

Catalyzing Learner Engagement Using Cutting-Edge Response Systems in Higher Education

Julie Schell*, Brian Lukoff*, Eric Mazur*

ABSTRACT

In this chapter, we introduce a new technology for facilitating and measuring learner engagement. The system creates a learning experience for students based on frequent feedback, which is critical to learning. We open by problematizing traditional approaches to learner engagement that do not maximize the potential of feedback and offer a research-based solution in a new classroom response system (CRS) two of the authors developed at Harvard University – Learning Catalytics. The chapter includes an overview of cognitive science principles linked to student learning and how those principles are tied to Learning Catalytics. We then provide an overview of the limitations of existing CRSs and describe how Learning Catalytics addresses those limitations. Finally, we describe how we used Learning Catalytics to facilitate and measure learner engagement in novel ways, through a pilot implementation in an undergraduate physics classroom at Harvard University. This pilot was guided by two questions: How can we use Learning Catalytics to help students engage with subject matter in ways that will help them learn? And how can we measure student engagement in new ways using the analytics built into the system? The objective of this chapter is to introduce Learning Catalytics as a new instructional tool and respond to these questions.

Correspondance to: Julie Schell, School of Engineering and Applied Sciences, Harvard University, 29 Oxford Street, Rm 293, Cambridge, MA 02138
E-mail: schell@seas.harvard.edu

*School of Engineering and Applied Sciences, Harvard University

This chapter is (c) Emerald Group Publishing and permission has been granted for this version to appear here (mazur.harvard.edu). Emerald does not grant permission for this chapter to be further copied/distributed or hosted elsewhere without the express permission from Emerald Group Publishing Limited.

Catalyzing Learner Engagement Using Cutting-Edge Response Systems in Higher Education

INTRODUCTION

“Any questions? Anybody?”

Throughout the globe, in every institutional type and across the disciplines, instructors use this line of interrogation as a universal attempt at eliciting learner engagement. The phrase represents a well-meaning attempt by faculty to address the shortcomings of the lecture method, which has dominated higher education for more than 500 years. Indeed, lecture remains the pedagogy of choice among faculty despite consensus that interactive approaches are more effective at maximizing student learning (see [Hake, 1998](#)).

Most of the time, faculty pleas for student engagement echo unanswered throughout lecture halls. Other times, only the bravest, most confident students respond. The danger in either situation is the potential for faculty to misinterpret lack of student responses to the ubiquitous “any questions” query as understanding or learning. Armed with this false sense of teaching effectiveness, faculty may journey on, continuing to deliver material unaware of the depth of student misunderstanding left in their lectures’ wake. Worse, students usually receive little to no feedback on their misunderstandings before their first exam or assessment, leaving them without the critical insights they need to learn effectively ([Ambrose et al., 2010](#); [Bransford, Brown, & Cocking, 2000](#)).

Education does not have to be this way. In the new millennium, instructors have at their disposal a large array of technologically-advanced resources for both facilitating meaningful learner engagement and providing frequent and useful feedback. In this chapter, we examine a new technology that offers the ability to enact and measure learner engagement in novel ways. This system, Learning Catalytics, is a new classroom response system (CRS).

According to Derek [Bruff \(2009\)](#) , “classroom response systems are instructional technologies that allow instructors to rapidly collect and

analyze student responses to questions posed during class” (p. 1). Clickers, or small, handheld devices that students use to submit responses to multiple-choice questions, are examples of the most popular implementation of CRSs in education. Studies across disciplines have demonstrated that using clickers in the classroom has a number of positive benefits for learner engagement, including student attendance and participation in class (Bruff, 2009). Clickers also provide instructors and students with a low-threshold system for soliciting and delivering frequent feedback on student learning.

Learning Catalytics is a CRS that addresses many of the limitations of clickers. Instead of being restricted to multiple-choice questions, instructors can create and pose a variety of closed- and open-ended questions (e.g., enter free text, sketch a graph, highlight a passage) to students who can then use any web-enabled device (laptop, smartphone, tablet, portable media player, etc.) to respond and review immediate feedback from instructors on their personal device (Fig. 1). Learning Catalytics can also automatically and intelligently group students for discussions of their responses.

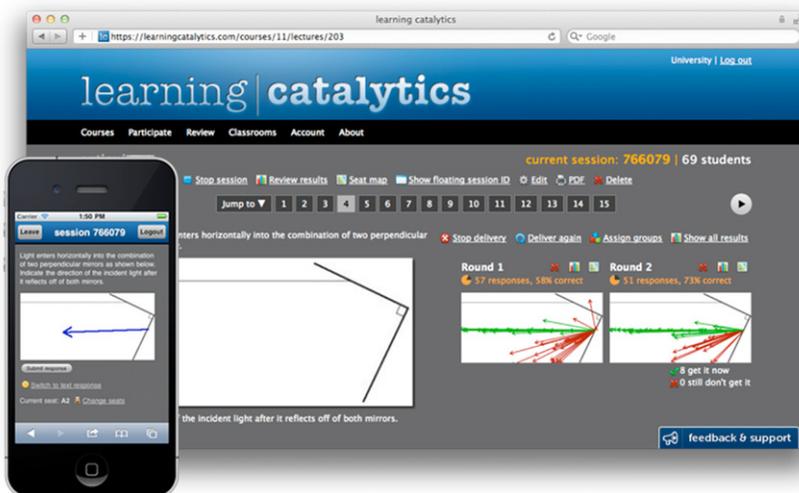


Fig. 1. Learning Catalytics, both from the instructor perspective (Large Window) and student perspective (iPhone Display). The student is responding to an open-ended graphical question by using a finger on the screen to indicate the direction of a vector. The instructor sees all student responses aggregated into a single view, and can see how student responses change from before (“Round 1”) to after (“Round 2”) discussion.

