey et al., 2017). Haak et al. (2011) also implemented a highly structured course design in an introductory biology course based on daily and weekly assessments in problem-solving, data analysis, and other higher-order cognitive skills. The design was associated with improved performance in all students enrolled in the course and reduced the performance gap between socioculturally disadvantaged and non-disadvantaged students. In a final example, Pape-Lindstrom et al. (2018) measured an increase in student performance in a community college biology course when they implemented frequent pre-class online and open-book reading assessments.

Overall, increased frequency of testing experiences appears to improve learning. The increased frequency of assessment has been reported as positively correlated with lower course-failure rates, higher course point totals, and higher scores on midterm assessments (Freeman et al., 2007) as well as increased academic motivation (Healy et al., 2017). Importantly, Leeming (2002) surveyed University of Memphis students and reported their greater satisfaction with courses that included more frequent assessments. Students also indicated that they learned more as a result. By simply increasing the frequency and number of exam experiences instructors can enhance the testing effect in the biology classroom for students.

Assessment Incentives

Incentives in terms of assessment scores and points are commonly used to motivate students and may have an influence on the testing effect. However, researchers typically have not treated classroom incentives per se as an experimental variable with regard to the testing effect. Hinze and Rapp (2014) awarded monetary compensation to lab study participants based on performance. Subjects scored relatively lower on high-stakes biology exams than low-stakes biology exams. The current application of incentives in the biology education classroom studies on the testing effect is varied and a clear understanding of the implications with regard to the testing effect is not well- defined in the literature. While researchers intuitively recommend low-stakes quizzing as an important safeguard against student test-anxiety in post-secondary biology, particularly on mid-course formative assessments, little classroom research has addressed this idea experimentally. In one exception, St. Clair et.al. (2020) found no difference between students' performance on exams with high- and low- incentive levels (21% vs 10% of the course points total on exams) in an introductory college-level biology course. Further research on classroom incentives and learning through the testing effect will be important to understand the interaction between more extreme levels of course incentives and the testing effect.

Traditional post-secondary courses frequently assess students in two ways: Formative assessment that communicates learning progress to both a student as well as the instructor

Traditional post-secondary courses frequently assess students in two ways: Formative assessment that communicates learning progress to both a student as well as the instructor, and summative assessment that in generally is performed at the end of a course to measure overall mastery of course material. A course instructor may remove the incentives from a formative assessment and make the experience voluntary. Student self-reported voluntary use of self-testing is correlated with increased student achievement (Hartwig & Dunlosky, 2012). Specifically, in university-level biology, Carpenter et al. (2017) reported that students who opted for quizzes as a review tool in a general-biology course scored higher on the initial examination in the course than those students who selected reading-based review instead. Subsequently, researchers promoted quiz completion using a classroom presentation of the differential outcome on examination performance. Increasingly more students participated in voluntary assessment practice prior to each subsequent examination, producing higher mean scores on the examinations. Others have modeled effective optional learning strategies (Rodriguez et al.,2018) and offered voluntary workshops promoting the testing effect as a learning strategy in large post-secondary biology courses (Stanger-Hall et al., 2011). Both strategies led to improved student learning. Peat and Franklin (2003) found no differences in the learning of students who participated in a voluntary quiz activity and those who did not. Yet, in a follow-up study, Peat et al. (2005) found an increase in mean summative exam score in voluntary second-year student participants compared to first year voluntary student participants.

Assessment Feedback

Informing students of the assessment items they missed and the correct answer to those items is generally referred to as feedback and is an influential factor in the testing effect (Kang et al, 2007). Retrieval practice is effective without feedback, but feedback enhances learning with the testing effect (Pashler et al., 2005; Lavigne & Risko, 2018). Jacoby et al. (2010) found that feedback enabled bird classification with fewer exam experiences. In another study, researchers displayed feedback to subjects following initial fill-in-the-blank items with process-based biology concepts (e.g., stages of mitosis) and found an increase in performance on final fill-in-the blank questions when feedback was given (Pan et al., 2019). The majority of studies on the testing effect and biology learning use feedback as a consistent part of testing, and the empirical evidence from laboratory studies seem to support this practice, yet relatively few studies exist that examine the effects of feedback on the testing effect specifically in the biology classroom. More work in this area is needed to understand the role of feedback on the testing effect in biology education.

