



Insights From the Science of Learning Can Inform Evidence-Based Implementation of Peer Instruction

Julie A. Schell^{1,2*} and Andrew C. Butler^{3,4}

¹ Department of Educational Leadership and Policy, The University of Texas at Austin, Austin, TX, United States, ² School of Design and Creative Technologies, The University of Texas at Austin, Austin, TX, United States, ³ Department of Education, Washington University in St. Louis, St. Louis, MO, United States, ⁴ Department of Psychological and Brain Sciences, Washington University in St. Louis, St. Louis, MO, United States

OPEN ACCESS

Edited by:

Rob Cassidy,
Concordia University, Canada

Reviewed by:

Tammy M. Long,
Michigan State University,
United States
Lisa McDonnell,
University of California, San Diego,
United States

*Correspondence:

Julie A. Schell
Julie.schell@austin.utexas.edu

Specialty section:

This article was submitted to
Digital Education,
a section of the journal
Frontiers in Education

Received: 06 December 2017

Accepted: 27 April 2018

Published: 28 May 2018

Citation:

Schell JA and Butler AC (2018)
Insights From the Science of Learning
Can Inform Evidence-Based
Implementation of Peer Instruction.
Front. Educ. 3:33.
doi: 10.3389/feduc.2018.00033

Peer Instruction is a popular pedagogical method developed by Eric Mazur in the 1990s. Educational researchers, administrators, and teachers laud Peer Instruction as an easy-to-use method that fosters active learning in K-12, undergraduate, and graduate classrooms across the globe. Research over the past 25 years has demonstrated that courses that incorporate Peer Instruction produce greater student achievement compared to traditional lecture-based courses. These empirical studies show that Peer Instruction produces a host of valuable learning outcomes, such as better conceptual understanding, more effective problem-solving skills, increased student engagement, and greater retention of students in science majors. The diffusion of Peer Instruction has been widespread among educators because of its effectiveness, simplicity, and flexibility. However, a consequence of its flexibility is wide variability in implementation. Teachers frequently innovate or personalize the method by making modifications, and often such changes are made without research-supported guidelines or awareness of the potential impact on student learning. This article presents a framework for guiding modifications to Peer Instruction based on theory and findings from the science of learning. We analyze the Peer Instruction method with the goal of helping teachers understand why it is effective. We also consider six common modifications made by educators through the lens of retrieval-based learning and offer specific guidelines to aid in evidence-based implementation. Educators must be free to innovate and adapt teaching methods to their classroom and Peer Instruction is a powerful way for educators to encourage active learning. Effective implementation, however, requires making informed decisions about modifications.

Keywords: Peer Instruction, cognitive science, retrieval practice, instructional design, Eric Mazur, research-based instructional strategies, learning science, active learning

INTRODUCTION

In today's classrooms, there is great demand for active learning among both students and educators. Calls for active learning are not new (see Eliot, 1909), but a recent surge of interest in this concept is transforming pedagogical practices in higher education. The inspiration for this movement comes in large part from the now well-established benefits for student achievement and motivation

produced by active learning environments (Bonwell and Eison, 1991; Braxton et al., 2000; National Research Council, 2000; Ambrose et al., 2010; Freeman et al., 2014). With a growing number of educators keenly aware of the limitations of “transmissionist” teaching methods, many of them are trying out new pedagogical methods that encourage active learning (Dancy et al., 2016).

Despite its popularity and general effectiveness, active learning is a broad concept and it is often vaguely defined, which leads to a great variability in its implementation within formal and informal education environments. We define active learning as a process whereby learners deliberately take control of their own learning and construct knowledge rather than passively receiving it (National Research Council, 2000). Active learning is not necessarily synonymous with liveliness or high levels of engagement, even if classrooms that feature active learning are often dynamic; and it is qualitatively different from more passive learning processes, such as listening to a lecture or reading a text, that primarily involve the transmission of information. Active learners construct meaning by integrating new information with existing knowledge, assess the status of their understanding frequently, and take agency in directing their learning. Even though control over learning ultimately resides with students, educators play a crucial role because they create classroom environments that can either foster or hinder active learning.

In this article, we explore the challenges faced by educators who want to effectively foster active learning using established pedagogical methods while retaining the ability to innovate and adapt those methods to the unique needs of their classroom. One challenge that educators face is that they often must teach themselves to use new methods that are very different from the teaching that they experienced as students. Moreover, graduate and post-doctoral education rarely focus on teaching, so most educators do not have any formal training to draw upon when trying to implement new methods or innovate. In addition, many educators who are trying new methods must do so with little or no feedback on effective implementation from more experienced teachers. Under these conditions, pedagogical improvement is exceedingly difficult, which makes it all the more impressive that the switch to active learning generally produces good results. Nevertheless, changes to pedagogy do not always result in positive effects. Indeed, when educators make modifications to established pedagogical methods, it may have the unintended consequence of limiting, inhibiting, or even preventing active learning. Thus, it is important for educators to understand how omitting or changing aspects of a pedagogical method might affect student learning and motivation.

We chose to focus on an established and popular pedagogical method called Peer Instruction, which researchers have demonstrated encourages active learning in a wide range of classrooms, disciplines, and fields (Mazur, 1997; Crouch and Mazur, 2001; Schell and Mazur, 2015; Vickrey et al., 2015; Müller et al., 2017). Eric Mazur developed Peer Instruction in the early 1990s at Harvard University (Mazur, 1997). The method is well-regarded in the educational research community for its demonstrated ability to stimulate active learning and achieve desired learning outcomes in a variety of educational contexts

(Vickrey et al., 2015; Müller et al., 2017). One of the key features of Peer Instruction is its flexibility that enables adaptation to almost any context and instructional design (Mazur, 1997). However, this flexibility comes with a potential cost in that modifications to the method may limit its effectiveness as it relates to active learning. Indeed, when educators modify Peer Instruction, they may be unaware that these modifications can disrupt the benefits of active learning (Dancy et al., 2016).

The primary goal of this paper is to provide Peer Instruction practitioners with an understanding of why the method is effective at fostering active learning so that they can make informed choices about how to innovate and adapt the method to their classroom. A secondary goal of this article is to respond to a need for explicit collaborations between educational researchers and cognitive scientists to help guide the implementation of innovative pedagogical methods (Henderson et al., 2015). Integrating basic principles from the science of learning into the classroom has been shown to increase learning in classrooms in ways that can easily scale and generalize to a variety of subjects (e.g., Butler et al., 2014). Unfortunately, the diffusion of general principles from the science of learning into the classroom has been much slower than innovative pedagogical methods that provide “off-the-shelf” solutions, such as Peer Instruction. Accordingly, analyzing such pedagogical methods to identify the mechanisms and basic principles that make them effective may be beneficial for both implementation in educational practice and scientific research on learning.

By way of providing the reader with an outline, our article begins with an overview of the Peer Instruction method, including a brief history and a description of the advice for implementation from the manual created by the developer (Mazur, 1997). Next, we provide an in-depth analysis of the efficacy of Peer Instruction by drawing upon theory and findings from the science of learning. Finally, we conclude with a discussion about the many common modifications users make to Peer Instruction. In this concluding section, we also provide clear recommendations for modifying Peer Instruction based on findings from the science of learning with a specific focus on a driving mechanism underlying the potent achievement outcomes associated with Mazur’s method—retrieval-based learning. Taken as a whole, we believe this article represents a novel, evidence-based approach to guiding Peer Instruction innovation and personalization that is not currently available in the literature.

PEER INSTRUCTION: A POPULAR PEDAGOGICAL METHOD THAT PROMOTES ACTIVE LEARNING

Mazur developed Peer Instruction in 1991 in an attempt to improve his Harvard undergraduates’ conceptual understanding of introductory physics (Mazur, 1997). Previously, Mazur’s teaching was lecture-based and his instructional design featured passive learning before, during, and after class. The impetus for the change in his teaching method came from David Hestenes and his colleagues who published the Force Concept Inventory (FCI)—a standardized test that evaluated students’ abilities to

solve problems based on their conceptual understanding of Newton's Laws, which is a foundational topic in introductory physics (Hestenes et al., 1992). In their classroom research using the FCI, Hestenes and colleagues found that most students could state Newton's Laws verbatim, but only a small percentage could solve problems that relied on mastery of the concept. Mazur learned about the FCI and decided to deliver the test to his students. To his surprise, the results were similar to Hestenes. After a brief period of questioning the validity of the test, Mazur became convinced that there was a serious gap in students' learning of physics in introductory college classrooms. The vast majority of physics education at the time was lecture-based. Mazur developed Peer Instruction to target the gap in conceptual understanding because he was convinced that it resulted from passive learning experiences and overreliance on transmission-based models of teaching.