In this chapter, we explore Learning Catalytics and present pilot research we conducted in one course to suggest a model for effective technology-enabled learner engagement in higher education. The pilot collected data from 90 students in an introductory physics course at Harvard University in the spring of 2011. With the exception of the first class, we used Learning Catalytics between three and ten times during every class meeting. All students in the pilot course had access to web-enabled devices, every class period – most students used their own devices, and we purchased a small number of used iPod Touches for students that did not own a device or did not bring it to class.

We begin with a brief review of the literature on effective learning and respond to our first research question by explaining how the system can promote intellectual growth using established principles in the cognitive and learning sciences. We also respond to the second research question by offering three examples of ways faculty can gain insights into learner engagement using Learning Catalytics. After reading this chapter, faculty will be able to understand how Learning Catalytics can be used to 1) get students engaged, 2) provide them useful feedback, and 3) how to use the metrics Learning Catalytics generates to measure student engagement in novel ways.

EFFECTIVE LEARNING

How can we use Learning Catalytics to help students engage with subject matter in ways that will help them learn? For at least a decade, cognitive scientists have argued for the need for college faculty to be familiar with basic theories about how people learn and construct knowledge (Halpern & Hakel, 2003; Miller, 2011). In a 2003 article in *Change Magazine*, Halpern and Hakel note that although the professional role of faculty includes a responsibility to help students learn, most instructors are not familiar with theories that explain how learning works and/or the cognitive strategies that are associated with effective learning. In order for faculty to truly harness the power of instructional technologies to maximize learning and learner engagement, it is essential for them to have familiarity with a few key cognitive science principles.

Piaget's (1985, 1971) explanations of cognitive development persist as a dominant framework for defining learning, despite his and others' acknowledgement of limitations in his theory and method (see Bruner, 1974; Posner, Strike, Hewson, & Gertzog, 1982 ; Vygotsky, 1998 cited in

Rieber, 1998). Piaget offered a clear definition of learning as represented by cognitive growth. We define effective learning as learning that best facilitates and maximizes cognitive growth. Bjork and Bjork (2011) define learning as “the more or less permanent change in knowledge or understanding that is the target of instruction” (p. 57) but emphasize “that it is something we must try to infer” (p. 57).

What promotes effective learning? One of the most enduring Piagetian concepts for explaining the mechanisms of cognitive growth is that it occurs when there is a conflict between prior or existing knowledge and new knowledge, or disequilibrium. Bjork and Bjork (2011) further suggest that learning is enhanced when learners are presented with “desirable difficulties,” or specific learning tasks that are not too hard but not too easy. Cognitive scientists have identified specific cognitive activities, such as metacognition and the related tasks of self-monitoring and self-regulation, that catalyze disequilibrium and in turn promote cognitive growth (Ambrose et al., 2010; Bransford et al., 2000; Flavell, 1979). When these cognitive activities involve desirable difficulties, they may be particularly beneficial for long-term retention of the knowledge learned (see Bjork & Bjork, 2011).

Metacognition, Self-Monitoring, and Self-Regulation

Educators commonly use the phrase “thinking about one’s own thinking” to define the concept of metacognition. Flavell (1979) introduced metacognition as “knowledge and cognition about cognitive phenomena” (p. 906), with knowing and understanding being examples of such phenomena. Individuals with high levels of metacognition monitor their own learning, are able to recognize gaps in their existing knowledge and understanding, and are able to accurately predict their abilities to perform knowledge tasks (self-monitoring) (see Ambrose et al., 2010). Feedback plays an important role in metacognition – students must have cues to alert them to their understandings and misunderstandings (Ambrose et al., 2010; Bransford et al., 2000). In addition metacognitively-adept individuals purposefully direct their learning in ways that will help them maximize their knowledge strengths and fill gaps, a process known as self-regulation (Ambrose et al., 2010; Bransford et al., 2000; Flavell, 1979; Schoenfeld, 1987, 2010).

Extensive research on metacognition indicates that when students engage in instructor-guided metacognitive practices, learning outcomes increase (Ambrose et al., 2010; Bransford et al., 2000; Palinscar & Brown, 1984;

Schoenfeld, 1987, 2010). But despite metacognition's fundamental relationship with effective cognitive growth, learning experts suggest that instruction often neglects to build students' metacognitive capacities (Ambrose et al., 2010).

A growing area of cognitive science research suggests that when students are asked to retrieve information, such as through testing, such action "is a powerful memory modifier" (Bjork, 1975 as cited in Bjork & Bjork, 2011). This research suggests that despite its powerful effects, retrieval-based testing is also underutilized as an instructional tool; typically, teachers use testing solely as a grading, versus a teaching or learning, instrument (Bjork & Bjork, 2011).

The effects of retrieval may be even more powerful when students are asked to "generate" an answer (Bjork & Bjork, 2011), rather than recognize it as correct (as they do when responding to multiple-choice questions) (Duchastel, 1981; Glover, 1989; Kang, McDermott, & Roediger, 2007; McDaniel, Derbish, & Morrisette, 2007). When students are asked questions that require them to retrieve and produce information, rather than simply retrieve and recognize it, they are engaging in effortful retrieval, a different and deeper exercise of self-monitoring (Duchastel, 1981). Before they can produce the information a question prompt elicits, they must ask themselves, "do I know this? Or not?" As they "generate" (Bjork & Bjork, 2011) rather than select a choice they believe is right, they may be further cued as to areas of their existing knowledge that are weak or strong. Once they finish their response to the prompt, students' awareness of their ability to answer those questions gives them immediate feedback as to what areas they might target for further practice. In addition, when students respond to questions using CRSs and faculty provide feedback on the right answer, they are further made aware of the strengths and limitations of their understanding. In this way, giving students opportunities to retrieve and produce responses may build their metacognitive skills.