How does assessment format and cognitive skill influence the testing effect?

Assessment Format

The initial test format may influence the final test success (Kang et al., 2007). According to Glover (1989), short answer, and fill-in-the-blank item formats both increased the testing effect over multiple-choice and true-or-false formats. However, Little and Bjork (2015) found that multiple-choice items were more effective when they contained strong distractor options and feedback. More specifically, Pagliarulo reported that multiple-choice and short answer assessment formats could be useful on complex biology content (2011). Hinze (2010) assessed post-secondary students on biology content in laboratory experiments and found that cued-recall assessment format (a sentence that includes pertinent content preceding the assessment item) improved performance on memory items while removing the cues from recall items made retrieval practice more difficult and less effective. Presumably, generating information on one's own, if successful, could increase the effectiveness of free-recall assessment items over cued-recall assessment items (Carpenter & DeLosh, 2006), yet there is an inherent balance between increasing effortful processing and overloading a student's ability to successfully do the task (see Pyc & Rawson, 2009). More research needs to be done specifically in post-secondary biology classrooms with regard to item type and the effectiveness of the testing effect to bring about student learning.

Assessment and cognitive skills

Although many instructors seek to develop high-level cognitive thinking in their students, most assessment items are specific to memory retrieval of subject content rather than application, analysis, evaluation, and creativity (Momsen et al., 2010). As such, quiz and exam questions that are related in content subject-material but do not focus on the same specific learning outcome or concept may not show a testing effect (Nguyen & McDaniel, 2015).

biology, it appears that the standard procedure of using quiz questions from the test bank provided by ancillary sources (e.g. textbook companies) may not benefit student performance unless the summative exam questions are closely tied to the targeted learning outcomes created for the course and taught by the instructor (Wooldridge et al., 2014). Instructors should pay specific attention to coordinating intended learning outcomes with assessment items to enable learning through the testing effect.

Researchers argue for the strength of the testing effect with complex material (Karpicke & Aue, 2015; Rawson, 2015; Burns, 2010). Jensen et al. (2014) found routine quizzing requiring application, analysis, and evaluation of biology material could be useful in promoting both conceptual and higher-order skills performance on the final exam in a biology class. Agarwal (2011) reported that a match in initial and final cognitive processing on assessment items (e.g. quizzed and tested on a specific skill) benefits long-term higher-order skills in learning biology. Further research is needed on the testing effect using complex material learning in biology in the post-secondary level classroom including valuable reasoning skills used in scientific discovery and problem solving.

What student characteristics influence the testing effect in biology?

Test Anxiety

Test anxiety is common among undergraduate students. In a survey, Gerwing (2015) found that 38.5% of student respondents reported test anxiety. High test anxiety typically is associated with poorer test performance, test avoidance, loss of motivation, decreased memory retrieval, and impaired attention (Wolf & Smith, 1995; Zeidner, 2005). In the laboratory, Tse and Pu (2012) replicated the testing effect using word pairs while also measuring attention to relevant detail and test anxiety. They found that students who scored lower on attention to relevant detail but higher in test anxiety made more errors on average on the final assessment. England et al. (2017) surveyed learners in undergraduate biology courses that featured active learning pedagogy including in-class clickers. High-test anxiety accompanied lower self-reported GPA and a weaker intention to persist in the biology major.

