

# Implementing Cognitive Science and Discipline-Based Education Research in the Undergraduate Science Classroom

Ido Davidesco<sup>\*\*\*</sup> and Catherine Milne<sup>†</sup>

<sup>†</sup>Department of Teaching and Learning and <sup>‡</sup>Department of Psychology, New York University, New York, NY 10003

## ABSTRACT

Cognitive science research on learning and instruction is often not directly connected to discipline-based research. In an effort to narrow this gap, this essay integrates research from both fields on five learning and instruction strategies: active retrieval, distributed (spaced) learning, dual coding, concrete examples, and feedback and assessment. These strategies can significantly enhance the effectiveness of science instruction, but they typically do not find their way into the undergraduate classroom. The implementation of these strategies is illustrated through an undergraduate science course for nonmajors called *Science in Our Lives*. This course provides students with opportunities to use scientific information to solve real-world problems and view science as part of everyday life.

## INTRODUCTION

Despite efforts to bridge the gap between research and practice in science teaching, implementing evidence-based pedagogies in undergraduate science classrooms remains a major challenge, often due to faculty's lack of knowledge, insufficient support, and resistance to change (Handelsman *et al.*, 2004; Henderson *et al.*, 2011; Brownell and Tanner, 2012; Dolan, 2015; Drinkwater *et al.*, 2017). Perhaps the most striking example is that, despite mounting evidence of the effectiveness of active learning, many undergraduate science courses still follow a traditional lecture-based model that reinforces students' prior negative experiences with science (National Science Board, 2000; Arum and Roksa, 2011; Freeman *et al.*, 2014).

This essay will demonstrate how learning strategies that emerge from cognitive science and discipline-based education research (DBER) can be implemented in the design of undergraduate science curricula. Owing to different goals, research practices, and terminology, cognitive science research on learning and instruction is often disconnected from DBER (i.e., research on learning within a discipline, such as biology or chemistry; National Research Council, 2012). For example, whereas cognitive science research is typically conducted in laboratory settings, DBER almost always takes place in classrooms (Mestre *et al.*, 2018). Based on recent calls for cross-disciplinary exchange (Coley and Tanner, 2012; Peffer and Renken, 2016; Mestre *et al.*, 2018), this essay aims to integrate cognitive science and DBER on five learning and instruction strategies: active retrieval, distributed (spaced) learning, dual coding, concrete examples, and feedback and assessment. These strategies will be illustrated through an undergraduate course for nonmajors called *Science in Our Lives*. This course was designed as an “umbrella” course with several sections, each taught over a semester, focusing on human health and disease, the human brain, and environmental and public health issues. The course was developed within a phenomenological feminist framework (e.g., Barad, 2007) with the goal of providing students with opportunities to use scientific information to solve real-world problems and view science as part of everyday life. As part of the course, students keep a log of their experiences as they are asking questions, making observations, and generating and interpreting data

Clark Coffman *Monitoring Editor*

Submitted Dec 9, 2018; Revised May 6, 2019;  
Accepted May 20, 2019

CBE Life Sci Educ September 1, 2019 18:es4  
DOI:10.1187/cbe.18-12-0240

\*Address correspondence to: Ido Davidesco  
(ido.davidesco@nyu.edu).

© 2019 I. Davidesco and C. Milne. CBE—Life Sciences Education © 2019 The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution–Noncommercial–Share Alike 3.0 Unported Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/3.0>).

“ASCB®” and “The American Society for Cell Biology®” are registered trademarks of The American Society for Cell Biology.

**TABLE 1. Learning and instruction strategies and their implementation in Science in Our Lives**

Strategy	Implementation in Science in Our Lives
Active retrieval	Weekly reflection cards, where students reflected on their learning and made connections to their day-to-day lives.
Distributed practice	Crosscutting concepts (e.g., acidity) were distributed across the semester and revisited in different contexts.
Dual coding	Integrating pictures and text in snaplogs; creating comics (Figure 1).
Concrete examples	Thinking about abstract concepts using real-world objects (Figure 2).
Feedback and assessment	Self-, peer- and coassessment, with students revising their work based on feedback.

(Milne, 2019). The use of student artifacts in this essay has been approved by the New York University Institutional Review Board committee (exempt; IRB-FY2018-2218).

The list of learning and instruction strategies that follows is far from exhaustive. We only discuss strategies that have been replicated across studies and can directly inform curricular design. We first review the evidence supporting each strategy and then demonstrate how it can be implemented through illustrative examples derived from *Science in Our Lives* (see Table 1). It should be emphasized, though, that we have not assessed how implementing these strategies affected student learning, which should be addressed in future research.

### ACTIVE RETRIEVAL

Testing is traditionally viewed as a way to measure learning, but a large body of research suggests that the act of taking a test can significantly enhance learning (for a review, see Brame and Biel, 2015; Karpicke, 2017). In a highly cited study, undergraduate students were first asked to read short passages and then to either review the passage three more times (SSSS group) or complete a recall task three times, writing down as much information from the passage as they could (STTT group). Memory retention was tested either 5 minutes or a week later. Repeated study led to better retention when the final test was administered immediately after the study phase. However, when the final test was delayed by a week, testing was much more effective: the STTT group outperformed the SSSS group by 21% (Roediger and Karpicke, 2006). These findings were shown to generalize across learner characteristics, materials, and educational contexts (Brame and Biel, 2015; Karpicke, 2017).