Learning Catalytics and Metacognition

Learning Catalytics, when used with question-based teaching methods, allows instructors to engage students in hundreds of metacognitive activities throughout a course by getting students to generate responses to a mini-test or question and then, following an intense discussion with one or more peers, reevaluate their original response in the context of the discussion and consider whether their own thoughts have changed. The wide array of constructed-response (non-multiple-choice) question types available in Learning Catalytics allows students to engage in the deeper metacognitive

tasks associated with effortful retrieval, rather than recognition (although the system can easily deliver multiple-choice questions as well). In addition, the system also provides another way for students to engage in self-monitoring: students can select an option on their devices to report to instructors right away, in class, “I get it now” or “I still don’t get it” after a round of voting and discussion. If students “still don’t get it” – or if they have a question they simply do not want to raise in front of the whole class – they can also self-direct in class by sending a question to the instructor, without waiting for their instructor’s “any questions” query and without having to raise their hands or say the question out loud in front of their peers.

Using this new CRS, faculty have multiple lines of evidence to determine whether their students are understanding or not: if they have data suggesting most students responded correctly, and they receive no “I still don’t get it” responses and no questions through the system, instructors can proceed with more confidence than is possible by simply soliciting “out loud” questions from the class.

Effective Learning Indicators

Actual cognitive growth occurs within many levels of the human brain, represented by subtle changes in brain structure (Zull, 2006). In particular, cognitive growth is evidenced by developments in the neocortex, the outer layer of the brain responsible for sensory perception (Zull, 2006). Of course, measuring effective learning by measuring changes to students’ brain structures poses a challenge for anyone who is not a neuroscientist. So what indicators of effective learning are more realistic to measure in educational settings? In this chapter, we discuss two: conceptual change and knowledge transfer.

Conceptual Change

Philosophers of science have suggested that conceptual change is an important indicator of learning (Posner et al., 1982). Conceptual change occurs when students, often aided by specific instructional strategies (Heron, Shaffer, & McDermott, n.d.), resolve conflicts between their preexisting conceptions and the concepts they are taught, and when these new conceptions are sustained (Posner et al., 1982). When students possess rigid preexisting conceptual frameworks – a particularly pervasive phenomenon in the science disciplines – it is not enough to simply resolve their

misunderstanding by offering an alternative framework (Posner et al., 1982), or by telling them the correct answer. Indeed, if simply telling students were enough to evoke conceptual change and cognitive growth more generally, empirical evidence would support the use of lecture as a more effective teaching method than current research suggests (see Hake, 1998; Mazur, 2009).

One research-based strategy for creating conceptual change is the elicit, confront, and resolve process developed by the McDermott Group at the University of Washington (Heron et al., n.d.). In this approach

the first step is to elicit a known difficulty by contriving a situation in which students are likely to make an error that exposes that particular difficulty. If the difficulty is sufficiently serious and not addressed, it may remain latent and arise in other contexts. It is therefore the responsibility of the instructor to insist that students confront and resolve underlying difficulties. (pp. 2, 3)

Although conceptual change is just one of many indicators of effective learning, it is an important one to consider because student conceptions are often powerfully resistant to change (Posner et al., 1982).

Knowledge Transfer

Wiggins and McTighe (2005) discuss another indicator of effective learning: knowledge transfer, or the ability to successfully navigate novel contexts using existing knowledge. For example, a student who uses her prior knowledge to solve new, foreign problems she has never seen before is exhibiting knowledge transfer. The ability to transfer knowledge may also be an indicator of conceptual change: presumably, a student who has experienced a permanent conceptual change would be better able to solve novel problems that require the use of new conceptual frameworks than a student who is still working from an old and “incorrect” framework (Posner et al., 1982). Bransford et al. (2000) cite an extensive research literature that suggests that the metacognitive practices of self-monitoring and self-regulation “have been shown to increase the degree to which students transfer their learning to new settings and events” (p. 12).

Learning Catalytics and Conceptual Change and Knowledge Transfer

Learning Catalytics offers numerous facilities for instructors to promote conceptual change and knowledge transfer. For example, instructors can use the system to construct numerous question types that elicit students’ misunderstanding; facilitate student activities, such as intelligently paired student discussion (where students are assigned to discuss their responses in

non-arbitrary groups that the instructor believes are most likely to benefit from a discussion with each other) that promote the confrontation of those misconceptions; and provide students with resolution of misconceptions, including explanations and the reasoning underlying correct answers.

In addition, instructors can use this CRS to provide students with opportunities to practice knowledge transfer during every class period, by posing specific questions that require them to apply what they know in new contexts. Finally, Learning Catalytics provides a robust feedback mechanism that immediately and frequently alerts students and instructors of student progress.

Feedback

While students can and do learn through passive engagement (Bonwell & Elison, 1991), it is no secret that they learn more when they are actively engaged. Bonwell and Elison (1991) posit that active learning occurs when students are involved in activities beyond passive listening and note taking, including reading, discussing, and writing, and when students receive “immediate feedback” from their instructor. And when it facilitates self-monitoring and self-regulation, feedback promotes student engagement; in other words, feedback helps engage students when it pulls them into activities that allow them to both self-assess their knowledge and consider what they need to do to improve their current learning state (Ambrose et al., 2010; Bransford et al., 2000). According to Ambrose et al. (2010), feedback is most effective when it signals to students the specific knowledge and skills they are expected to learn; is given to students at times and frequencies that are aligned with when they are most likely to need and use it; and provides additional and specific opportunities for students to practice.