The Peer Instruction Method

In 1997, Mazur published *Peer Instruction: A User's Manual* in which he describes the seven steps that constitute the method (Mazur, 1997, page 10). The seven steps are the following:

1. Question posed (1 min)
2. Students given time to think (1 min)
3. Students record individual answers [*optional*]
4. Students convince their neighbors—peer instruction (1–2 min)
5. Students record revised answers [*optional*]
6. Feedback to teacher: Tally of answers
7. Explanation of correct answer (2+ min)

As can be gleaned from the list, the Peer Instruction method involves a structured series of learning activities. The overall learning objective is the improvement of conceptual understanding, or in Mazur's words: "The basic goals of Peer Instruction are to exploit student interaction during lectures and focus students' attention on underlying concepts" (Mazur, 1997, p. 10). Accordingly, the method begins with the teacher focusing students' attention by posing a conceptual question called a ConcepTest that is generally in a multiple-choice format (but increasingly short answer format is being used), and then the remaining activities build on this question. The method is designed to take between 5 and 15 min depending on the complexity of the concept and whether all of the seven steps are used.

Given the central importance of the ConcepTest to Peer Instruction, it is no surprise that the efficacy of the method depends upon the quality of the question. Although a ConcepTest is a question, not all questions are a ConcepTest—a ConcepTest has specific features that distinguish it from other types of questions. First, as an assessment item, a ConcepTest is designed to test and build students' conceptual understanding rather than factual or procedural knowledge. Another distinct feature of a ConcepTest is the list of multiple choice alternatives. A well-designed, multiple choice ConcepTest will follow published guidelines for designing effective multiple choice questions (Haladyna et al., 2002). In particular, the teacher will construct

the responses by including a correct answer and viable distractors that elicit common misconceptions about the concept.

After the teacher poses the ConcepTest (Step 1), she gives students time to think and construct an answer based on their current understanding (Step 2). The teacher then directs students to record and display their answers to the teacher using a classroom response method (Step 3). The response method can be low-tech (e.g., hand signals, flashcards, or student whiteboards) or high-tech (e.g., clickers, text messages, or cloud-based courseware). The "modality" in which students record and/or display their answer is not critical—the key is that students generate and commit to a response (Lasry, 2008). That said, the higher-tech response systems (clickers, web-based response systems) have benefits to consider. For students, the systems record answers for later review and provide greater anonymity than using hand signals or flashcards. For teachers, the higher-tech systems enable the analysis of student responses that may inform teacher behavior and future assessment planning based on the pattern of answer choices (Schell et al., 2013). For example, students may surprise the teacher if the majority chooses a distractor as the right answer, thereby prompting the teacher to modify her teaching plan.

Once the teacher collects the responses, she reviews them without disclosing, displaying, or sharing the correct answer or the frequency of choices among the students. Next, the teacher cues students to "turn to their neighbor" to use reasoning to convince their peer of their answer (Step 4). If their neighbor has the same answer, Mazur recommends cueing students to find someone with a different answer (Mazur, 2012). Students then engage in a brief discussion in pairs where they have the opportunity to recall their response as well as justify why they responded the way they did. Mazur emphasizes that during the discussion students must defend their answers with reasoning based on what they have previously heard, read, learned, or studied. After the discussion is complete, the teacher gives students time to think about their final answer—whether they want to keep the same answer or change answers. Once they have had a moment to think, students record their final responses (Step 5), which are communicated to the teacher using the same classroom response method (Step 6).

The teacher closes the series of activities by finally revealing and explaining the correct answer (Step 7). Some teachers display the pre-post response frequencies so students can see how their answers changed (often, in the direction of the correct answer) and how many others selected specific answer choices. After revealing the correct answer, some teachers ask for explanations from representatives from each answer choice to explain their reasoning. Students are often willing to explain their reasoning despite the revelation that their response was incorrect. The purpose of this additional exercise is to help students interrogate and resolve any potential misconceptions that led them to select one of the distractors. Hearing the correct answer explained by their peers can be more effective because other novices may be able to better communicate it than the teacher who is an expert (Mazur, 1997).

Finally, it is important to note some of the key features of the method that are critical to the efficacy of Peer Instruction. In a recent article, Dancy et al. (2016, p. 010110-5) analyzed the

method in consultation with Mazur and other experienced Peer Instruction practitioners. They identified nine key features of Peer Instruction based on their research:

1. Instructor adapts instruction based on student responses
2. Students are not graded on in-class Peer Instruction activities
3. Students have a dedicated time to think and commit to answers independently
4. The use of conceptual questions
5. Activities draw on student ideas or common difficulties
6. The use of multiple choice questions that have discrete answer options
7. Peer Instruction is interspersed throughout class period
8. Students discuss their ideas with their peers
9. Students commit to an answer after peer discussion

These features, which are present in the original Peer Instruction user manual (Mazur, 1997), have proven to be essential to the success of the method.

Diffusion of Peer Instruction

Over the past quarter-century, the use of Peer Instruction has expanded far beyond Ivy League undergraduate physics education. Educators from wildly diverse contexts have used the method to engage hundreds of thousands of students in active learning. For example, middle school, high school, undergraduate, and graduate students studying Biology, Chemistry, Education, Engineering, English, Geology, US History, Philosophy, Psychology, Statistics, and Computer Science, in a variety of countries in Africa, Australia, Asia, Europe, North America, and South America, have all experienced Mazur's Peer Instruction (Mazur, 1997; Schell and Mazur, 2015; Vickrey et al., 2015; Müller et al., 2017). The widespread adoption of Peer Instruction by a diverse array of educators over the past 25 years has prompted a new area of research and large body of scholarship. Studies that support the efficacy of Peer Instruction run the gamut from applied research in a single classroom (Mazur, 1997) to multi-course, large-sample investigations (Hake, 1998), comparisons across institutional types (Fagen et al., 2002; Lasry et al., 2008), and meta-analyses covering a variety of educational contexts (Vickrey et al., 2015; Müller et al., 2017).

The consensus woven through the fabric of over two and a half decades of scholarship is that when compared to traditional lecture-based pedagogy, Peer Instruction leads to positive outcomes for multiple stakeholders, including teachers, institutions, disciplines, and (most importantly) students. For example, large-sample studies of Peer Instruction report that teachers observe lower failure rates even in challenging courses (Porter et al., 2013). On a more structural level, researchers have also demonstrated that Peer Instruction may offer a high impact solution to stubborn educational problems, such as retention of STEM majors and reduction of the gender gap in academic performance in science (Lorenzo et al., 2006; Watkins and Mazur, 2013). Peer Instruction efficacy is not limited to STEM courses. For example, Draper and Brown (2004) and Stuart et al. (2004) investigated the use of Peer Instruction in the humanities. And Chew (2004, 2005) has studied Peer Instruction use in the social sciences. Both Stuart and Chew observed

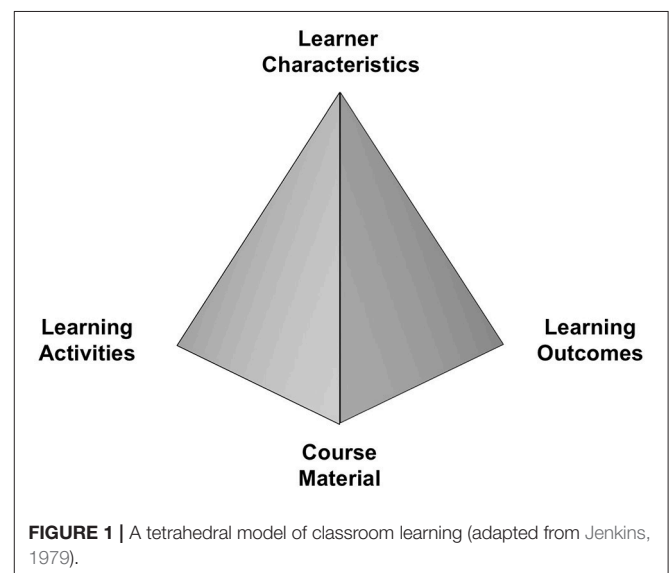
positive outcomes. Finally, the benefits of Peer Instruction are most notable for students. In particular, research has shown that learners in Peer Instruction courses develop more robust quantitative problem-solving skills, more accurate conceptual knowledge, increased academic self-efficacy, and an increased interest in and enjoyment of their subject (Hake, 1998; Nicol and Boyle, 2003; Porter et al., 2013; Watkins and Mazur, 2013; Vickrey et al., 2015; Müller et al., 2017). However, this literature is limited in the sense that it mainly focuses on educational outcomes that result from the use of Peer Instruction without considering why and how the method produces those outcomes. In the remainder of this article, we contribute such an analysis through the lens of the science of learning.

WHY IS PEER INSTRUCTION EFFECTIVE? PERSPECTIVES FROM THE SCIENCE OF LEARNING

We now turn to analyzing why Peer Instruction is an effective teaching method for fostering active learning by drawing upon theory and findings from the science of learning. As a framework for presenting our analysis, we have grouped the key aspects of Peer Instruction into four general categories of factors that form the context that an educator must consider in order to facilitate student learning in any course (see **Figure 1**). Learner objectives (Course Material and Skill), learner characteristics, learner activities, and learner outcomes.