By contrast, moderate test anxiety may enhance assessment performance (Keeley et al., 2008). A majority of students report decreased test anxiety when they use retrieval practice (Agarwal et al., 2014) and low-stakes in-class quizzing (Khanna, 2015) to prepare for a summative course assessment. A clear picture of the relationship between individual learners' test anxiety, the testing effect, and biology material in a college classroom is weakly defined in the literature, partially due to the variable nature of individual students and their reaction to test experiences. Clearly isolating variables in the ecologically complex classroom is challenging yet needed to further clarify the mechanisms surrounding these commonly experienced pedagogical tools.

Individual Student Differences in Academic Performance

Researchers have begun to study individual student performance differences and the testing effect. While some researchers have demonstrated the benefit of quizzing in biology to students of diverse academic abilities (Orr & Foster, 2013; Pape-Lindstrom et al, 2014), Hubbard and Couch (2018) found that the use of in-class clickers benefited high-performing students more than low-performing students. Carpenter et al. (2016) found that among undergraduate biology learners, all benefited from the use of frequent assessment, but high-performing students benefited more from it than mid- and low-performing students. Bailey et al. (2017) studied increased quiz frequency and categorized students into learning history. They found that mid- and late-learners (those who did not show mastery until the second half of the course) comprised 24% of the class and specifically benefited more from the increased assessment frequency. Butler (2010) found that repeated testing produced improved average success on assessment items with biology inference questions if prior learning of individuals was included in the model. Individual differences in student prior learning and academic ability may impact the outcome of learning biology through the testing effect.

Conclusion

The application of research findings from cognitive psychology to post-secondary classrooms may yield significant benefits in STEM education reform. The evidence supporting the testing effect in particular may enable learning for students in post-secondary biology classrooms. Increased test frequency influences the testing effect in the biology classroom and is the most obvious recommendation for immediate classroom application of the testing effect. Exam feedback has effectively been shown to be an influential moderator of the testing effect in the literature, though no specific study in this search has experimentally applied feedback to a biology classroom.

Course incentives, such as points or stakes also may affect the result of the testing effect in the biology classroom. Current application of incentives in biology education research is varied and there is not a clear understanding of the interaction with the testing effect and points. Cognitive psychology researchers recommend low-stakes quizzing as an important preventive for student test-anxiety in post-secondary biology, though little classroom research has addressed this idea experimentally. Further empirical work on classroom incentives, such as points and learning through the testing effect will be important to understand the interaction.

The influence of individual learner achievement on the testing effect in biology is also a consideration in the success of students in STEM and its application to the postsecondary classroom overall. Students experience test-anxiety, but the influence of anxiety on the testing effect in post-secondary biology is not thoroughly demonstrated. Early work shows that most students can benefit from testing, though it seems that learners embody varied characteristics that may modify learning through the testing effect. Continued experimental application of the testing effect in classroom settings may illuminate mechanistic boundaries to varied learners.

The connection between content, cognitive process and coordinated learning outcomes on the initial and final exam may impact the testing effect. Biology instructors can accentuate exam experiences by ensuring that the required cognitive skills in both formative and summative exam items align with the designated learning outcomes in the course. Practice retrieval, and in many cases, practice processing, on an exam, if successful, could increase the effectiveness of learning through tests, yet there may be an balance between increasing effortful processing and overloading a student's ability to successfully do the task.

Research efforts will continue to illuminate and support reforms in STEM education. The testing effect is a promising principle of learning that has the potential to aid post-secondary biology teaching. Effective instruction needed for deep application and conceptual knowledge in biology education will require further understanding of the mechanistic boundaries of the testing effect as they apply to the biology classroom.

References

Agarwal, P. (2011) Examining the Relationship Between Fact Learning and Higher Order Learning via Retrieval Practice . (Doctoral Dissertation). Retrieved from https://openscholarship.wustl.edu/etd/546

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. doi:10.1016/j.jarmac.2014.07.002

Agarwal, P. K., Karpicke, J. D., Kang, S. H. K., Roediger, H. L., & McDermott, K. B. (2008). Examining the testing effect with open- and closed-book tests. Applied Cognitive Psychology, 22(7), 861–876. doi.org/10.1002/acp.1391