Learning Catalytics and Feedback

Learning Catalytics is, at its core, a feedback engine. When instructors use the CRS to pose conceptually-based questions, they signal to students several levels of expectation: the general topic of importance (which also occurs in traditional lecture); the deeper conceptual understanding the instructor expects students to build; and what students should be able to do with that understanding. Second, in traditional educational settings, instructors only periodically check in on students’ understanding through formal quizzes and exams. In those settings, students usually do not engage in an “elicit, confront, resolve” process or receive corrective feedback on

their responses in a timely manner. Timely, immediate feedback is important in order to convert student misunderstandings into learning (Epstein et al., 2002). Learning Catalytics allows instructors to give students immediate feedback on their performance on conceptual tests and collects performance data in a student dashboard for later review. Students can review such feedback for later directed practice and self-study.

Peer Instruction

Learning Catalytics is designed around the approach of Peer Instruction (Mazur, 1997). The next section discusses the Learning Catalytics technology in more detail, but first we present an overview of Peer Instruction and some of the research results that have led thousands of instructors around the world to adopt this pedagogy in their own classrooms. Peer Instruction emphasizes teaching by questioning. Instructors pose questions to students in class; for each question, students commit to an initial answer, and then students are tasked with discussing their responses with a neighbor and trying to convince that neighbor that their response is correct. After the discussion, students are prompted to respond to the question again. Typically, second round (post-discussion) responses are correct at a much higher rate than first round (pre-discussion) responses, suggesting that discussions are a productive way for students to learn (Lasry, Charles, Whittaker, & Lautman, 2009; Smith et al., 2009).

Different types of groups can be effective at promoting learning. If a question has a correct answer (which is not true of all questions, especially outside of STEM disciplines), we consider a group effective when its members answer correctly in the second round, especially after answering incorrectly in the first round. Surprisingly, there is evidence (see Smith et al., 2009) that groups where all students initially give incorrect responses can be effective. Even in groups where one student has initially responded correctly, none of the group members are given right/wrong feedback about their responses prior to their discussions. As a result, students usually must argue based on the merits of their responses and reasoning; since group members do not know who gave the correct response (if anyone did), a student cannot simply blindly copy another student's response.

Peer Instruction as it is described above is one of the most popular of a class of so-called "interactive teaching" or "interactive engagement" methods (see Henderson & Dancy, 2009); in a meta-analysis of about 6000 students, Hake (1998) identified a large number of classes that administered conceptual inventories both at the beginning and end of the

course and found that interactive engagement methods produce student conceptual gains that are on average about double those for traditional teaching methods. Work on Peer Instruction specifically has shown substantial and stark differences in conceptual gains over time for two courses that switched between traditional lectures and Peer Instruction (Crouch & Mazur, 2001). Research also demonstrates that Peer Instruction contributes to increased student retention in STEM majors (Watkins, 2010).

Limitations of Existing Classroom Response Systems

The effectiveness of Peer Instruction at both the micro (individual question response pattern) and macro (conceptual gain) levels suggests that Peer Instruction provides a solid pedagogical foundation for building an interactive teaching technology. Traditionally, students respond to questions in Peer Instruction by using clickers, by using flashcards (by holding up one off our or five flashcards either labeled A through E or differently colored), or by placing a number of fingers in front of one's chest so that they are visible by the instructor but by no one else. The latter two methods have logistical disadvantages. For example, using flashcards or raised hands does not allow the instructor to collect, save, or accurately analyze responses – instructors can only take a visual survey. Clickers, on the other hand, quickly and automatically collect student response data and, when used well, can promote student engagement in retrieval practice (see Bjork & Bjork, 2011). When it comes to conceptual gain, however, according to one study, clickers do not seem to have any appreciable impact when compared to flashcards (Lasry, 2008).

However, both clickers and non-technology CRS methods like flashcards share a number of limitations. First, they are usually limited to multiple-choice questions. Constructed-response questions – questions where students must generate a response from scratch, rather than simply recognize one of the given response options as correct – can push students to engage in different cognitive processes than multiple-choice questions, although this depends on the ways in which the question author makes use of a particular question format (Martinez, 1999). Perhaps even more importantly, question content and format make a statement to students about what is important, and few instructors are likely to agree with the notion that the goal of their course is to teach students how to answer multiple-choice questions correctly, no matter how cleverly constructed. In addition, multiple-choice questions are particularly difficult to write, because they require the instructor to develop appropriate distractors (Haladyna, 2004). Most instructors are not

professional question writers and as a result the writing of good multiple-choice questions is particularly challenging.

A second limitation of clickers is that while they provide a mechanism for collecting response data, they do not guide the instructor to create an effective interactive teaching environment. For example, the traditional approach to Peer Instruction – telling students to simply “turn to your neighbor” – frequently results in the formation of groups that are created out of convenience: students often turn to their friends, or they simply turn to the same neighbor each time, and as a result often miss out on a potentially more pedagogically beneficial discussion that they might have with a different group.

In addition, the use of clickers or any non-technological CRS method requires that instructors closely monitor the timing of questioning so that the class moves at a good pace while also monitoring the incoming responses to plan the next phase of instruction. Two common approaches are to either wait for the last few students to respond, making the majority of the class bored and fidgety, or to use a timer (e.g., 30 seconds per question), which generally will not match the actual time needed on a broad variety of questions.