Why is retrieval more effective than restudying? The effort invested in retrieving information from memory can strengthen knowledge and make it more easily accessible (Bjork and Bjork, 2011). Another explanation is that, when students are provided with a cue and requested to recall target information, they generate additional memory items that are semantically related to the cue, resulting in better recall in a subsequent test (Carpenter, 2009). More recently, it has been proposed that retrieval updates the episodic context in which information is encoded, making it easier to recover the information in a subsequent test (Karpicke *et al.*, 2014).

In the context of higher education, it has been demonstrated that frequent classroom quizzing can enhance students' performance at an end-of-semester exam (Leeming, 2002; Lyle and Crawford, 2011; McDaniel *et al.*, 2012; Orr and Foster, 2013; Pennebaker *et al.*, 2013; Butler *et al.*, 2014; Batsell *et al.*, 2017) (for a review, see Brame and Biel, 2015). *Science in our Lives*, however, was intentionally designed to move away from traditional testing (Schinske and Tanner, 2014; see *Feedback and Assessment*). How can retrieval practice be implemented in a

course that does not include any quizzes or exams? Instead of weekly quizzes, students in the course were asked to post weekly reflections on the course online forum, where they reflected on their learning and made connections between course topics and their everyday experiences. For example, when learning about air quality, students could reflect on how air quality might be tested in their own environments and what impact air quality could have on their health. This process encouraged students to actively retrieve information that was taught in class and make meaningful connections between different topics. In the spirit of active retrieval, students were encouraged to complete this task without reviewing their class notes. These weekly reflections led to a culminating event at the end of the semester, when students were invited to reflect on their learning as a whole (see *Feedback and Assessment*).

Another way in which active retrieval can be incorporated into a college course is by using clicker systems (Caldwell, 2007; Martyn, 2007; Mayer *et al.*, 2009). Clicker systems allow instructors to record students' responses to multiple-choice questions and provide immediate feedback by displaying the distribution of answers and discussing the correct answer. In addition, instead of using traditional tests, instructors can use ungraded (Khanna, 2015) or collaborative quizzing (Wissman and Rawson, 2016). For example, students can take a short quiz and then exchange quizzes and grade each other or have a discussion with a partner and then go back to their quizzes and revise their answers. These quizzes do not necessarily need to be collected by the instructor (Pandey and Kapitanoff, 2011). Finally, drawing concept maps from memory can be another effective way to retrieve information and reinforce learning (Blunt and Karpicke, 2014). This strategy was used in *Science in Our Lives*, as students were asked to work in teams to build a mind map of the course. This activity was designed to encourage students to think back on what they had learned and make connections between discrete but related topics.

### DISTRIBUTED (SPACED) PRACTICE

Although students tend to "cram" their learning just before a final, distributing learning over time has been shown across many studies to produce better long-term retention than massed learning (Cepeda *et al.*, 2006; Benjamin and Tullis, 2010; Delaney *et al.*, 2010). In a meta-analysis of 254 studies, Cepeda *et al.* (2006) reported that performance on a final recall test was significantly higher after spaced study (47%) compared with massed study (37%). This effect is robust to different lags, student characteristics, learning materials, and criterion tasks (Dunlosky *et al.*, 2013).

Most of these studies, however, were conducted in laboratory environments. One of the few field-based studies compared two sections of a statistics undergraduate course, one taught

over 6 months and the other over 8 weeks (Budé *et al.*, 2011). The 6-month course consisted of the same elements as the 8-week course, but the elements were far more spaced out in time. Scores on a conceptual understanding test as well as on the final test were significantly higher in the 6-month course than those of the 8-week course. Importantly, the two groups achieved comparable scores on a control test in another course, thus demonstrating that the effect was not driven by pre-existing differences between the groups (Budé *et al.*, 2011). In another study, introducing daily or weekly preparatory and review assignments in an introduction to biology undergraduate course led to improved student achievement, especially among students from underrepresented racial/ethnic groups. These assignments were hypothesized to help students distribute their learning throughout the semester and engage with content both before and after class (Eddy and Hogan, 2014).

The distributed practice effect can be explained in various ways. Related to the testing effect, when learning is spaced over time, each learning session reminds the learner of the previous session, leading to active retrieval. Further, retrieving information that has been learned in the past requires more effort, which can strengthen the memory trace. In contrast, retrieving information that has just been learned can give students the false impression that they know the material better than they actually do (Bahrick and Hall, 2005).

In *Science in Our Lives*, learning was organized along several crosscutting concepts that were distributed across the semester, such that each time a topic was revisited, students were already familiar with it to some extent. For example, the concept of “acidity” was first introduced in week 3, when students tested the pH of fruit. Then, at home, they observed how slices of an apple turn brown and explored the science behind this phenomenon. In week 6, students made their own pH indicator from red cabbage and learned about acidity and its involvement in digestion. This concept was then revisited in week 11, when students learned about ocean acidification and observed how seashells dissolve in an acidic solution.