Objectives

One of the first steps in designing any course should be the development of specific, achievable student learning objectives (Tyler, 1949; Wiggins and McTighe, 2005). Ideally, the process of developing such learning objectives grows out of a careful analysis of the goals of the course material in the context of the broader curriculum and the skills and knowledge that students



need to acquire to achieve these goals. In large part, the creation of Peer Instruction was born out of a recognition that the skills and knowledge that students acquired in introductory physics courses were qualitatively different from what is needed to progress in physics education. More specifically, the key insight was that students were acquiring procedural skills and knowledge but lacked the conceptual understanding to effectively use them, which is a common issue in many STEM disciplines (e.g., Rittle-Johnson et al., 2015). In addition to content-specific learning objectives, Mazur (1997) also emphasizes the importance of domain-general objectives that active learning can help achieve, such as critical thinking and metacognitive monitoring. Mazur states that Peer Instruction, “forces the students to think through the arguments being developed and provides them (as well as the teacher) with a way to assess their understanding of the concept” (p. 10). As a result, Peer Instruction fosters critical thinking in the domain of study and metacognition. Indeed, these cognitive skills are essential components of active learning – it is impossible to monitor and direct one’s own learning without them. When students receive feedback throughout each cycle of Peer Instruction on how well they “understand” the concepts, they can direct their efforts toward learning concepts they are struggling with. In sum, a clear sense of the skills and knowledge that students need to acquire is critical to selecting the learning activities and outcome measures that will be appropriate for any given group of students.

Activities

Educators have a multitude of instructional activities from which to choose in order to facilitate student learning and active learning more specifically (see Hattie, 2009). Importantly, there are substantial differences among this broad array activities in terms of how they affect student learning, and thus selecting an effective learning activity depends upon the learning objective (Koedinger et al., 2012). In addition, the effectiveness of a given learning activity can also differ as a function of where students are in the process of learning, so it is also imperative to consider how to structure and scaffold learning as student knowledge and skills progress. The complexity underlying how learning occurs and the need to align teaching accordingly can seem daunting to educators (Koedinger et al., 2013), which is one reason that Peer Instruction is so useful. That is, Peer Instruction provides educators with a well-structured method that includes a potent mix of effective learning activities that are designed to foster active learners.

One key to understanding the utility of any learning activity is to analyze the types of cognitive processes that are required to perform the task. Although a multitude of basic cognitive processes are engaged during learning, educators are understandably more interested in types of processing that facilitate the construction of meaning from information (Craik and Lockhart, 1972; Craik and Tulving, 1975). At this more complex level, there are many ways in which people can process information (e.g., Packman and Battig, 1978; Hunt and Einstein, 1981). One framework that can inform the analysis of the cognitive processes that are engaged by a particular learning activity is the updated Bloom’s taxonomy of educational

objectives (see too Bloom, 1956; Anderson et al., 2001). Does the activity involve application, analysis, classification, evaluation, comparison, etc.? The reason that such analysis is important is that the cognitive processes that are used during the activity will dictate what is learned and as such, how students will direct further learning. With this idea in mind, a clear advantage of Peer Instruction is that it provides educators with great flexibility in deciding how students should process the information during learning. For example, the question posed on a ConcepTest, whether in multiple choice or constructed response format, could induce students to engage any one or multiple processes described in Bloom’s taxonomy.

While on the topic of cognitive processing, one critical distinction is that learning activities differ in the extent to which they involve perceiving and encoding new information relative to retrieving and using information that has already been stored in memory. In more simplistic terms, this distinction is between how much the activity involves “putting information in” vs. “getting information out.” Many of the learning activities traditionally used in college courses predominantly involve perceiving and encoding new information—listening to a lecture, reading a textbook, watching video, etc. One of the key innovations in Peer Instruction is to introduce more activities that require students to retrieve and use information (e.g., ConcepTests), a change that is reflective of a broader movement toward active learning in STEM courses (Freeman et al., 2014). Perceiving and encoding new information is imperative during the initial stages of learning. However, after students have some knowledge to work with, it is often much more effective for them to engage in activities that require them to retrieve and use that knowledge (Roediger and Karpicke, 2006; for review see Dunlosky et al., 2013; Rowland, 2014).

Retrieval practice is a low-threshold instructional activity in that it is simple and easy to implement for teachers. Engaging in retrieval practice has both direct and indirect effects on learning. The direct effect stems from the fact that retrieving information from memory changes memory, and thus causes learning (Roediger and Butler, 2011). Retrieval practice has been shown to improve long-term retention (e.g., Larsen et al., 2013) and transfer of learning to new contexts (e.g., Butler, 2010; for review see Carpenter, 2012). In addition, the indirect effects are numerous—students are incentivized to keep up with material outside of class (Mawhinney et al., 1971) and they become less anxious about assessments (Agarwal et al., 2014), among other benefits. When educators use Peer Instruction following Mazur’s protocol, students engage in more than three distinct retrieval practice opportunities in a single cycle (see above section on Peer Instruction Method, Steps 2,4, and 5). In short, retrieval practice is a critical mechanism in the Peer Instruction method that facilitates the development of deeper understanding that enables students to transfer their knowledge to new contexts (e.g., solve problems, analyze new ideas). We discuss retrieval as a mechanism for learning in Peer Instruction in the next section on common modifications.

Having students engage in activities that require retrieving and using recently acquired knowledge also has another important indirect benefit—it provides feedback to both students and

educators (Black and Wiliam, 1998; Hattie, 2009). Feedback is one of the most powerful drivers of learning because it enables students to check their understanding and address any potential gaps (Bangert-Drowns et al., 1991; Butler and Winne, 1995; Hattie and Timperley, 2007). In particular, explanation feedback promotes the development of deeper understanding (Butler et al., 2013). Equally important is the information that is provided to educators about the current state of student understanding, which enables them to circle back and address misunderstandings. In comparison to the traditional lecture method, Peer Instruction is rich with opportunities for feedback from student-to-student, teacher-to-student, and student-to-teacher (i.e., in addition to the metacognitive benefits of feedback that results from retrieval practice). The student-to-student feedback may be particularly valuable given the benefits of collaborative learning (Nokes-Malach et al., 2015). As described in the Peer Instruction manual, students can often explain concepts better to each other than their teacher can, providing both valuable feedback and new information (Smith et al., 2009). In addition, the act of explaining to someone else is a powerful learning event as well, so both students benefit.

Finally, it is important to consider how activities are structured in order to continuously facilitate learning as the acquisition of knowledge and skills progresses. Peer Instruction does a good job of scaffolding student learning—pre-class readings and reading quizzes prepare students to learn in class, lectures present new information that extends from the readings, ConcepTests provide further practice and an opportunity for feedback. All of these activities are aligned and build upon each other. This structure also incorporates many basic principles from the science of learning that are known to promote long-term retention and the development of understanding. For example, learning is spaced or distributed over time rather than massed (Dempster, 1989; Cepeda et al., 2006) and variability is introduced during the learning of a particular piece of concept or skill by using different examples, contexts, or activities (Glass, 2009; e.g., Butler et al., 2017). Variation of this sort is particularly useful for honing students' abilities to be active learners who transfer their knowledge across contexts (Butler et al., 2017). A single cycle of Peer Instruction, which could be as short as 2–3 min, is packed with variation in learning activities. For example, students think on their own, retrieve, discuss, retrieve again, and then receive feedback on their responses.

Learner Characteristics

Perhaps the most important set of factors that influence learning in any course are the characteristics of the learners—their individual knowledge and experiences, expectations, interests, goals, etc. Individual differences play a major role in determining student success in STEM disciplines (Gonzalez and Kuenzi, 2012), and yet it is this aspect of the learning context that is so often ignored in large introductory STEM courses. One of the reasons that Peer Instruction is so effective is that it directly addresses this issue in that it is “student-centered.” Throughout the Peer Instruction manual there is a consistent focus on the student experience when explaining the methodology. In addition, the rationale for focusing on students is bolstered by

insightful anecdotes and observations: “Students' frustration with physics—how boring physics must be when it is reduced to a set of mechanical recipes that do not even work all the time!” (Mazur, 1997, p. 7). Taken as a whole, the manual makes clear that student engagement is essential to the successful implementation of Peer Instruction.