Finally, some of the existing CRS technology makes it challenging for instructors to make the most of the formative assessment data collected. Non-technology CRS methods do not save any information about student responses, of course: the instructor can only make a mental note of the approximate percentages of students holding up each kind of flashcard. But it can even be challenging to make use of the data collected using some kinds of existing response technologies: the most useful data is often the longitudinal data of student performance between class sessions, and this data is not always presented in pedagogically useful ways.

Learning Catalytics

Learning Catalytics’ CRS features are designed to help instructors create more authentic and engaging interactions with students in the classroom, regardless of class size. Learning Catalytics is web-based and hosted in the cloud; all data is stored centrally and securely. Instructors control the delivery of course material using a computer or iPad, and students participate in class using any modern web-enabled device:

- Laptop computers
- Smartphones, including iPhone, Android, and Blackberry phones

- Tablets, including iPad
- E-readers, including Amazon Kindle
- Any other modern web-enabled device (e.g., iPod Touch)

Both students and instructor access Learning Catalytics through their device's web browser, so there is no software to install or configure. Because these devices provide the opportunity for complex two-way interactions, they provide a medium for Learning Catalytics to overcome the limitations described in the previous section and create a flexible platform for formative assessment, Peer Instruction, and other interactive teaching pedagogies.

Learning Catalytics was first piloted in a medium-sized physics class (about 90 students) taught at Harvard University in Spring 2011. During the 2011–2012 academic year, Learning Catalytics was used by a number of instructors in both secondary and higher education across different disciplines and in both small and large classes.

Instructor Workflow

In typical use, the instructor uses the laptop computer used to display slides by connecting the laptop to the classroom projector and placing the laptop into dual monitor mode. The instructor is provided with an instructor “control panel” – updated in real time with information about student responses – as well as a “student window” that displays the current question without prematurely revealing information about student responses before the instructor is ready to make that data available to students. The instructor configures the display so that the control panel displays on the laptop monitor and the student window displays on the classroom projector.

Typically, instructors prepare the questions that will be posed to students in advance of class. During class, instructors can deliver questions to students in any order. When ready to pose a question, the instructor clicks a “Deliver” button and the question is posted in the student window and simultaneously pushed out to student devices. As students respond, the control panel updates in real time with a visual representation of student responses. (The format of the visual representation depends on the format of the question; for a simple multiple-choice question a histogram of student responses is shown, but for constructed-response questions the visual representation will be some sort of appropriate aggregation of student responses.) While students are responding, the instructor may also call up a

seating chart showing the arrangement of seats in the classroom, with seats colored red, yellow, or green to indicate which students have responded to the question correctly, partially correctly, or incorrectly.

An instructor wishing to engage students in Peer Instruction discussions can click a button to automatically assign groups based on a grouping policy; for example, the instructor may want to pair students so that at least one student in the pair responded to the question correctly. (At this point, students have not been given any feedback about the correctness of their responses, so even if students know the grouping policy they do not know whether it is they or their partner that responded correctly.) Learning Catalytics uses the geographical location of students within the classroom when assigning students to groups, and issues personalized messages on the students' devices indicating who they should turn to for discussion of their responses. Students are then given a second opportunity to respond to the same question.

Finally, when ready to "wrap up" the discussion of a question, the instructor can push the visual representation of student responses to each student's device. This is the same visual representation that previously was visible only to the instructor. Based on this visual display, the instructor can engage the entire class in a discussion of the question and the responses received. The most common use of this display is to explain to students what the right answer is and why; for constructed-response questions, student responses provide a window into the specific misconceptions that students brought to the question, and instructors can use this data to address those misconceptions in their explanation. For questions where there is no right answer, the visual representation of student responses might be used to generate debate about the question; since students have all responded to the question individually prior to the debate, they already have a "stake" in the debate, energizing the discussion.

For constructed-response questions, another use of this visual representation of student responses is to highlight a particular response given by one or more students, and to engage the class in a discussion of that response. For example, when each student has sketched a graph in response to a question prompt, one possible visual representation is a set of "thumbnails" of each student graph. The instructor can select one particular graph, enlarge it, and publicly analyze its features; since the students do not know the identity of the author of the graph, the focus is on the response itself and not on the student that constructed it. In this way, students can feel more comfortable about actively participating in class.

Student Workflow

From the student's perspective, Learning Catalytics is used in a manner similar to a traditional clicker. This simple design was intentional, as it allows students to be up and running using the system with no help from the instructor – essential in a large class where it is simply impractical to work with students individually to help them configure their devices. When students enter the classroom, they point their devices to lcatalytics.com and they are prompted to log in and enter the number of the class session to join. After logging in, they are prompted to indicate where in the classroom they are sitting (Fig. 2.) – this information is used both to group students for discussion and to aid the instructor in managing and

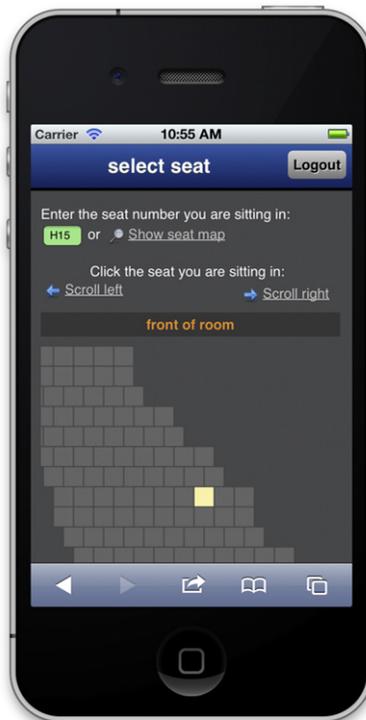


Fig. 2. A student selecting a seat when logging into Learning Catalytics.

monitoring the class. Once a student selects a seat, the student's device indicates that they should wait for a question to be presented. When the instructor is ready to pose a question to the class, it is "pushed" to student devices automatically and the question appears immediately on each student's device.