Another crosscutting concept was “using measurement to make observations.” The goal here was to help students understand that, although humans make observations of the world every day of their lives, science values observations in very specific ways: first, observations are purposeful, as science seeks to address a specific question; and second, instruments should be calibrated such that observations can be compared across contexts (e.g., comparing ozone levels inside and outside buildings). To address this notion of measurement and observation, throughout the course, students constructed different tools (e.g., camera obscura, spectroscope, and Schoenbein ozone-testing paper). As they used these tools to collect data, they were encouraged to note how the phenomena that they observed varied based on the instruments with which they interacted. They learned how tools could be calibrated to compare observations of similar phenomena across different contexts. These experiences were distributed throughout the entire course with the hope that students would gain confidence in their ability to collect and interpret data.

Student assessment was also distributed throughout the semester. Because students tend to cram for exams and are often unaware of the benefits of distributed learning (Pyc and Rawson, 2012; Wissman *et al.*, 2012), instead of having one

final test, various forms of assessment were used across the semester. These included weekly reflection cards, lab reports, and classroom presentations (see *Feedback and Assessment*).

## DUAL CODING

Pictures are more likely to be remembered than words, a phenomenon called “the picture superiority effect” (Paivio and Csapo, 1969, 1973). The explanation for this phenomenon is that pictures are dually encoded—using both verbal and image codes—whereas words are primarily coded verbally. Image and verbal codes are processed by separate brain systems, and therefore they have additive and independent effects on memory retention (Paivio, 1986; Mayer and Moreno, 2003; Weinstein *et al.*, 2018). This finding led to the development of dual-coding theory, according to which encoding the same information using multiple representations enhances learning and memory (Paivio, 1971, 1986). Dual-coding theory has direct implications for teaching and learning: instructors should encourage students develop multiple representations (e.g., auditory and visual; image and text) of the same material (Sadoski, 2005; Weinstein *et al.*, 2018).

Dual coding should not be confused with the notion of “learning styles.” Recent surveys conducted in several different countries demonstrated that a vast majority of teachers believe that students learn best when information is presented in their preferred modality (i.e., visual, auditory, or tactile; Howard-Jones, 2014). This common belief, however, is not supported by research (Pashler *et al.*, 2008; Rohrer and Pashler, 2012; Howard-Jones, 2014). For example, in a recent study, participants’ preference for verbal and visual information was assessed using a questionnaire. They then learned a list of picture pairs and a list of word pairs while providing subjective ratings of their learning. It was found that individuals who learned information in their preferred style thought that they had learned better, but their objective performance was unaffected (Knoll *et al.*, 2017). In contrast to learning styles, dual coding suggests that multiple representations of the same information can enhance learning, regardless of individual preferences (Weinstein *et al.*, 2018).

Throughout their learning in *Science in Our Lives*, students were encouraged to integrate pictures and text. Students were asked to use their phones to take pictures of what they were observing and then post these pictures on the course website and explain their findings in their own words. In the last semester, we replaced the handwritten weekly reflections with snaplogs (Bramming *et al.*, 2012). In the spirit of the popular messaging app, students took pictures of their classroom activities and added text to describe these activities (see Figure 1).

Dual coding is not limited to accompanying text with pictures. It was recently demonstrated that creating drawings of to-be-remembered information improves memory retention more than writing. In line with dual-coding theory, it has been claimed that drawing facilitates the integration of semantic, visual, and motor aspects of a memory trace (Wammes *et al.*, 2016). In science classrooms, students are typically only asked to interpret visual information rather than create their own representations. Drawing is important, because it is integral to the practice of science. It can help students enhance their observational skills and better construct their knowledge (for a review, see Quillin and Thomas, 2015). Accordingly, in *Science in Our*