Nevertheless, it is possible for a pedagogy to be “student-centered” and yet ineffective on this front; what sets Peer Instruction apart is that it is consistent with many principles and best practices from research on student motivation. Chapter 3 of the manual, which focuses on student motivation, begins with advice about “setting the tone” that addresses student expectations and beliefs about learning. One theme that emerges is that students should embrace the idea that learning is challenging and requires effort and strategic practice (i.e., a growth mindset; Yeager and Dweck, 2012). Students who adopt such a mindset often show greater resilience and higher achievement (e.g., Blackwell et al., 2007). Another theme that emerges is about the importance of students coming to value what they are learning in the course and the methodology used for learning. People's perceptions about the value of an activity (e.g., self-relevance, interest, importance, etc.) can have a strong effect on their motivation to engage in that activity (Harackiewicz and Hulleman, 2010; Cohen and Sherman, 2014). Examples from the manual include Mazur's introductory questionnaire that probes student goals and interests and the explanation provided on the first day of class about why the course is being taught in this manner. A third theme that emerges is the benefit of creating a cooperative learning environment rather than a competitive one. Classrooms that foster cooperation lead students to adopt mastery learning goals (i.e., rather than performance goals) and produce greater achievement relative to classrooms that foster competition (Johnson et al., 1981; Ames, 1992). Numerous aspects of Peer Instruction help produce a cooperative environment, from the student-to-student peer instruction at the core of the pedagogy to the use of an absolute grading scale that enables everyone to succeed.

Outcomes

The purpose of any course is to facilitate learning that will endure and transfer to new situations. In education, summative assessment provides a proximal measure of learning that is assumed to predict future performance (Black, 2013). As such, it is imperative that the nature of the assessment used reflect such future performance to the extent that it is possible. The assessment tools used within Peer Instruction and afterwards to evaluate its effectiveness are derived from a careful analysis of what students must know and do in future courses. The result of this analysis is mix of different types of assessments each designed to measure a different aspect of the knowledge and skills that students need to acquire. The use of one or more diagnostic tests that tap fundamental concepts in the discipline are recommended (e.g., the FCI and the Mechanics Baseline Test in physics). Course exams are meant to feature different types of questions, such as conceptual essays and conventional problems, that engage students in types of cognitive processing (see discussion of learning activities above; Anderson et al., 2001).

Importantly, the assessment tools used in Peer Instruction are not only aligned with the future, but also with the activities that are used to facilitate student learning. As discussed above, the cognitive processes that students engage during activity determines what is learned; however, a student's ability to demonstrate that learning depends upon the nature of the assessment task. Performance tends to be optimized when the processes engaged during learning match the processes required for the assessment, a concept known as *transfer-appropriate processing* (Morris et al., 1977; for a review see Roediger and Challis, 1989). When there is a mismatch in cognitive processing (e.g., learning involved application but the test requires evaluation), then assessment can fail to accurately measure student learning.

Finally, it is critical to remember that every assessment provides the opportunity to both measure learning and facilitate learning. Every question that a student answers, regardless of whether it is in the context of a low-stakes ConcepTest or a high-stakes exam, provides summative information (i.e., measuring learning up until that point), formative information (feedback for the student and teacher), and an opportunity to retrieve and use knowledge that directly causes learning. Thus, assessment is learning and learning is assessment, and this inherent relationship makes it even more imperative that assessment reflect what students must be able to know and do in future.

In summary, Peer Instruction is an effective pedagogy because it utilizes many principles and best practices from the science of learning, while also allowing flexibility with respect to implementation. No laws of learning exist (McKeachie, 1974; Roediger, 2008), and thus facilitating student learning involves considering each category of factors shown in **Figure 1** in the context of the other three categories to optimize learning (see McDaniel and Butler, 2011). By allowing flexibility, Peer Instruction enables educators to foster active learning in ways that are optimal for their particular context. In the next section, we use the insights about Peer Instruction gleaned from the science of learning to evaluate the potential impact of common modifications to the method made by teachers. Our goal is to provide evidence-based guidance for how to make decisions about modifying Peer Instruction in ways that will not undermine student learning and motivation.

IMPLEMENTING PEER INSTRUCTION: RECOMMENDED GUIDELINES ON COMMON MODIFICATIONS BASED ON RETRIEVAL-ENHANCED LEARNING

Teachers commonly modify their use of Peer Instruction (Turpen and Finkelstein, 2007; Dancy et al., 2016; Turpen et al., 2016). In physics education, where Peer Instruction has been most widely practiced, Dancy et al. (2016) found that teachers often make modifications to Mazur's method by omitting one or more of the seven steps outlined in the original user manual. In addition, Dancy et al. found that teachers also modify the nine key features identified through their analysis (see above section on The Peer

Instruction Method). Teachers gave variety of reasons, both personal and structural, for their modifications. Some teachers revealed that they modified the method because they did not have a clear understanding of it (e.g., they often confused Peer Instruction with general use of peer-to-peer engagement). Other teachers reported making modifications due to concerns about the limited time to cover content during class time or a perceived difficulty with motivating students to engage in the method. Finally, many teachers stated they modified the Peer Instruction method by omitting key steps and features because they were unaware that eliminating them might negatively affect learning, motivation, or other desired outcomes (Dancy et al., 2016; Turpen et al., 2016). Taken as a whole, studies on teacher implementation of Peer Instruction indicate that the common changes made to the method are not informed by the science of learning, educational research on active learning in the classroom, or even the literature on Peer Instruction itself.

The overwhelmingly positive results produced by Peer Instruction despite the prevalence of relatively uninformed modifications to the method is intriguing. This finding speaks to the robust effectiveness of Peer Instruction because a potent cocktail of mechanisms for learning remain even if one aspect of the method is removed. For example, eliminating one of the many retrieval attempts in the 7-step cycle still leaves many opportunities for retrieval practice. However, it also obscures the possible reductions in effectiveness of the method that such changes might cause. Much of the literature on Peer Instruction is built on studies in which the method is implemented in full fidelity or modified by researchers who have carefully designed the modification. The subset of studies in which modifications have been made to the method usually find positive results, but the magnitude of the observed effects may be lower, indicating an overall reduction in effectiveness. Of course, modifications could also maintain or even improve the effectiveness of the method. However, we argue that the changes to the Peer Instruction most likely to improve the effectiveness of the method are ones that are supported by theory, findings, and evidence from the science of learning and classroom research.

In this final section, we aim to help Peer Instruction practitioners understand how their choices with respect to common modifications could affect active learning in their classroom. More specifically, we provide answers to the following two questions: If a Peer Instruction user wishes to promote active learning in their classroom, what should they understand about common modifications to the method? What are some other modifications teachers can make that would be aligned with the science of learning? We focus on the concept of retrieval-based learning in order to further explicate one of the key mechanisms that drives learning in Peer Instruction. We hone in on retrieval to explain Peer Instruction effectiveness and to guide implementation for two reasons. First, as aforementioned, Peer Instruction is packed with retrieval events. Second, retrieval is one of the most firmly established mechanisms for causing student learning, retention of learning regardless of complexity of the material, and the ability to transfer learning to new contexts (Roediger and Butler, 2011). Many of the modifications made to the method reduce the number of opportunities for students

to retrieve and use their knowledge. The remainder of the paper is dedicated to describing six common modifications made to Peer Instruction and discussing the potential effects of these changes. The result is a set of detailed decision-making guidelines supported by the science of learning with clear recommendations for modifying Peer Instruction.

Retrieval-Based Learning: A Key Mechanism in Peer Instruction

As explained above, retrieval practice is one of the most robust and well-established active-learning strategies in the science of learning (for review see Roediger and Karpicke, 2006; Roediger and Butler, 2011; Carpenter, 2012; Dunlosky et al., 2013; Rowland, 2014), and it pervades the Peer Instruction method. Retrieval involves pulling information from long-term memory into working memory so that it can be re-processed along with new information for a variety of purposes. The cue used to prompt a retrieval attempt (e.g., the question, problem, or task) determines in large part what knowledge is retrieved and how it is re-processed. The information can be factual, conceptual, or procedural in nature, among other types and aspects of memory. Thus, depending on the cue, retrieval can be used for anything from rote learning (e.g., the recall of a simple fact) to higher-order learning (e.g., re-construction of a complex set of knowledge in order to analyze a new idea). As people attempt to retrieve a specific piece of information from memory, they also activate related knowledge, making it easier to access this other knowledge if needed and integrate new information into existing knowledge structures. In the foregoing discussion of common modifications, we refer to the act of *attempting* to pull knowledge from memory as a retrieval opportunity. It is important to note that such an attempt to retrieve can be a potent learning event even if retrieval is unsuccessful. Science of learning researchers have demonstrated that even when students fail to generate the correct knowledge or make an error, the mere act of trying to retrieve potentiates (or facilitates) subsequent learning, especially when feedback is provided after the attempt (Metcalfe and Kornell, 2007; Arnold and McDermott, 2013; Hays et al., 2013).

The effectiveness of retrieval-based learning can be enhanced in several ways depending on how retrieval practice is implemented and structured. In our subsequent analysis of modifications to Peer Instruction, we will focus on four specific ways to make instruction that employs retrieval practice more effective:

- 1) **Feedback**—Retrieval practice is beneficial to learning even without feedback (e.g., Karpicke and Roediger, 2008), but it becomes even more effective when feedback is provided (Kang et al., 2007; Butler and Roediger, 2008)
- 2) **Repetition**—A single retrieval opportunity can be effective, but retrieval practice becomes even more effective when students receive multiple opportunities to pull information from memory and use it (Wheeler and Roediger, 1992; Pyc and Rawson, 2007).
- 3) **Variation**—Verbatim repetition of retrieval practice can be useful and effective for memorizing simple pieces of

information (e.g., facts, vocabulary, etc.), but introducing variation in how information is retrieved and used can facilitate the development of deeper understanding (Butler et al., 2017).