When the instructor is ready to wrap up the discussion surrounding a particular question, the class results are pushed to student devices, so that students can inspect the results on their own screens. While the instructor is engaging in a wrap-up activity for a question – typically a whole-class discussion of the different answers – students also have the opportunity to press buttons labeled "I get it now" and "I still don't get it" to provide asynchronous feedback to the instructor. Tallies of the numbers of students pressing each of these buttons are made available to the instructor in real time, so the instructor can calibrate the post-question activity properly: if the instructor has just finished an explanation of the correct reasoning behind a question and is ready to move on, a plurality of students indicating "I still don't get it" can prompt the instructor to circle back and perhaps spend more time on the concepts underlying that particular question; conversely, if the instructor is about to start a long discussion of a concept because of poor student performance, a chorus of students indicating "I get it now" may suggest that a shorter activity is warranted.

Students also have the ability to provide asynchronous feedback in another way. At all times during the class, students have access to a button labeled "Send a message to the instructor." This allows the student to asynchronously send a question or comment that shows up in an unobtrusive way on the instructor's screen. This encourages students who are reluctant to raise their hand in class – or students that have a comment or question that may not be appropriate to raise directly – to communicate with the instructor, and the asynchronous nature of the communication means that the instructor can respond to the question as they see fit: in the moment, later in the class period, or even by email to the student after the lecture ends.

After class, students can log back into the system to review what happened in class: for each question, they can see the question itself and their response(s) to that question. Instructors can optionally write an explanation of the correct answer to accompany each question, and students will see this explanation when reviewing questions only after class. This makes the active learning activities that take place in the classroom more permanent, as the tool becomes a study aid for students to use when working on homework or preparing for exams.

Overcoming the Limitations of Existing Technologies

Going Beyond Multiple-Choice Questions

Learning Catalytics permits instructors to ask questions in a wide variety of constructed-response formats, incorporating question formats that are adaptable to many different disciplines and questioning styles. The goal in designing Learning Catalytics is not to provide different question formats simply for variety's sake, but to provide question formats that will facilitate higher-order cognitive activities and make asking certain kinds of questions as natural as possible.

Many of the different question formats in Learning Catalytics were developed as we piloted the system; we analyzed existing multiple-choice questions that the instructor planned to use in class and developed formats to make those questions as open as possible. We did this by creating an environment where we asked students to construct a response rather than selecting from among several given choices. For example, one multiple-choice question asked students which direction an electric field is pointing in, and provided response options such as "the positive x-axis," "the negative x-axis," "the positive y-axis," etc. By converting this question to a constructed-response question where students simply click (or tap their screen) to indicate a direction, we can shift the focus from selecting an answer (which can involve a cognitive process adapted from standardized test-taking strategy) to the problem itself. By providing this more authentic environment for responding to the question, we were able to elicit misconceptions that were not possible in the multiple-choice format (i.e., from students whose response did not match one of the original response options), and this can improve the in-class interactions when this question is used in class.

Fig. 3 shows four different examples of constructed-response questions: the larger images show the instructor control panel display – which updates with responses in real time – and the smaller iPhone images show what a student sees.

- Direction questions (upper left), where students indicate the direction of an arrow or vector. This question, based on an existing activity ([TERC, n.d.](#)), asks students to study a satellite image and determine in which direction the winds blow. Examining the results show that there are clusters of responses, which, a priori, may not have been clear to a question author.
- Sketch questions (upper right), where students sketch a graph or diagram. In the implementation shown here, instructors can see thumbnails of

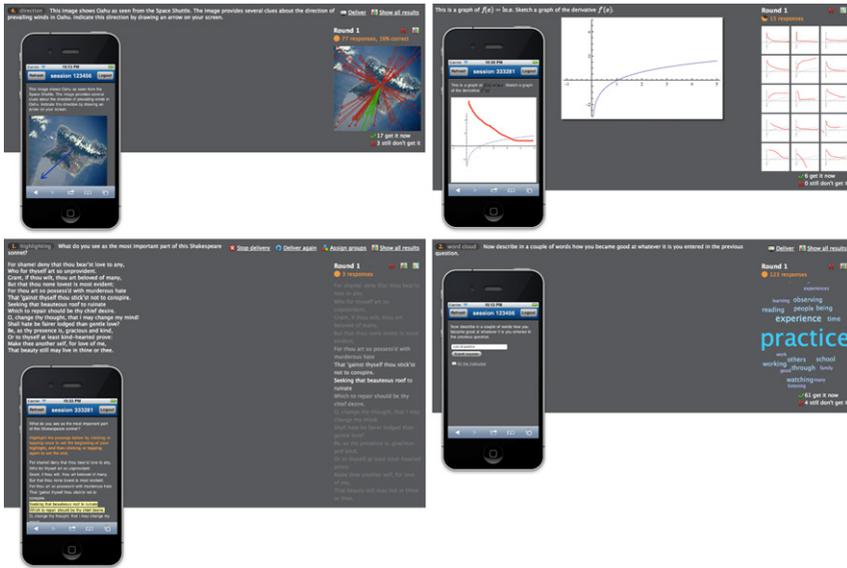


Fig. 3. Four examples of different types of open-ended questions that can be posed to students using Learning Catalytics. The larger view shows the instructor control panel view of the question, with a set of sample results, and the smaller view shows the student response view.

individual sketches; another version – designed for larger classes where it is not feasible to look at all student sketches at once – allows instructors to see a single “composite sketch” where all sketches are semi-transparently overlaid on top of one another. This latter mode is particularly useful for getting a sense of the common features of student sketches.