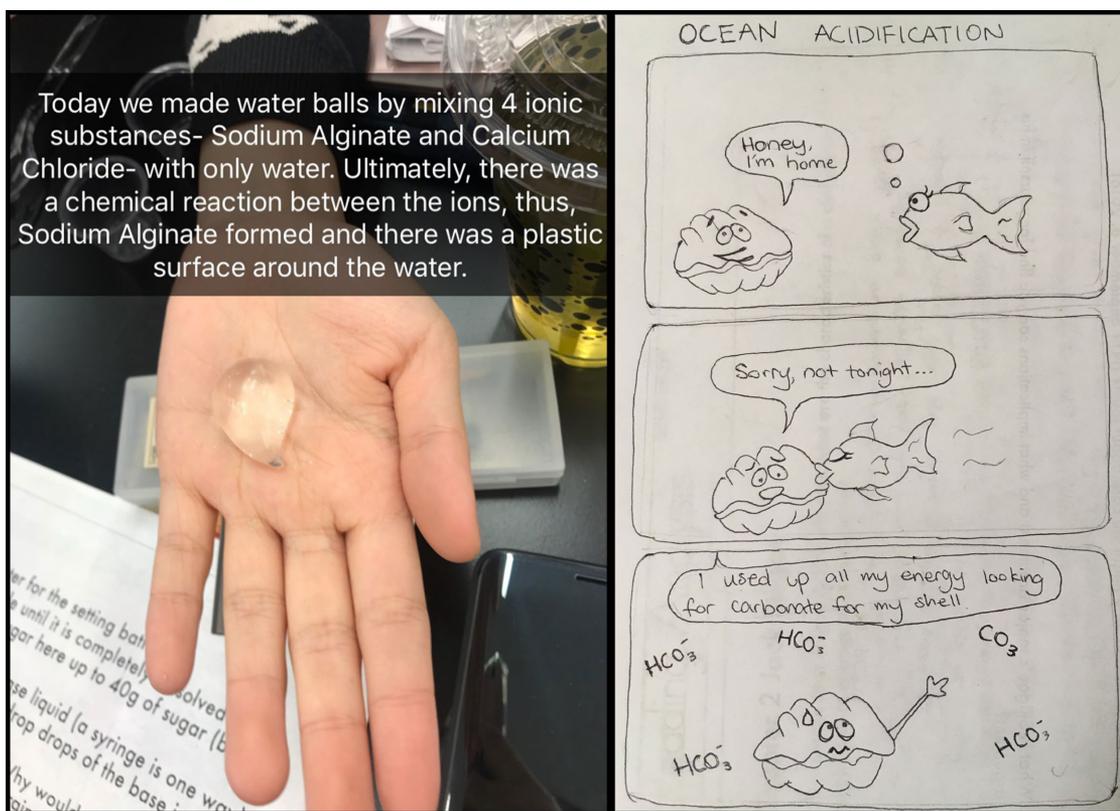


FIGURE 1. Integrating text with images. Left, an example of a snaplog; right, comics made by a student to illustrate ocean acidification.

Lives, students created their own science notebooks at the beginning of the semester. These notebooks were used for note-taking (in light of research that shows that handwritten notes are advantageous compared with laptop note-taking; Mueller and Oppenheimer, 2014) as well as for drawing. In lab reports, students were asked to accompany their explanations with drawings and sketches and, before constructing an instrument (e.g., camera obscura), students drew a model and labeled it.

Comic books are another good example of dual coding. They allow for the integration of text and images into a coherent story. Comics can engage students, motivate them to learn, and boost memory (Tatalovic, 2009). In light of research on the value of comic books in engaging students in science learning (Hosler and Boomer, 2011), students in Science in Our Lives were encouraged to create comics to demonstrate their learning (see Figure 1 for an example).

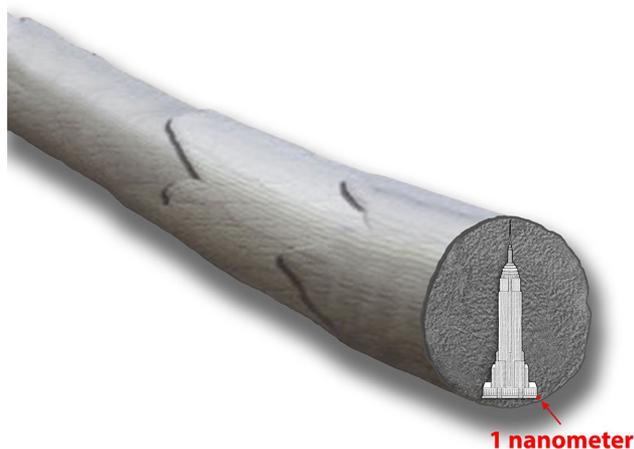
### CONCRETE EXAMPLES

Instructors and textbooks typically provide students with concrete examples to support the comprehension and retention of abstract concepts. This practice is consistent with research in cognitive psychology. This body of research suggests that concrete words are better remembered than abstract words (Gorman, 1961) and that it is easier to learn associations between high-imaginability words (e.g., “car” and “gate”) than low-imaginability words (like “aspect” and “fraction”; Madan *et al.*, 2010; Caplan and Madan, 2016). Concrete examples

can activate prior knowledge, which can, in turn, facilitate learning (Reed and Evans, 1987). Indeed, it has been demonstrated that college students learn concepts better with illustrative examples compared with additional study of the concept definitions. Interestingly, performance was comparable when the definitions preceded the examples and when the examples preceded the definitions (Rawson *et al.*, 2015). A related topic of investigation is how effective examples provided by an instructor are compared with examples generated by the student. In a recent study, for each concept that was learned, students either received four examples, generated four examples, or received two examples and generated two examples. Across two experiments, the results indicated that concept learning was better for provided examples (Zamary and Rawson, 2018).

Can concrete examples hinder the transfer of knowledge? One concern when using concrete examples is that students will remember the surface details of an example rather than the abstract concept. Kaminski *et al.* (2008) argued that, in the case of math learning, abstract examples are better than concrete examples, as concrete examples make it harder for students to apply their knowledge to new contexts. However, there is some controversy about the distinction between concrete and abstract examples in this study (Reed, 2008). It seems that, rather than a dichotomy, there is a continuum between concrete examples and abstract concepts, and more research is needed to clarify the ideal point along this continuum (Weinstein *et al.*, 2018).

In Science in Our Lives, as the name of the course suggests, we tried to make science concrete and relevant to students’



**FIGURE 2.** Thinking about nanoscale using everyday objects. If we scale down the Empire State Building to the width of a human hair, a nanometer will be just a quarter of an inch dot at the bottom of the skyscraper. Based on: [www.youtube.com/watch?v=IC3ActKc3U](http://www.youtube.com/watch?v=IC3ActKc3U).

everyday lives. For example, as part of a unit on vision, students learned that the wavelength of light is measured in nanometers. Acknowledging the challenges involved in teaching students about nanoscale (Jones *et al.*, 2013), we asked students to visualize the wavelength of light using real-life objects (see Figure 2). When learning about air pollution, students calculated their own carbon footprints (e.g., as a function of the car that they drive and how many flights they took in the past year), and discussed ways to reduce their negative impact. Finally, in their weekly reflections, students were specifically instructed to generate examples from their daily lives and relate them to topics learned in class.