- 4) **Spacing**—When repeated, retrieval practice is more effective when it is spread out or distributed over time, even if the interval between attempts is just a few minutes (Kang et al., 2014).

Of course, these four ways can also be used in various combinations, which creates the potential for even greater effectiveness.

Peer Instruction involves numerous retrieval opportunities that are implemented and structured in a way that would enhance the benefits of such retrieval practice. Many of the common modifications to Peer Instruction involve eliminating opportunities for retrieval practice in ways that might reduce active learning. The simplest recommendations for guiding Peer Instruction modification through the lens of retrieval-based learning are to consider increasing the number of opportunities to engage in retrieval practice, implement and structure retrieval practice in effective ways (e.g., provide feedback), and avoid omitting the retrieval opportunities present in the original method (Mazur, 1997). With that advice in mind, we now turn to analyzing some of the common modifications to Peer Instruction.

Common Modification #1: Skipping Initial Individual Thought and Response

One of the most common modifications to Peer Instruction is skipping the first retrieval event (Steps 2 and 3) and moving right into the peer discussion (Step 4) (Turpen and Finkelstein, 2009; Vickrey et al., 2015). In this modification scenario, teachers pose the question or ConcepTest, but they immediately direct students to turn to their neighbor to discuss instead of giving students time to think and respond on their own. Nicol and Boyle (2003) report that students prefer Peer Instruction when the initial individual think and response steps are included, but there are additional, more important reasons to keep the first response in the Peer Instruction cycle.

Through the lens of retrieval-based learning, skipping the initial opportunity for students to generate a response is problematic for several reasons. First, there is a learning benefit to students from attempting to retrieve information without immediate feedback, even if they are not able to generate the correct response. Second, a prominent finding from the science of learning literature is that repeated retrieval of the same question enhances learning (see Roediger and Butler, 2011). Removing the first response reduces the benefits of engaging in multiple rounds of retrieval practice on the same question throughout the Peer Instruction cycle. Finally, removing the first retrieval attempt eliminates a powerful opportunity for students to engage in metacognitive monitoring about their current understanding of the content being tested. Fostering student metacognition is critical to helping students direct their subsequent learning behavior. In summary, we offer the following guideline for Common Modification #1: Removing the first “think and response” steps eliminates a key retrieval practice

opportunity and thereby reduces a key opportunity for active learning produced by Peer Instruction. Avoid this modification unless it is absolutely necessary or if you plan to replace the omitted retrieval with another equally powerful learning activity.

Common Modification #2: Revealing the Frequency of Responses Before Peer Discussion

Another common modification to Peer Instruction is revealing the results of the initial thought and response (Steps 2 and 3) before the peer discussion begins (Step 4) without revealing the correct answer (Vickrey et al., 2015). For example, some educators using clickers or other voting devices will show the results on the screen via a projector; or if using flashcards, they will reveal by verbal description the frequency of student responses after the first round (e.g., 70% of students voted A, 20% voted for B, 5% for C, and 5% for D). Some researchers have found that revealing the results of the vote (but not the correct answer) before peer discussion biases student responses to the most commonly chosen answer even if that answer is incorrect (Perez et al., 2010; Vickrey et al., 2015). However, a smaller study in chemistry education did not find a student bias when the responses were revealed before (see Vickrey et al., 2015). Although the effects of this modification deserve further investigation, we think that it is helpful to consider how it might influence retrieval-based learning in Peer Instruction. By showing the distribution of responses in the class, students may misinterpret this information as feedback and think that the most popular answer choice is correct. Such a misinterpretation could potentially confuse students or even lead them to acquire a misconception. In addition, the benefits of retrieval practice are enhanced when there is a delay between the retrieval attempt and corrective feedback (e.g., Butler and Roediger, 2008). By contrast, providing students with the class response frequencies right after the initial individual thought and response essentially constitutes immediate feedback. In summary, we offer the following guideline for Common Modification #2: Educators who elect to reveal the response frequencies before peer discussion may confuse students and negate the benefits of delaying feedback (e.g., time for students to reflect on their understanding), so we recommend not revealing students' answers after the first response round.

Common Modification #3: Refashioning Question Design

Educators use many different types and formats of questions during Peer Instruction cycles that do not always align with the original conceptualization of a ConcepTest, which is a multiple-choice test designed to build conceptual understanding (Mazur, 1997). Popular modifications include switching from multiple choice to constructed response format and using types of questions that are not necessarily aimed at conceptual understanding (Smith et al., 2009; Vickrey et al., 2015). Routinely, Peer Instruction practitioners also fill class time with administrative questions, such as polling to record attendance, using questions that require recall of basic facts to determine

if students completed pre-assigned homework, or to check if students are listening during a lecture. The consensus from reviews of Peer Instruction efficacy is that questions that are challenging and involve higher-order cognition (e.g., application, analysis; see Anderson et al., 2001) are correlated with larger gains in learning than questions that require the recall of basic facts (Vickrey et al., 2015). As such, modification recommendations for Peer Instruction tend to emphasize that ConcepTest questions should tap higher-order cognition and not recall of basic facts. For the most part, theory and findings from the science of learning would agree with these recommendations. However, it is important for educators to consider the learning objectives of the course and each particular class when creating or selecting questions. If mastery of basic knowledge (e.g., vocabulary, facts) is important then giving students retrieval practice through Peer Instruction on such information is useful. Indeed, improving students' basic knowledge can form a strong foundation that enables them to effectively engage in higher-order cognition. Nevertheless, it is probably best that retrieval practice of such basic knowledge be given outside of class time and the use of ConcepTests focused on engaging students in higher-order cognition during class when the teacher and peers are available to aide in understanding.

With respect to format, Peer Instruction researchers emphasize that writing multiple-choice questions with viable distractors is one of the key elements that represent fidelity of implementation, but practitioners often lament that multiple-choice questions are difficult to construct. Although there are clear benefits to the use of multiple-choice format (e.g., ease of grading responses), the type of question being asked is much more important for learning than the format of the question (McDermott et al., 2014; Smith and Karpicke, 2014). In summary, we offer the following guideline for Common Modification #3: Feel free to be creative with the ConcepTest using different formats and types of questions, but it is probably best if ConcepTest questions posed during class time engage students in higher-order cognition. And, because even one act of retrieval can significantly enhance students' knowledge retention, ConcepTests or other Peer Instruction questions should always be aligned with specific learning objectives and not content teachers do not really want students to remember or use in the future.

Common Modification #4: Skipping Peer Discussion

Some Peer Instruction practitioners elect to skip peer discussion and only require a single round for individual thought and response. However, peer discussion represents an important learning opportunity for students because it requires them to engage in many different higher-order cognitive processes. When following Mazur's protocol, students must first retrieve their response to the ConcepTest, which provides another opportunity for retrieval practice. Next, they must discuss it with their partner, a complex interaction which involves explaining the rationale for why their answer is the correct answer, considering another point of view and (potentially) new information,

thinking critically about competing explanations, and updating knowledge (if the response was incorrect). Although he does not detail it in the Peer Instruction manual (Mazur, 1997), Mazur now recommends educators to instruct students not to just “turn to your neighbor and convince them you are right” but to “find someone with a different answer and convince them you are right.” Note that Smith et al. (2009) found that “peer discussion enhances understanding, even when none of the students in a discussion group originally knows the correct answer” (p. 010104-1). The task of convincing someone else about the correctness of a response may require retrieving other relevant knowledge (e.g., course content, source information about where they learned it), and thus it might be considered additional retrieval practice that is distinct but related to the ConcepTest question itself. Peer discussion also allows students to practice a host of domain-general skills, such as logical reasoning, debating, listening, perspective-taking, metacognitive monitoring, and critical thinking. Removing such a rich opportunity for active learning seems like it would have negative consequences, and indeed it does: Smith et al. (2009) found that the inclusion of peer discussion was related to larger gains in learning relative to its omission. That said, Mazur does endorse skipping peer discussion if during the first “think and respond” rounds, more than 70% of students respond correctly OR less than 30% of students correctly (Mazur, 2012). In summary, we offer the following guideline for Common Modification #4: Eliminating peer discussion removes the central feature of Peer Instruction, one that contains a cocktail of potent mechanisms for learning, especially variation in retrieval practice. Because there are benefits to peer discussion even when students have the wrong answer, we recommend always including peer discussion. In cases where the majority of the students have responded correctly, consider shortening the discussion period.