Bailey, E. G., Jensen, J., Nelson, J., Wiberg, H. K., & Bell, J. D. (2017). Weekly Formative Exams and Creative Grading Enhance Student Learning in an Introductory Biology Course. CBE - Life Sciences Education, 16(1), ar2. doi: 10.1187/cbe.16-02-0104

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. doi:10.3102/00346543061002213

Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. Assessment in Education: Principles, Policy & Practice, 5(1), 7–74. doi:10.1080/0969595980050102

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. doi:10.1037/a0019902

Butler, A. C., Karpicke, J. D., & Roediger, H. L. I. (2008). Correcting a metacognitive error: Feedback increases retention of low-confidence correct responses. Journal of Experimental Psychology: Learning, Memory, and Cognition, 34(4), 918–928. doi:10.1037/0278-7393.34.4.918

Carpenter, S.K. & Delosh, E.L. (2006). Impoverished cue support enhances subsequent retention: Support for the elaborative retrieval explanation of the testing effect. Memory & Cognition, 34(2), 268-276. https://doi.org/10.3758/BF03193405

Carpenter, S. K., Lund, T. J., Coffman, C. R., Armstrong, P. I., Lamm, M. H., & Reason, R. D. (2016). A classroom study on the relationship between student achievement and retrieval-enhanced learning. Educational Psychology Review, 28(2), 353–375. doi:10.1186/s41235-017-0078-z

Carpenter, S. K., Rahman, S., Lund, T. J. S., Armstrong, P. I., Lamm, M. H., Reason, R. D., & Coffman, C. R. (2017). Students' Use of Optional Online Reviews and Its Relationship to Summative Assessment Outcomes in Introductory Biology. CBE Life Sciences Education, 16(2). doi:10.1187/cbe.16-06-0205

Carrier, M., & Pashler, H. (1992). The influence of retrieval on retention. Memory & Cognition, 20(6), 633–642. doi:10.3758/BF03202713

Daniel, D. B. (2012). Promising principles: Translating the science of learning to educational practice. Journal of Applied Research in Memory and Cognition, 1(4), 251–253. https://doi.org/10.1016/j.jarmac.2012.10.004

England, B. J., Brigati, J. R., & Schussler, E. E. (2017). Student anxiety in introductory biology classrooms: Perceptions about active learning and persistence in the major. PLOS ONE, 12(8), e0182506. doi;10.1371/journal.pone.0182506

Gerwing, T. G., Rash, J. A., Allen Gerwing, A. M., Bramble, B., & Landine, J. (2015). Perceptions and Incidence of Test Anxiety. Canadian Journal for the Scholarship of Teaching and Learning, 6(3) 1–17. doi:10.5206/cjsotl-rcacea.2015.3.3

Glover, J. A. (1989). The "testing" phenomenon: Not gone but nearly forgotten. Journal of Educational Psychology, 81(3), 392–399. doi:10.1037/0022-0663.81.3.392

Haak, D. C., HilleRisLambers, J., Pitre, E., & Freeman, S. (2011). Increased structure and active learning reduce the achievement gap in introductory biology. Science, 332(6034), 1213–1216. doi:10.1126/science.1204820

Hartwig, M. K., & Dunlosky, J. (2012). Study strategies of college students: Are self-testing and scheduling related to achievement? Psychonomic Bulletin & Review, 19(1), 126–134. doi:10.3758/s13423-011-0181-y

Healy, A. F., Jones, M., Lalchandani, L. A., & Tack, L. A. (2017). Timing of quizzes during learning: Effects on motivation and retention. Journal of Experimental Psychology: Applied, 23(2), 128–137. doi:10.1037/xap0000123

Hinze, S. R. (2010). The influence of retrieval practice on memory and comprehension of science texts (Order No. 3431210). Available from ProQuest Dissertations & Theses Global; Social Science Premium Collection. (807672924). Retrieved from https://lib.byu.edu/remoteauth/?url=https://search.proquest.com/docview/807672924?accountid=4488