### FEEDBACK AND ASSESSMENT

Feedback can have a major impact on students' learning and skill development (Hattie and Timperley, 2007). However, feedback and assessment are typically the topics with which university students are least satisfied. For example, in the 2017 national student survey in the United Kingdom, only 73% were satisfied with the quality of assessment and feedback they receive in their universities (Higher Education Funding Council for England, 2017). Paradoxically, students recognize the importance of feedback, but they tend not to use it effectively to feed forward (Withey, 2013).

What are the barriers that prevent university students from effectively engaging with feedback? On the basis of student interviews, Winstone *et al.* (2017) suggested four psychological processes that can inhibit the use of feedback: 1) awareness of what the feedback means and its purpose; 2) cognizance of strategies by which the feedback could be used; 3) agency to implement strategies; and 4) volition to examine feedback and implement it. Students often have unrealistic expectations about feedback. They desire immediate transfer of feedback, rather than focus on long-term goals (Price *et al.*, 2010). They also expect feedback that lists exactly what they should do. Additionally, students report difficulties understanding

feedback due to the use of terminology (Winstone *et al.*, 2016, 2017).

Two common problems students encounter with feedback are lack of opportunities to implement the feedback and limited transferability of feedback to future work (Gleaves *et al.*, 2008). If feedback is only given on a final submission, students cannot use it to improve the quality of their work. Furthermore, even if feedback is provided throughout the course, students often find it difficult to use previous feedback to inform their future work.

In *Science in Our Lives*, students received regular feedback from the instructor. They received feedback on their weekly reflection cards and snaplogs, where they were encouraged to think more deeply about a scientific phenomenon and how it is reflected in their day-to-day lives. Students also received detailed feedback on their lab reports and were requested to revise and resubmit their work according to the instructor's feedback. All lab reports were structured in the same way and had the same requirements to help students transfer what they had learned from the instructor's feedback.

Furthermore, based on prior research, the course emphasized self-assessment. Self-assessment leads to more reflection on students' own work and higher sense of responsibility for the learning process (Mahlberg, 2015). Even though the accuracy of self-assessment can be a concern, it seems to improve over time, especially when students receive feedback from the instructor on their self-assessments (Dochy *et al.*, 1999; Simkin, 2015). Accordingly, students in *Science in Our Lives* were asked to submit weekly reflections on their learning and received feedback from the instructor on these reflections. At the end of the semester, students prepared a presentation in a Pecha Kucha style (i.e., 20 slides that are presented for 20 seconds each), in which they summarized and reflected on what they have learned throughout the course. They were also given the opportunity to grade their course performances, but they were asked to justify their grades and provide sufficient evidence.

The course also emphasized peer feedback. Research on peer assessment in higher education indicates that it can be a valuable assessment tool and can promote students' involvement. Students tend to perceive peer assessment as sufficiently fair and accurate, even though it can be influenced, for example, by friendship marking (Dochy *et al.*, 1999; Liu and Carless, 2006; Freeman and Parks, 2010; Panadero *et al.*, 2013). There is also evidence that peer assessment can lead to better course performance (Pelaez, 2002; Sun *et al.*, 2015). A recent study reported that peer- and self-grading of practice exams in an undergraduate introduction to biology course can be equally effective, and thus they can be implemented side-by-side (Jackson *et al.*, 2018). Accordingly, students in *Science in Our Lives* were asked not to grade one another, but instead to provide feedback on one another's work (Schinske and Tanner, 2014). For example, they were expected to post comments on their peers' weekly reflections on the course website and provide feedback on one another's research projects.

### CONCLUSIONS AND NEXT STEPS

This essay highlighted several learning and instruction strategies that have been studied both within cognitive science and DBER. These strategies can directly inform the design of undergraduate science courses and can potentially improve student

learning. This essay outlined how these strategies can be implemented in an undergraduate science course, but the effectiveness of the proposed implementation has not yet been assessed and should be addressed in future studies.

In this essay, we made an effort to integrate research findings from both cognitive science and DBER. Whereas DBER is typically conducted in classrooms within a specific discipline, cognitive science research is mostly done in controlled laboratory environments. The difference in theoretical and methodological approaches limits the communication and collaboration across the cognitive science research and DBER communities (Coley and Tanner, 2012; Peffer and Renken, 2016). As acknowledged by cognitive scientists, more field-based research is needed to validate the efficacy of learning strategies that arise from laboratory research (Dunlosky *et al.*, 2013; Weinstein *et al.*, 2018). Therefore, we believe that collaborations between cognitive scientists and discipline-based education researchers are crucial.

It should be noted that the strategies highlighted in this essay can be combined to amplify their impact. For example, spaced retrieval practice is a technique that integrates spaced learning and retrieval practice: students actively retrieve material that was previously learned rather than reviewing it passively (Kang, 2016). There is evidence that spaced retrieval practice (with feedback) is more effective than spaced rereading for middle school students (Carpenter *et al.*, 2009). Similarly, concrete examples can be presented both visually and verbally, taking advantage of dual coding (Weinstein *et al.*, 2018).