Common Modification #5: Skipping Final Individual Thought and Response (Step 5)

In Peer Instruction, some teachers may skip the final individual response round (Step 5). In this scenario, teachers deliver the ConcepTest question, solicit individual thinking and responses, engage students in peer discussion, but then move directly to an explanation of the correct answer. Although it is less common than skipping the initial individual “think and response” rounds, some teachers eliminate this step if they need to save time or a large percentage of students are correct on their initial response. Like skipping the peer discussion round when a large percentage (over 70%) of students’ initial responses are correct, skipping the final response round in the same situation is endorsed by Mazur (2012). Skipping the final “think and respond” rounds eliminates an opportunity for repeated, spaced retrieval practice. Importantly, retrieval practice is substantively distinct from rote repetition—students have been exposed to new information in the interim between retrieval attempts and thus the second retrieval attempt represents a learning event that can facilitate the updating of knowledge. Such knowledge updating is likely to occur regardless of whether students’ responses were correct

or incorrect initially because either way they are being exposed to new information during peer discussion. In summary, we offer the following guideline for Common Modification #5: The time saved by skipping the final individual thought and response probably does not outweigh the benefits of repeated spaced retrieval practice, but a potential alternative would be shift its timing by asking students to provide their final answer and an explanation for it after class as homework (i.e., further increasing the spacing between retrieval attempts, which would be beneficial).

Common Modification #6: Skipping the Explanation of the Correct Answer

Occasionally, educators choose to eliminate the final step of the Peer Instruction method—the explanation of the correct answer (Step 7). However, this step is critically important, especially if steps 1–6 reveal that student understanding is poor, because of the powerful effects of explanation feedback on student understanding (for review see Hattie and Timperley, 2007; e.g., Butler et al., 2013). In separate studies on Peer Instruction each in a different discipline, Smith et al. (2011) and Zingaro and Porter (2014) observed larger gains in learning when an explanation was provided relative to when it was not. An ideal implementation of this final step might proceed as follows: Once students have recorded their final response, the teacher reveals the correct answer, provides explanatory feedback, and then potentially engages students in additional learning activities if the desired level of mastery has not been achieved. However, there is ample room for flexibility and customization in how explanatory feedback is provided. When using Peer Instruction, the first author often implements the final step by asking student representatives from each answer choice to again retrieve their answers and explain the rationale for supporting their response. The following script illustrates this version of Step 7:

Teacher: “The correct answer was C; can I get a volunteer who answered differently to explain their thinking?”

Student: “[Provides one or two explanations for answer choice A]”

Teacher: “[Takes the opportunity to address misconceptions underlying answer choice A]. How about a volunteer who chose B or who can understand why someone else might do so?”

Student: “[Provides one or two explanations for answer choice B]”

Teacher: “[Takes the opportunity to address misconceptions underlying answer choice A]. Thank you, how about answer C? Why did you select C?”

Student: “[Provides one or two explanations for answer choice C]”

Teacher: “[Takes the opportunity to address misconceptions underlying answer choice C and provides the final explanation]”

A script for constructed responses rather than multiple choice questions would be analogous, but the teacher might specify several possible answers generated by students instead of the multiple-choice alternatives (A, B, C, etc.). It is also worth noting that this particular implementation of the final explanatory feedback step adds yet another repeated, spaced retrieval

opportunity to the original method. However, students who volunteer to explain their response in front of a large group are engaging in learning event that is somewhat different from the other retrieval attempts that occurred earlier and thus it incorporates valuable variation in retrieval practice as well. In summary, we offer the following guideline for Common Modification #6: The final step of Peer Instruction invites opportunities for innovation and customization, but the one modification that we discourage is the elimination of explanatory feedback. That said, teachers should feel free to customize their approach to this explanation, such as through the above script, demonstrations, discussion, simulations, and more.

CONCLUSION

Teaching is an incredibly personal endeavor. Part of the beauty of teaching is the opportunity it provides an educator to breathe unique life into a subject to which they have dedicated their careers. Thus, it seems both natural and important for teachers to be able to personalize the way they teach so that it fits within their teaching context. Given the desire for personalization in teaching, it is imperative to allow flexibility in the use of instructional methods developed by others. A key characteristic of innovations that scale, pedagogical or otherwise, is the innovation's capacity for reinvention or customization in ways the developer did not anticipate (Rogers, 2003). Indeed, experts who study the uptake of pedagogical innovation report that teachers "rarely use a research-based instructional strategy 'as is.' They almost always use it in ways different from the recommendations of the developer" (Dancy et al., 2016, p. 12; see too Vickrey et al., 2015).

Yet, allowing the flexibility for teachers to modify instructional methods also comes with a potential cost because modifications can reduce the efficacy of the method. If a teacher using a modified version of a method observes limited or no improvement in learning outcomes, their tweaked version may lead to the erroneous conclusion that the method itself does not work; and if teachers sense the new method they have adopted does not work, they may choose to return to more familiar pedagogical habits that encourage passivity in students and yield middling results for learning (Vickrey et al., 2015; Dancy et al., 2016).

The potential for evidence-based pedagogical methods to produce poor results due to modifications creates a tension

between the need to personalize teaching and the need to follow protocols that are designed to produce specific learning outcomes. We believe this tension can be resolved if teachers understand why a method is effective at facilitating learning so that they can make informed decisions about potential modifications. To this end, we have provided an analysis of why Peer Instruction is effective through the lens of the science of learning and clear guidelines regarding common modifications of the method. Peer Instruction is a remarkably flexible, easy-to-use, high-impact pedagogy that has been shown to foster active learning in a variety of contexts. By simply following the original method described by Mazur (1997), educators can infuse the state-of-the-art learning science in their classrooms and be assured they are using practices demonstrated to foster active learning. Nevertheless, the personal nature of teaching guarantees that teachers will modify Peer Instruction. We love the spirit of teaching improvement and innovation that educators are embracing, and we encourage them to make their choices by evaluating evidence from the science of learning while also considering their own unique classroom context.

AUTHOR CONTRIBUTIONS

JS and AB co-developed the concept of the paper. JS lead contributions to the Introduction, the first two sections of the paper, the section on common modification, and the conclusion and provided feedback on the remainder of the paper. AB led the section on perspectives from the science of learning and made substantive conceptual and written contributions to all sections of the paper.

ACKNOWLEDGMENTS

JS would like to acknowledge Eric Mazur for consultation on the ConcepTest portion of this paper and for his development of Peer Instruction and Charles Henderson, Melissa Dancy, Chandra Turpen, and Michelle Smith for their scholarship on educator modifications of Peer Instruction. While co-writing this article, the second author (AB) was supported by the James S. McDonnell Foundation Twenty first Century Science Initiative in Understanding Human Cognition—Collaborative Grant No. 220020483.

REFERENCES

- Agarwal, P. K., D'Antonio, L., Roediger, H. L., McDermott, K. B., and McDaniel, M. A. (2014). Classroom-based programs of retrieval practice reduce middle school and high school students' test anxiety. *J. Appl. Res. Mem. Cogn.* 3, 131–139. doi: 10.1016/j.jarmac.2014.07.002
- Ambrose, S. A., Bridges, M. W., DiPietro, M., Lovett, M. C., and Norman, M. K. (2010). *How Learning Works: Seven Research-Based Principles for Smart Teaching*. San Francisco, CA: John Wiley & Sons.
- Ames, C. (1992). Classrooms: goals, structures, and student motivation. *J. Educ. Psychol.* 84, 261–271. doi: 10.1037/0022-0663.84.3.261
- Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., et al. (Eds.) (2001). *A Taxonomy For Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York, NY: Longman, Inc.
- Arnold, K. M., and McDermott, K. B. (2013). Test-potentiated learning: distinguishing between direct and indirect effects of tests. *J. Exp. Psychol. Learn. Mem. Cogn.* 39, 940–945. doi: 10.1037/a0029199
- Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. A., and Morgan, M. (1991). The instructional effect of feedback in test-like events. *Rev. Educ. Res.* 61, 213–238. doi: 10.3102/00346543061002213
- Black, P. (2013). "Formative and summative aspects of assessment: theoretical and research foundations in the context of pedagogy." in *Handbook of Research on Classroom Assessment*, eds James H. McMillan Sage. (Thousand Oakes, CA: SAGE Publications, Inc), 167–178.