Hinze, S. R., & Rapp, D. N. (2014). Retrieval (Sometimes) Enhances Learning: Performance Pressure Reduces the Benefits of Retrieval Practice. Applied Cognitive Psychology, 28(4), 597–606. doi:10.1002/acp.3032

Hubbard, J. K., & Couch, B. A. (2018). The positive effect of in-class clicker questions on later exams depends on initial student performance level but not question format. Computers & Education, 120, 1–12. doi:10.1016/j.compedu.2018.01.008

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. doi:10.1037/a0020636

Jensen, J. L., McDaniel, M. A., Woodard, S. M., & Kummer, T. A. (2014). Teaching to the Test or Testing to Teach: Exams Requiring Higher Order Thinking Skills Encourage Greater Conceptual Understanding. Educational Psychology Review, 26(2), 307–329. doi:10.1007/s10648-013-9248-9

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. doi:10.1080/09541440601056620

Karpicke, J. (2012). Retrieval-Based Learning: Active Retrieval Promotes Meaningful Learning. Current Directions in Psychological Science, 21(3), 157-163. Retrieved from http://www.jstor.org/stable/23213127

Karpicke, J. D., & Aue, W. R. (2015). The Testing Effect Is Alive and Well with Complex Materials. Educational Psychology Review, 27(2), 317–326. doi:10.1007/s10648-015-9309-3

Karpicke, J. D., & Roediger, H. L., III. (2008). Repeated Retrieval during Learning Is the Key to Long-Term Retention. Journal of Memory and Language, 57(2), 151–162. doi:10.1016/j.jml.2006.09.004

Keeley, J., Zayac, R., & Correia, C. (2015). Curvilinear Relationships Between Statistics Anxiety and Performance Among Undergraduate Students: Evidence for Optimal Anxiety. Statistics Education Research Journal, 7(1), 1-15.

Khanna, M. M. (2015). Ungraded Pop Quizzes: Test-Enhanced Learning Without All the Anxiety. Teaching of Psychology, 42(2), 174–178. doi:10.1177/0098628315573144

Lavigne, E., & Risko, E. F. (2018). Optimizing the use of interpolated tests: The influence of interpolated test lag. Scholarship of Teaching and Learning in Psychology, 4(4), 211–221. doi:10.1037/stl0000118

Leeming, F. C. (2002). The Exam-A-Day Procedure Improves Performance in Psychology Classes. Teaching of Psychology, 29(3), 210–212. doi:10.1207/S15328023TOP2903_06

Little, J. L., & Bjork, E. L. (2015). Optimizing multiple-choice tests as tools for learning. Memory & Cognition, 43(1), 14–26. doi:10.3758/s13421-014-0452-8

Marsh, E. J., Roediger, H. L., Bjork, R. A., & Bjork, E. L. (2007). The memorial consequences of multiple-choice testing. Psychonomic Bulletin & Review, 14(2), 194–199. doi:10.3758/BF03194051

McDaniel, M. A., & Fisher, R. P. (1991). Tests and test feedback as learning sources. Contemporary Educational Psychology, 16(2), 192–201. doi:10.1016/0361-476X(91)90037-L

Momsen, J. L., Long, T. M., Wyse, S. A., & Ebert-May, D. (2010). Just the Facts? Introductory Undergraduate Biology Courses Focus on Low-Level Cognitive Skills. CBE—Life Sciences Education, 9(4), 435–440. doi:10.1187/cbe.10-01-0001

Nguyen, K., & McDaniel, M. A. (2015). Using Quizzing to Assist Student Learning in the Classroom: The Good, the Bad, and the Ugly. Teaching of Psychology, 42(1), 87–92. doi:10.1177/0098628314562685

Oliver, M., Renken, M., & Williams, J. J. (2018). Revising Biology Misconceptions Using Retrieval Practice and Explanation Prompts.