It should also be stressed that the list of strategies reviewed here is far from being exhaustive. We chose to focus on strategies that were found to be effective both in laboratory settings and in classrooms. We also chose to highlight strategies that are typically not implemented in undergraduate science classrooms (Dunlosky *et al.*, 2013). Faculty are typically not trained in evidence-based pedagogies, and due to lack of time and incentives, many of them resort to familiar lecture-based practices (Brownell and Tanner, 2012; Dolan, 2015; Kang, 2016). Therefore, wider dissemination of cognitive science and DBER in undergraduate education is essential.

## ACKNOWLEDGMENTS

This work was supported by the National Science Foundation under Grant #1661016. We thank E. Laurent for her assistance in preparing the manuscript.

## REFERENCES

- Arum, R., & Roksa, J. (2011). *Academically adrift: Limited learning on college campuses*. Chicago: University of Chicago Press.
- Bahrhik, H. P., & Hall, L. K. (2005). The importance of retrieval failures to long-term retention: A metacognitive explanation of the spacing effect. *Journal of Memory and Language*, 52(4), 566–577.
- Barad, K. (2007). *Meeting the universe halfway: Quantum physics and the entanglement of matter and meaning*. Durham, NC: Duke University Press.
- Batsell, W. R., Jr., Perry, J. L., Hanley, E., & Hostetter, A. B. (2017). Ecological validity of the testing effect: The use of daily quizzes in introductory psychology. *Teaching of Psychology*, 44(1), 18–23.
- Benjamin, A. S., & Tullis, J. (2010). What makes distributed practice effective? *Cognitive Psychology*, 61(3), 228–247.
- Bjork, E., & Bjork, R. (2011). Making things hard on yourself, but in a good way. In Gernsbacher, M. A., Pew, R. W., Hough, L. M., & Pomerantz, J. R. (Eds.), *Psychology in the real world* (pp. 56–64). New York: Worth.
- Blunt, J. R., & Karpicke, J. D. (2014). Learning with retrieval-based concept mapping. *Journal of Educational Psychology*, 106(3), 849–858.
- Brame, C. J., & Biel, R. (2015). Test-enhanced learning: The potential for testing to promote greater learning in undergraduate science courses. *CBE—Life Sciences Education*, 14(2), es4.
- Bramming, P., Gorm Hansen, B., Bojesen, A., & Gylling Olesen, K. (2012). (Im)perfect pictures: Snaplogs in performativity research. *Qualitative Research in Organizations and Management: An International Journal*, 7(1), 54–71.
- Brownell, S. E., & Tanner, K. D. (2012). Barriers to faculty pedagogical change: Lack of training, time, incentives, and... tensions with professional identity? *CBE—Life Sciences Education*, 11(4), 339–346.
- Budé, L., Imbos, T., van de Wiel, M. W., & Berger, M. P. (2011). The effect of distributed practice on students' conceptual understanding of statistics. *Higher Education*, 62(1), 69–79.
- Butler, A. C., Marsh, E. J., Slavinsky, J., & Baraniuk, R. G. (2014). Integrating cognitive science and technology improves learning in a STEM classroom. *Educational Psychology Review*, 26(2), 331–340.
- Caldwell, J. E. (2007). Clickers in the large classroom: Current research and best-practice tips. *CBE—Life Sciences Education*, 6(1), 9–20.
- Caplan, J. B., & Madan, C. R. (2016). Word imageability enhances association-memory by increasing hippocampal engagement. *Journal of Cognitive Neuroscience*, 28(10), 1522–1538.
- Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(6), 1563.
- Carpenter, S. K., Pashler, H., & Cepeda, N. J. (2009). Using tests to enhance 8th grade students' retention of US history facts. *Applied Cognitive Psychology*, 23(6), 760–771.
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, 132(3), 354–380.
- Coley, J. D., & Tanner, K. D. (2012). Common origins of diverse misconceptions: Cognitive principles and the development of biology thinking. *CBE—Life Sciences Education*, 11(3), 209–215.
- Delaney, P. F., Verkoeijen, P. P., & Spigler, A. (2010). Spacing and testing effects: A deeply critical, lengthy, and at times discursive review of the literature. In Ross, B. H. (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 53, pp. 63–147). San Diego, CA: Elsevier/Academic Press.
- Dochy, F., Segers, M., & Sluismans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education*, 24(3), 331–350.
- Dolan, E. L. (2015). Biology education research 2.0. *CBE—Life Sciences Education*, 14, ed1.
- Drinkwater, M. J., Matthews, K. E., & Seiler, J. (2017). How is science being taught? Measuring evidence-based teaching practices across undergraduate science departments. *CBE—Life Sciences Education*, 16(1), ar18.
- Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, 14(1), 4–58.
- Eddy, S. L., & Hogan, K. A. (2014). Getting under the hood: How and for whom does increasing course structure work? *CBE—Life Sciences Education*, 13(3), 453–468.
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences USA*, 111(23), 8410–8415.
- Freeman, S., & Parks, J. W. (2010). How accurate is peer grading? *CBE—Life Sciences Education*, 9(4), 482–488.
- Gleaves, A., Walker, C., & Grey, J. (2008). Using digital and paper diaries for assessment and learning purposes in higher education: A case of critical reflection or constrained compliance? *Assessment & Evaluation in Higher Education*, 33(3), 219–231.
- Gorman, A. M. (1961). Recognition memory for nouns as a function of abstractness and frequency. *Journal of Experimental Psychology*, 61(1), 23–29.