- Black, P., and Wiliam, D. (1998). Assessment and classroom learning. *Assess. Educ. Princ. Pol. Pract.* 5, 7–74. doi: 10.1080/0969595980050102
- Blackwell, L. S., Trzesniewski, K. H., and Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: a longitudinal study and an intervention. *Child Dev.* 78, 246–263. doi: 10.1111/j.1467-8624.2007.00995.x
- Bloom, B. S. (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals*. Essex, England: Harlow.
- Bonwell, C. C., and Eison, J. A. (1991). *Active Learning: Creating Excitement in the Classroom*. 1991 ASHE-ERIC Higher Education Reports. ERIC Clearinghouse on Higher Education, The George Washington University, One Dupont Circle, Suite 630, Washington, DC 20036-1183.
- Braxton, J. M., Milem, J. F., and Sullivan, A. S. (2000). The influence of active learning on the college student departure process: toward a revision of Tinto's theory. *J. High. Educ.* 71, 569–590. doi: 10.2307/2649260
- Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *J. Exp. Psychol. Learn. Mem. Cogn.* 36, 1118–1133. doi: 10.1037/a0019902
- Butler, A. C., Black-Maier, A. C., Raley, N. D., and Marsh, E. J. (2017). Retrieving and applying knowledge to different examples promotes transfer of learning. *J. Exp. Psychol. Appl.* 23, 433–446. doi: 10.1037/xap0000142
- Butler, A. C., Godbole, N., and Marsh, E. J. (2013). Explanation feedback is better than correct answer feedback for promoting transfer of learning. *J. Educ. Psychol.* 105, 290–298. doi: 10.1037/a0031026
- Butler, A. C., Marsh, E. J., Slavinsky, J. P., and Baraniuk, R. G. (2014). Integrating cognitive science and technology improves learning in a STEM classroom. *Educ. Psychol. Rev.* 26, 331–340. doi: 10.1007/s10648-014-9256-4
- Butler, A. C., and Roediger, H. L. III. (2008). Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Mem. Cogn.* 36, 604–616. doi: 10.3758/MC.36.3.604
- Butler, D. L., and Winne, P. H. (1995). Feedback and self-regulated learning: a theoretical synthesis. *Rev. Educ. Res.* 65, 245–281. doi: 10.3102/00346543065003245
- Carpenter, S. K. (2012). Testing enhances the transfer of learning. *Curr. Dir. Psychol. Sci.* 21, 279–283. doi: 10.1177/0963721412452728
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., and Rohrer, D. (2006). Distributed practice in verbal recall tasks: a review and quantitative synthesis. *Psychol. Bull.* 132, 354–380. doi: 10.1037/0033-2909.132.3.354
- Chew, S. L. (2004). Using ConcepTests for formative assessment. *Psychol. Teach. Netw.* 14, 10–12. Available online at: <http://www.apa.org/ed/precollege/ptn/2004/01/issue.pdf>
- Chew, S. L. (2005). “Student misperceptions in the psychology classroom,” in *Essays from excellence in teaching*, eds B. K. Saville, T. E. Zinn and V. W. Hevern (Available online at: Society for the Teaching of Psychology Web site: <https://teachpsych.org/ebooks/tia2006/index.php>)
- Cohen, G. L., and Sherman, D. K. (2014). The psychology of change: self-affirmation and social psychological intervention. *Annu. Rev. Psychol.* 65, 333–371. doi: 10.1146/annurev-psych-010213-115137
- Craik, F. I., and Lockhart, R. S. (1972). Levels of processing: a framework for memory research. *J. Verbal Learn. Verbal Behav.* 11, 671–684. doi: 10.1016/S0022-5371(72)80001-X
- Craik, F. I. M., and Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *J. Exp. Psychol. Gen.* 104, 268–294. doi: 10.1037/0096-3445.104.3.268
- Crouch, C. H., and Mazur, E. (2001). Peer Instruction: ten years of experience and results. *Am. J. Phys.* 69, 970–977. doi: 10.1119/1.1374249
- Dancy, M., Henderson, C., and Turpen, C. (2016). How faculty learn about and implement research-based instructional strategies: the case of Peer Instruction. *Phys. Rev. Phys. Educ. Res.* 12, 010110. doi: 10.1103/PhysRevPhysEducRes.12.010110
- Dempster, F. N. (1989). Spacing effects and their implications for theory and practice. *Educ. Psychol. Rev.* 1, 309–330. doi: 10.1007/BF01320097
- Draper, S. W., and Brown, M. I. (2004). Increasing interactivity in lectures using an electronic voting system. *J. Comput. Assist. Learn.* 20, 81–94. doi: 10.1111/j.1365-2729.2004.00074.x
- Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., and Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychol. Sci. Public Interest* 14, 4–58. doi: 10.1177/1529100612453266
- Eliot, C. W. (1909). *Educational Reform: Essays and Addresses*. New York, NY: Century.
- Fagen, A. P., Crouch, C. H., and Mazur, E. (2002). Peer Instruction: results from a range of classrooms. *Phys. Teacher* 40, 206–209. doi: 10.1119/1.1474140
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., et al. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceed. Natl. Acad. Sci.* 111, 8410–8415. doi: 10.1073/pnas.1319030111
- Glass, A. L. (2009). The effect of distributed questioning with varied examples on exam performance on inference questions. *Educ. Psychol. Int. J. Exp. Educ. Psychol.* 29, 831–848. doi: 10.1080/01443410903310674
- Gonzalez, H. B., and Kuenzi, J. J. (2012). *Science, Technology, Engineering, and Mathematics (STEM) Education: A Primer*. Washington, DC: Congressional Research Service, Library of Congress.
- Hake, R. R. (1998). Interactive-engagement versus traditional methods: a six-thousand-student survey of mechanics test data for introductory physics courses. *Am. J. Phys.* 66, 64–74. doi: 10.1119/1.18809
- Haladyna, T. M., Downing, S. M., and Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Appl. Measure. Educ.* 15, 309–344. doi: 10.1207/S15324818AME1503_5
- Harackiewicz, J. M., and Hulleman, C. S. (2010). The importance of interest: the role of achievement goals and task values in promoting the development of interest. *Soc. Personal. Psychol. Compass* 4, 42–52. doi: 10.1111/j.1751-9004.2009.00207.x
- Hattie, J. (2009). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. Oxford, UK: Routledge.
- Hattie, J., and Timperley, H. (2007). The power of feedback. *Rev. Educ. Res.* 77, 81–112. doi: 10.3102/003465430298487
- Hays, M. J., Kornell, N., and Bjork, R. A. (2013). When and why a failed test potentiates the effectiveness of subsequent study. *J. Exp. Psychol. Learn. Mem. Cogn.* 39, 290–296. doi: 10.1037/a0028468
- Henderson, C., Mestre, J. P., and Slakey, L. L. (2015). Cognitive science research can improve undergraduate STEM instruction: what are the barriers? *Pol. Insights Behav. Brain Sci.* 2, 51–60. doi: 10.1177/2372732215601115
- Hestenes, D., Wells, M., and Swackhamer, G. (1992). Force concept inventory. *Phys. Teacher* 30, 141–158.
- Hunt, R. R., and Einstein, G. O. (1981). Relational and item-specific information in memory. *J. Verbal Learn. Verbal Behav.* 20, 497–514. doi: 10.1016/S0022-5371(81)90138-9
- Jenkins, J. J. (1979). “Four points to remember: a tetrahedral model of memory experiments,” in *Levels of Processing in Human Memory*, eds L. S. Cermak and F. I. M. Craik (Hillsdale, NJ: Erlbaum), 429–446.
- Johnson, D. W., Maruyama, G., Johnson, R., Nelson, D., and Skon, L. (1981). Effects of cooperative, competitive, and individualistic goal structures on achievement: a meta-analysis. *Psychol. Bull.* 89, 47–62. doi: 10.1037/0033-2909.89.1.47
- Kang, S. H., Lindsey, R. V., Mozer, M. C., and Pashler, H. (2014). Retrieval practice over the long term: should spacing be expanding or equal-interval? *Psychon. Bull. Rev.* 21, 1544–1550. doi: 10.3758/s13423-014-0636-z
- Kang, S. H., McDermott, K. B., and Roediger, H. L. III (2007). Test format and corrective feedback modify the effect of testing on long-term retention. *Eur. J. Cogn. Psychol.* 19, 528–558. doi: 10.1080/09541440601056620
- Karpicke, J. D., and Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science* 319, 966–968. doi: 10.1126/science.1152408
- Koedinger, K. R., Booth, J. L., and Klahr, D. (2013). Instructional complexity and the science to constrain it. *Science* 342, 935–937. doi: 10.1126/science.1238056
- Koedinger, K. R., Corbett, A. T., and Perfetti, C. (2012). The knowledge-learning-instruction framework: bridging the science-practice chasm to enhance robust student learning. *Cogn. Sci.* 36, 757–798. doi: 10.1111/j.1551-6709.2012.01245.x
- Larsen, D. P., Butler, A. C., and Roediger, H. L. III (2013). Comparative effects of test-enhanced learning and self-explanation on long-term retention. *Med. Educ.* 47, 674–682. doi: 10.1111/medu.12141
- Lasry, N. (2008). Clickers or flashcards: is there really a difference? *Phys. Teach.* 46, 242–244. doi: 10.1119/1.2895678
- Lasry, N., Mazur, E., and Watkins, J. (2008). Peer Instruction: from Harvard to the two-year college. *Am. J. Phys.* 76, 1066–1069. doi: 10.1119/1.2978182