Orr, R., & Foster, S. (2013). Increasing Student Success Using Online Quizzing in Introductory (Majors) Biology. CBE-Life Sciences Education, 12(3), 509. doi:10.1187/cbe.12-10-0183

Pagliarulo, C. L. (2011). Testing Effect and Complex Comprehension in a Large Introductory Undergraduate Biology Course (Doctoral Dissertation). Retrieved from https://repository.arizona.edu/handle/10150/202773

Pan, S. C., Hutter, S. A., D'Andrea, D., Unwalla, D., & Rickard, T. C. (2019). In search of transfer following cued recall practice: The case of process-based biology concepts. Applied Cognitive Psychology, 1-17. doi:10.1002/acp.3506

Pape-Lindstrom, P., Eddy, S., & Freeman, S. (2018). Reading Quizzes Improve Exam Scores for Community College Students. CBE Life Sciences Education, 17(2). doi:10.1187/cbe.17-08-0160

Pashler, H., Bain, P., Bottge, B., Graesser, A., Koedinger, K., McDaniel, M., and Metcalfe, J. (2007) Organizing Instruction and Study to Improve Student Learning (NCER 2007-2004). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ncer.ed.gov.

Pashler, H., Cepeda, N. J, Wixted, J. T, & Rohrer, D. (2005). When does feedback facilitate learning of words? Journal of Experimental Psychology-Learning Memory and Cognition, 31(1), 3-8. Retrieved from https://escholarship.org/uc/item/7fg8h6zq

Pyc, M. A., & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? Journal of Memory and Language, 60(4), 437–447. doi.org/10.1016/j.jml.2009.01.004

Rawson, K. A. (2015). The Status of the Testing Effect for Complex Materials: Still a Winner. Educational Psychology Review, 27(2), 327–331. doi:10.1007/s10648-015-9308-4

Roediger, H. L., & Karpicke, J. D. (2006). Test-Enhanced Learning: Taking Memory Tests Improves Long-Term Retention. Psychological Science, 17(3), 249–255. doi:10.1111/j.1467-9280.2006.01693.x

Roediger, H., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. Trends in Cognitive Sciences, 15(1), 20–27. doi:10.1016/j.tics.2010.09.003

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. doi:10.1037/a0037559

Schwieren, J., Barenberg, J., & Dutke, S. (2017). The Testing Effect in the Psychology Classroom: A Meta-Analytic Perspective. Psychology Learning & Teaching, 16(2), 179–196. doi:10.1177/1475725717695149

Stanger-Hall, K. F., Shockley, F. W., & Wilson, R. E. (2011). Teaching Students How to Study: A Workshop on Information Processing and Self-Testing Helps Students Learn. CBE-Life Sciences Education, 10(2), 187–198. doi.org/10.1187/cbe.10-11-0142

St.Clair, B.E., Putnam, P., Miller, H.L., Larsen, R., & Jensen, J.L. (in-press) Test-enhanced learning and assessment incentives: a study in a post-secondary biology course. CBE-LSE

Tse, C.-S., & Pu, X. (2012). The effectiveness of test-enhanced learning depends on trait test anxiety and working-memory capacity. Journal of Experimental Psychology: Applied, 18(3), 253–264. doi:10.1037/a0029190

Wooldridge, C. L., Bugg, J. M., McDaniel, M. A., & Liu, Y. (2014). The testing effect with authentic educational materials: A cautionary note. Journal of Applied Research in Memory and Cognition, 3(3), 214–221. doi:10.1016/j.jarmac.2014.07.001

Wise, S. L., & DeMars, C. E. (2005). Low Examinee Effort in Low-Stakes Assessment: Problems and Potential Solutions. Educational Assessment, 10(1), 1–17. doi:10.1207/s15326977ea1001_1

Wolf, L. F., & Smith, J. K. (1995). The Consequence of Consequence: Motivation, Anxiety, and Test Performance. Applied Measurement in Education, 8(3), 227. doi:10.1207/s15324818ame0803_3

Zeidner, M. (1998). Test Anxiety: The State of the Art. Springer Science & Business Media.