- Handelsman, J., Ebert-May, D., Beichner, R., Bruns, P., Chang, A., DeHaan, R., ... & Wood, W. B. (2004). Scientific teaching. *Science*, *304*(5670), 521–522.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112.
- Henderson, C., Beach, A., & Finkelstein, N. (2011). Facilitating change in undergraduate STEM instructional practices: An analytic review of the literature. *Journal of Research in Science Teaching*, *48*(8), 952–984.
- Higher Education Funding Council for England. (2017). *National student survey results 2017*. Retrieved January 10, 2019, from <https://webarchive.nationalarchives.gov.uk/20180319125655/http://www.hefce.ac.uk/It/nss/results/2017/>
- Hosler, J., & Boomer, K. (2011). Are comic books an effective way to engage nonmajors in learning and appreciating science? *CBE—Life Sciences Education*, *10*(3), 309–317.
- Howard-Jones, P. A. (2014). Neuroscience and education: Myths and messages. *Nature Reviews Neuroscience*, *15*(12), 817–822.
- Jackson, M. A., Tran, A., Wenderoth, M. P., & Doherty, J. H. (2018). Peer vs. self-grading of practice exams: Which is better? *CBE—Life Sciences Education*, *17*(3), ar44.
- Jones, M. G., Blonder, R., Gardner, G. E., Albe, V., Falvo, M., & Chevrier, J. (2013). Nanotechnology and nanoscale science: Educational challenges. *International Journal of Science Education*, *35*(9), 1490–1512.
- Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2008). The advantage of abstract examples in learning math. *Science*, *320*(5875), 454–455.
- Kang, S. H. (2016). Spaced repetition promotes efficient and effective learning: Policy implications for instruction. *Policy Insights from the Behavioral and Brain Sciences*, *3*(1), 12–19.
- Karpicke, J. D. (2017). Retrieval-based learning: A decade of progress. In Wixted, J. T. (Ed.), *Cognitive psychology of memory, Vol. 2 of Learning and memory: A comprehensive reference* (Byrne, J. H., Series Ed.) (pp. 487–514). Oxford, UK: Academic Press.
- Karpicke, J. D., Lehman, M., & Aue, W. R. (2014). Retrieval-based learning: An episodic context account. In Ross, B. H. (Ed.), *Psychology of learning and motivation* (Vol. 61, pp. 237–284). San Diego, CA: Academic Press.
- Khanna, M. M. (2015). Ungraded pop quizzes: Test-enhanced learning without all the anxiety. *Teaching of Psychology*, *42*(2), 174–178.
- Knoll, A. R., Otani, H., Skeel, R. L., & Van Horn, K. R. (2017). Learning style, judgements of learning, and learning of verbal and visual information. *British Journal of Psychology*, *108*(3), 544–563.
- Leeming, F. C. (2002). The exam-a-day procedure improves performance in psychology classes. *Teaching of Psychology*, *29*(3), 210–212.
- Liu, N.-F., & Carless, D. (2006). Peer feedback: The learning element of peer assessment. *Teaching in Higher Education*, *11*(3), 279–290.
- Lyle, K. B., & Crawford, N. A. (2011). Retrieving essential material at the end of lectures improves performance on statistics exams. *Teaching of Psychology*, *38*(2), 94–97.
- Madan, C. R., Glaholt, M. G., & Caplan, J. B. (2010). The influence of item properties on association-memory. *Journal of Memory and Language*, *63*(1), 46–63.
- Mahlberg, J. (2015). Formative self-assessment college classes improves self-regulation and retention in first/second year community college students. *Community College Journal of Research and Practice*, *39*(8), 772–783.
- Martyn, M. (2007). Clickers in the classroom: An active learning approach. *Educuse Quarterly*, *30*(2), 71–74.
- Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, *38*(1), 43–52.
- Mayer, R. E., Stull, A., DeLeeuw, K., Almeroth, K., Bimber, B., Chun, D., ... & Zhang, H. (2009). Clickers in college classrooms: Fostering learning with questioning methods in large lecture classes. *Contemporary Educational Psychology*, *34*(1), 51–57.
- McDaniel, M. A., Wildman, K. M., & Anderson, J. L. (2012). Using quizzes to enhance summative-assessment performance in a Web-based class: An experimental study. *Journal of Applied Research in Memory and Cognition*, *1*(1), 18–26.
- Mestre, J. P., Cheville, A., & Herman, G. L. (2018). Promoting DBER—cognitive psychology collaborations in STEM education. *Journal of Engineering Education*, *107*(1), 5–10.
- Milne, C. (2019). Intra-actions that matter: Building for practice in a liberal arts science course. In Milne, C., & Scantlebury, K. (Eds.), *Material practice and materiality: Too long ignored in Science Education?* (pp. 64–82). Dordrecht, Netherlands: Springer.
- Mueller, P. A., & Oppenheimer, D. M. (2014). The pen is mightier than the keyboard: Advantages of longhand over laptop note taking. *Psychological Science*, *25*(6), 1159–1168.
- National Research Council. (2012). *Discipline-based education research: Understanding and improving learning in undergraduate science and engineering*. Washington, DC: National Academies Press.
- National Science Board. (2000). *Science and engineering indicators—2000*. Washington, DC: U.S. Government Printing Office.
- Orr, R., & Foster, S. (2013). Increasing student success using online quizzing in introductory (majors) biology. *CBE—Life Sciences Education*, *12*(3), 509–514.
- Paivio, A. (1971). Imagery and language. In Segal, S. J. (Ed.), *Imagery: Current cognitive approaches* (pp. 7–32). New York: Academic Press.
- Paivio, A. (1986). *Mental representations: A dual code approach*. New York: Oxford University Press.
- Paivio, A., & Csapo, K. (1969). Concrete image and verbal memory codes. *Journal of Experimental Psychology*, *80*(2), 279–285.
- Paivio, A., & Csapo, K. (1973). Picture superiority in free recall: Imagery or dual coding? *Cognitive Psychology*, *5*(2), 176–206.
- Panadero, E., Romero, M., & Srijbos, J.-W. (2013). The impact of a rubric and friendship on peer assessment: Effects on construct validity, performance, and perceptions of fairness and comfort. *Studies in Educational Evaluation*, *39*(4), 195–203.
- Pandey, C., & Kapitanoff, S. (2011). The influence of anxiety and quality of interaction on collaborative test performance. *Active Learning in Higher Education*, *12*(3), 163–174.
- Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2008). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest*, *9*(3), 105–119.
- Peffer, M., & Renken, M. (2016). Practical strategies for collaboration across discipline-based education research and the learning sciences. *CBE—Life Sciences Education*, *15*(4), es11.
- Pelaez, N. J. (2002). Problem-based writing with peer review improves academic performance in physiology. *Advances in Physiology Education*, *26*(3), 174–184.
- Pennebaker, J. W., Gosling, S. D., & Ferrell, J. D. (2013). Daily online testing in large classes: Boosting college performance while reducing achievement gaps. *PLoS ONE*, *8*(11), e79774.
- Price, M., Handley, K., Millar, J., & O'Donovan, B. (2010). Feedback: All that effort, but what is the effect? *Assessment & Evaluation in Higher Education*, *35*(3), 277–289.
- Pyc, M. A., & Rawson, K. A. (2012). Are judgments of learning made after correct responses during retrieval practice sensitive to lag and criterion level effects? *Memory & Cognition*, *40*(6), 976–988.
- Quillin, K., & Thomas, S. (2015). Drawing-to-learn: A framework for using drawings to promote model-based reasoning in biology. *CBE—Life Sciences Education*, *14*(1), es2.
- Rawson, K. A., Thomas, R. C., & Jacoby, L. L. (2015). The power of examples: Illustrative examples enhance conceptual learning of declarative concepts. *Educational Psychology Review*, *27*(3), 483–504.
- Reed, S. K. (2008). Concrete examples must jibe with experience. *Science*, *322*(5908), 1632–1633.
- Reed, S. K., & Evans, A. C. (1987). Learning functional relations: A theoretical and instructional analysis. *Journal of Experimental Psychology: General*, *116*(2), 106–118.
- Roediger, H. L., III, & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*(3), 249–255.
- Rohrer, D., & Pashler, H. (2012). Learning styles: Where's the evidence? *Online Submission*, *46*(7), 634–635.
- Sadoski, M. (2005). A dual coding view of vocabulary learning. *Reading & Writing Quarterly*, *21*, 221–238.
- Schinske, J., & Tanner, K. (2014). Teaching more by grading less (or differently). *CBE—Life Sciences Education*, *13*(2), 159–166.