- Lorenzo, M., Crouch, C. H., and Mazur, E. (2006). Reducing e Gender Gap In e Physics Classroom. *Am. J. Phys.* 74, 118–122. doi: 10.1119/1.2162549
- Mawhinney, V. T., Bostow, D. E., Laws, D. R., Blumenfeld, G. J., and Hopkins, B. L. (1971). A comparison of students studying-behavior produced by daily, weekly, and three-week testing schedules. *J. Appl. Behav. Analysis* 4, 257–264. doi: 10.1901/jaba.1971.4-257
- Mazur, E. (1997). *Peer Instruction: A User's Manual*. Upper Saddle River, NJ: Prentice Hall.
- Mazur, E. (2012). *Peer Instruction Workshop*. Available online at: <https://www.smu.edu/Provost/CTE/Resources/Technology/~media/63F908C3C1E84A5F93D4ED1C2D41469A.ashx>.
- McDaniel, M. A., and Butler, A. C. (2011). “A contextual framework for understanding when difficulties are desirable,” in *Successful Remembering and Successful Forgetting: Essays in Honor of Robert A. Bjork*, eds A. S. Benjamin (New York: Psychology Press), 175–199.
- McDermott, K. B., Agarwal, P. K., D'antonio, L., Roediger, H. L. III, and McDaniel, M. A. (2014). Both multiple-choice and short-answer quizzes enhance later exam performance in middle and high school classes. *J. Exp. Psychol. Appl.* 20, 3–21. doi: 10.1037/xap0000004
- McKeachie, W. J. (1974). Instructional psychology. *Annu. Rev. Psychol.* 25, 161–193. doi: 10.1146/annurev.ps.25.020174.001113
- Metcalfe, J., and Kornell, N. (2007). Principles of cognitive science in education: the effects of generation, errors, and feedback. *Psychon. Bull. Rev.* 14, 225–229. doi: 10.3758/BF03194056
- Morris, C. D., Bransford, J. D., and Franks, J. J. (1977). Levels of processing versus transfer-appropriate processing. *J. Verbal Learn. Verbal Behav.* 16, 519–533. doi: 10.1016/S0022-5371(77)80016-9
- Müller, M. G., Araujo, I. S., Veit, E. A., and Schell, J. (2017). Uma revisão da literatura acerca da implementação da metodologia interativa de ensino Peer Instruction (1991 a 2015). *Rev. Bras. Ensino Fís.* 39:e3403. doi: 10.1590/1806-9126-rbef-2017-0012
- National Research Council (2000). *How People Learn: Brain, Mind, Experience, and School: Expanded Edition*. Washington, DC: National Academies Press.
- Nicol, D. J., and Boyle, J. T. (2003). Peer Instruction versus class-wide discussion in large classes: a comparison of two interaction methods in the wired classroom. *Stud. High. Educ.* 28, 457–473. doi: 10.1080/0307507032000122297
- Nokes-Malach, T. J., Richey, J. E., and Gadgil, S. (2015). When is it better to learn together? Insights from research on collaborative learning. *Educ. Psychol. Rev.* 27, 645–656. doi: 10.1007/s10648-015-9312-8
- Packman, J. L., and Battig, W. F. (1978). Effects of different kinds of semantic processing on memory for words. *Mem. Cogn.* 6, 502–508. doi: 10.3758/BF03198238
- Perez, K. E., Strauss, E. A., Downey, N., Galbraith, A., Jeanne, R., and Cooper, S. (2010). Does displaying the class results affect student discussion during Peer Instruction? *CBE Life Sci. Educ.* 9, 133–140. doi: 10.1187/cbe.09-11-0080
- Porter, L., Bailey Lee, C., and Simon, B. (2013). “Halving fail rates using Peer Instruction: a study of four computer science courses,” in *Proceeding of the 44th ACM Technical Symposium on Computer Science Education* (Denver, CO: ACM), 177–182.
- Pyc, M. A., and Rawson, K. A. (2007). Examining the efficiency of schedules of distributed retrieval practice. *Mem. Cogn.* 35, 1917–1927. doi: 10.3758/BF03192925
- Rittle-Johnson, B., Schneider, M., and Star, J. R. (2015). Not a one-way street: bidirectional relations between procedural and conceptual knowledge of mathematics. *Educ. Psychol. Rev.* 27, 587–597. doi: 10.1007/s10648-015-9302-x
- Roediger, H. L. III (2008). Relativity of remembering: why the laws of memory vanished. *Annu. Rev. Psychol.* 59, 225–254. doi: 10.1146/annurev.psych.57.102904.190139
- Roediger, H. L. III, and Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends Cogn. Sci.* 15, 20–27. doi: 10.1016/j.tics.2010.09.003
- Roediger, H. L. III, and Karpicke, J. D. (2006). The power of testing memory: basic research and implications for educational practice. *Perspect. Psychol. Sci.* 1, 181–210. doi: 10.1111/j.1745-6916.2006.00012.x
- Roediger, H. L. III, Weldon, M. S., and Challis, B. H. (1989). “Explaining dissociations between implicit and explicit measures of retention: a processing account,” in *Varieties of Memory and Consciousness: Essays in Honor of Endel Tulving*, eds H. L. Roediger, III and F. I. M. Craik (Hillsdale, NJ: Erlbaum), 3–41.
- Rogers, E. M. (2003). *Diffusion of Innovations 5th Edn*. New York, NY: Free Press.
- Rowland, C. A. (2014). The effect of testing versus restudy on retention: a meta-analytic review of the testing effect. *Psychol. Bull.* 140, 1432–1463. doi: 10.1037/a0037559
- Schell, J., Lukoff, B., and Mazur, E. (2013). “Catalyzing learner engagement using cutting-edge classroom response systems in higher education,” in *Increasing Student Engagement and Retention using Classroom Technologies: Classroom Response Systems and Mediated Discourse Technologies*, eds C. Wankel, P. Blessinger (Bingley, UK: Emerald Group Publishing Limited), 233–261.
- Schell, J., and Mazur, E. (2015). “Flipped the chemistry classroom with Peer Instruction,” in *JChemistry Education: Best Practices, Opportunities and Trends, 1*, eds J. García-Martínez and E. Serrano-Torregrosa (Wiley-VCH), 319–334.
- Smith, M. A., and Karpicke, J. D. (2014). Retrieval practice with short-answer, multiple-choice, and hybrid tests. *Memory* 22, 784–802. doi: 10.1080/09658211.2013.831454
- Smith, M. K., Wood, W. B., Adams, W. K., Wieman, C., Knight, J. K., Guild, N., et al. (2009). Why peer discussion improves student performance on in-class concept questions. *Science* 323, 122–124. doi: 10.1126/science.1165919
- Smith, M. K., Wood, W. B., Krauter, K., and Knight, J. K. (2011). Combining peer discussion with instructor explanation increases student learning from in-class concept questions. *CBE Life Sci. Educ.* 10, 55–63. doi: 10.1187/cbe.10-08-0101
- Stuart, S. A., Brown, M. I., and Draper, S. W. (2004). Using an electronic voting system in logic lectures: one practitioner's application. *J. Comput. Assist. Learn.* 20, 95–102. doi: 10.1111/j.1365-2729.2004.00075.x
- Turpen, C., Dancy, M., and Henderson, C. (2016). Perceived affordances and constraints regarding instructors' use of Peer Instruction: implications for promoting instructional change. *Phys. Rev. Phys. Educ. Res.* 12:010116. doi: 10.1103/PhysRevPhysEducRes.12.010116
- Turpen, C., and Finkelstein, N. D. (2007). “Understanding how physics faculty use Peer Instruction,” in *AIP Conference Proceedings* (Melville, NY: AIP), 204–207.
- Turpen, C., and Finkelstein, N. (2009). Not all interactive engagement is the same: variations in physics professors' implementation of Peer Instruction. *Phys. Rev. Spec. Top Phys. Educ. Res.* 5:020101. doi: 10.1103/PhysRevSTPER.5.020101
- Tyler, R. W. (1949). *Basic Principles of Curriculum and Instruction*. Chicago, IL: The University of Chicago Press.
- Vickrey, T., Rosploch, K., Rahmanian, R., Pilarz, M., and Stains, M. (2015). Research based implementation of Peer Instruction: a literature review. *CBE Life Sci. Educ.* 14:es3. doi: 10.1187/cbe.14-11-0198
- Watkins, J., and Mazur, E. (2013). Retaining students in science, technology, engineering, and mathematics (STEM) majors. *J. Coll. Sci. Teach.* 42, 36–41. Available online at: <https://eric.ed.gov/?id=EJ1011746>
- Wheeler, M. A., and Roediger, H. L. III (1992). Disparate effects of repeated testing: reconciling Ballard's (1913) and Bartlett's (1932) results. *Psychol. Sci.* 3, 240–246.
- Wiggins, G., and McTighe, J. (2005). *Understanding by Design 2nd Edn*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Yeager, D. S., and Dweck, C. S. (2012). Mindsets that promote resilience: when students believe that personal characteristics can be developed. *Educ. Psychol.* 47, 302–314. doi: 10.1080/00461520.2012.722805
- Zingaro, D., and Porter, L. (2014). Peer instruction in computing: the value of instructor intervention. *Comput. Educ.* 71, 87–96. doi: 10.1016/j.compedu.2013.09.015

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Schell and Butler. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.