- Highlighting questions (lower left), where students highlight part of a text passage. Students might be asked to identify the part of the passage that indicates something in particular, to find a case of a particular literary structure, or simply to indicate what they find to be the most salient, interesting, or inspiring part of the text. The visual representation of student responses is a heat map where the brightness of each part of the text is in proportion to the number of students that highlighted it.
- Word cloud questions (lower right), where students enter a textual response and a word cloud is displayed where words are sized relative to their popularity in student responses. Word clouds such as these provide a

quick way to process textual responses in a large class where it is not feasible to read each individual student response.

As described previously, instructors can encourage more productive Peer Instruction discussions by automatically assigning students to groups for discussion. Our informal observations suggest that when students receive the grouping directions on their device (e.g., talk to Alice in front of you and Bob to your left), the social experience can be different from when the instructor gives generic instructions to “talk to your neighbor”: providing explicit directions has the potential to break the ice for students who do not know each other and get students engaged in discussion quickly. In this way, the instructor obtains more control over the way groups are formed, because the grouping parameters are controlled by the instructor.

As a result, the instructor is freed up to use the time during Peer Instruction discussions for more interesting pedagogical purposes. The instructor can consult the graphical seating chart on their personal screen and note if there are clusters of students in certain areas of the room that all responded incorrectly before discussion; during discussion, the instructor can seek out those students or groups to listen in to their discussions to better understand student misconceptions – or to drop a hint or otherwise help guide the discussion.

The instructor can also ask the system to manage the time used for questioning. Normally, instructors and students are shown a timer on the projector screen that tracks how much time has already elapsed in that round of questioning; students are also shown a pie chart that tracks the fraction of class that has already responded to the question. If desired, Learning Catalytics can make use of a predictive model that watches the response data and at some point switches the count-up timer to a count-down timer; the amount of time initially shown is calculated based on the response pacing so that the remaining time should permit the vast majority of students to respond to the question. When the timer counts down to zero, students will either be automatically grouped for Peer Instruction or simply shown the results of the questioning, depending on parameters set by the instructor. In this way, instructors are freed from having to carefully monitor and manage the time, and can more productively take the time during student responding to go around the room and monitor student discussions or simply take a

moment to think about how they want to structure the remainder of the class period.

Providing Actionable Data Analytics for Instructors

When instructors use Learning Catalytics, a large amount of formative assessment data is collected about students, both individually and as a whole. One focus of the design of Learning Catalytics is to make that data easily actionable by aggregating it in useful ways. In particular, it allows the instructor to easily browse through the data and take advantage of the hierarchical nature of the assessment measurements. Fig. 4 shows the course

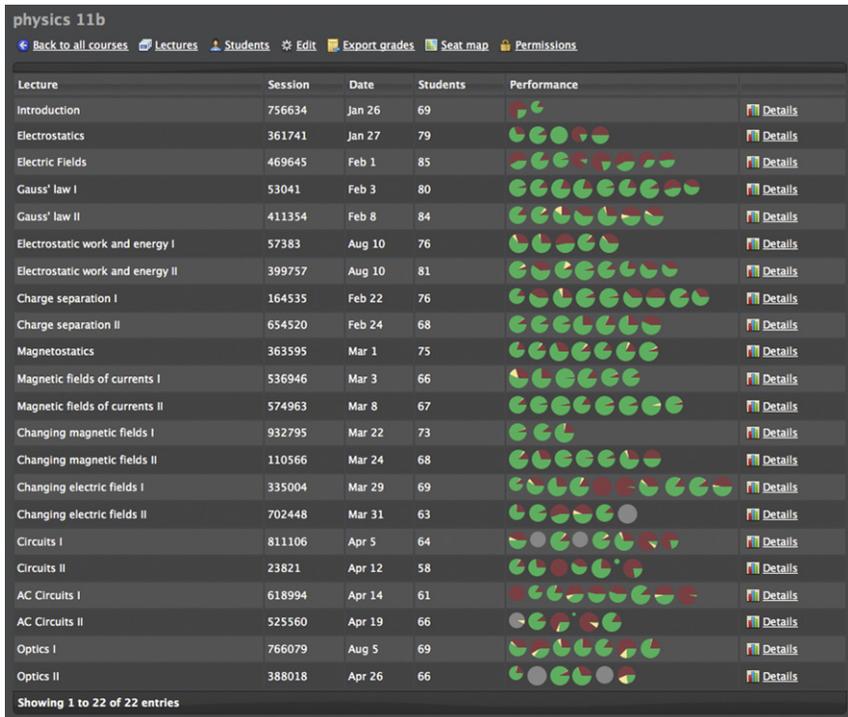


Fig. 4. The course dashboard. Each row represents a class session, and each small pie chart represents a question asked during that session.

dashboard for our pilot Learning Catalytics course; each row in the table corresponds to a class session. Each small pie chart represents a question that was asked during that lecture, with the size of the pie proportional to the number of students participating; if the question was asked twice with a Peer Instruction discussion in between, then the post-discussion results are shown. When it appears onscreen, green corresponds to students who answered the question correctly, red corresponds to students who answered incorrectly, yellow corresponds to students who answered incorrectly but subsequently pressed the “I get it now” button, and gray corresponds to questions where the response was not scored as correct or incorrect (questions where there is no right answer, or question formats where Learning Catalytics is not currently able to automatically score student responses).