- Simkin, M. G. (2015). Should you allow your students to grade their own homework? *Journal of Information Systems Education*, 26(2), 147–153.
- Sun, D. L., Harris, N., Walther, G., & Baiocchi, M. (2015). Peer assessment enhances student learning: The results of a matched randomized crossover experiment in a college statistics class. *PLoS ONE*, 10(12), e0143177.
- Tatalovic, M. (2009). Science comics as tools for science education and communication: A brief, exploratory study. *Journal of Science Communication*, 8(4), 1–17.
- Wammes, J. D., Meade, M. E., & Fernandes, M. A. (2016). The drawing effect: Evidence for reliable and robust memory benefits in free recall. *Quarterly Journal of Experimental Psychology*, 69(9), 1752–1776.
- Weinstein, Y., Madan, C. R., & Sumeracki, M. A. (2018). Teaching the science of learning. *Cognitive Research: Principles and Implications*, 3(1), 2.
- Winstone, N. E., Nash, R. A., Rowntree, J., & Menezes, R. (2016). What do students want most from written feedback information? Distinguishing necessities from luxuries using a budgeting methodology. *Assessment & Evaluation in Higher Education*, 41(8), 1237–1253.
- Winstone, N. E., Nash, R. A., Rowntree, J., & Parker, M. (2017). "It'd be useful, but I wouldn't use it": Barriers to university students' feedback seeking and recipience. *Studies in Higher Education*, 42(11), 2026–2041.
- Wissman, K. T., & Rawson, K. A. (2016). How do students implement collaborative testing in real-world contexts? *Memory*, 24(2), 223–239.
- Wissman, K. T., Rawson, K. A., & Pyc, M. A. (2012). How and when do students use flashcards? *Memory*, 20(6), 568–579.
- Withey, C. (2013). Feedback engagement: Forcing feed-forward amongst law students. *Law Teacher*, 47(3), 319–344.
- Zamary, A., & Rawson, K. A. (2018). Which technique is most effective for learning declarative concepts—provided examples, generated examples, or both? *Educational Psychology Review*, 30(1), 275–301.