By scanning the course dashboard, the instructor is quickly able to identify topics that students found particularly challenging or where the questions were simply particularly difficult (in this example, AC circuits), or whether there are trends in student participation over time. Hovering the mouse over any of the pie charts shows the question itself, and clicking on a pie chart allows the instructor to drill down into the individual student-level responses for the corresponding question. In this way, an instructor can review the course at a glance to help plan future instruction over the remainder of the semester, or to identify which topics would be ripe for review.

Similar dashboard functionality is available at the student level. An instructor can view a graphical display of an individual student’s performance throughout the semester, and can use these data to zero in on the student’s particular difficulties in the course, determine if there are performance or participation trends over time, and quickly and visually evaluate how well students work in Peer Instruction groups.

Using Learning Catalytics to Measure Learner Engagement

Typically, response technologies are used to measure student understanding (e.g., what misconceptions do students hold about this difficult idea?) or sentiment (e.g., what is the range of opinions in the class about this controversial topic?), and in the pilot physics course the former was the target of interest. However, we wanted to use the formative assessment data not just to understand student learning but also to begin to understand student engagement and metacognition. In this section we describe three

ways that we can use the formative assessment data to start to illuminate these more elusive constructs.

Degree of Engagement

As a first pass at identifying low engagement, we examined student response times – the amount of time that elapses between the time a question is made available to students and the time a particular student actually responds to the question. Since in the pilot course students received participation credit for responding to the questions, we hypothesized that low-engagement students might simply respond to questions without putting a great deal of thought into them (e.g., randomly, or with only cursory reasoning). Such students might frequently enter a response only a few seconds after the question has been read – or wait until there are only a few seconds remaining to enter a response. In the pilot course, we measured for each student the average fraction of the available time that the student waited before responding. A value of 50% indicates that the student waited on average half the available time before responding, a value of 0% indicates that the student is always the first to respond, and a value of 100% indicates that the student is always the last to respond. Fig. 5 shows the distribution of average response waiting periods. Students at the tails of this distribution are unusually fast or slow to respond, which may suggest disengagement – although we were surprised to discover that in our pilot class the distribution is relatively smooth and no students in particular stand out.

Another possibility for measuring degree of engagement might be to compare responses to multiple-choice questions with responses to open-ended questions, to look for students that appear to be participating in multiple-choice questions but that do not give coherent responses to open-ended questions.

Knowing What You Don't Know

We can also use the data to identify (lack of) student metacognition. As [Kruger and Dunning \(1999\)](#) found in their classic study, people that are low-performing in a particular area often believe that they are higher performing than they are in reality. A Peer Instruction environment ([Mazur, 1997](#)) where students respond twice – and discuss their responses with each other

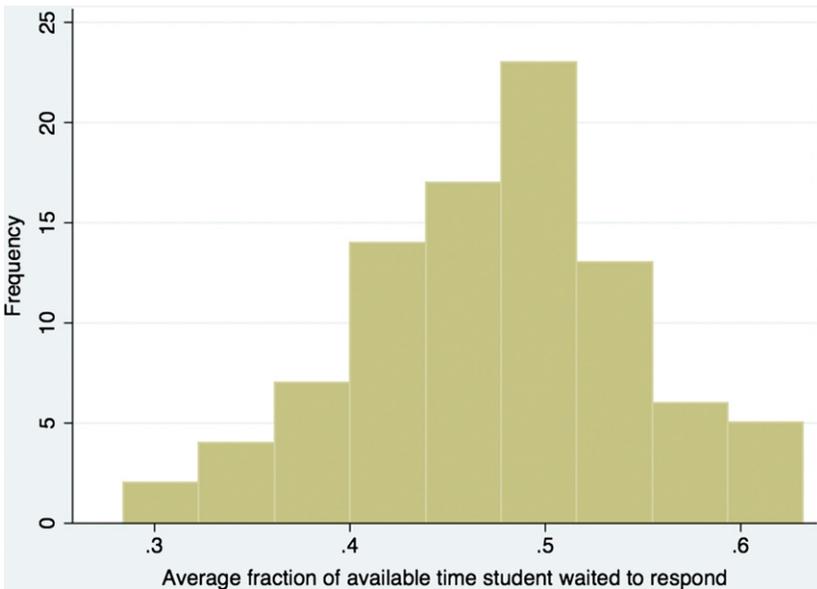


Fig. 5. Distribution of average waiting periods (the amount of time students waited before responding to a question) in our pilot course.

after their first response – provides a valuable opportunity to obtain a measurement of students’ metacognition. In particular, we can identify rigid learners by finding students who frequently respond incorrectly twice even after having a discussion with another student who responded correctly. Fig. 6 plots a histogram of the proportion of the time that students fail to learn from a knowledgeable partner among the $N = 48$ students in the pilot class that responded incorrectly before discussion at least 10 times over the course of the semester. Using this metric, we can clearly identify a small number of students (5 out of the 48) that failed to learn from a more knowledgeable partner at least 50% of the time. This suggests that some small proportion of students fairly consistently do not recognize when their partner has the correct answer or reasoning. Of course, this is only a small part of a student’s metacognitive toolset, but this metric could help identify a metacognitive deficiency in a rigid learner that is also performing poorly along traditional metrics (e.g., exam performance).

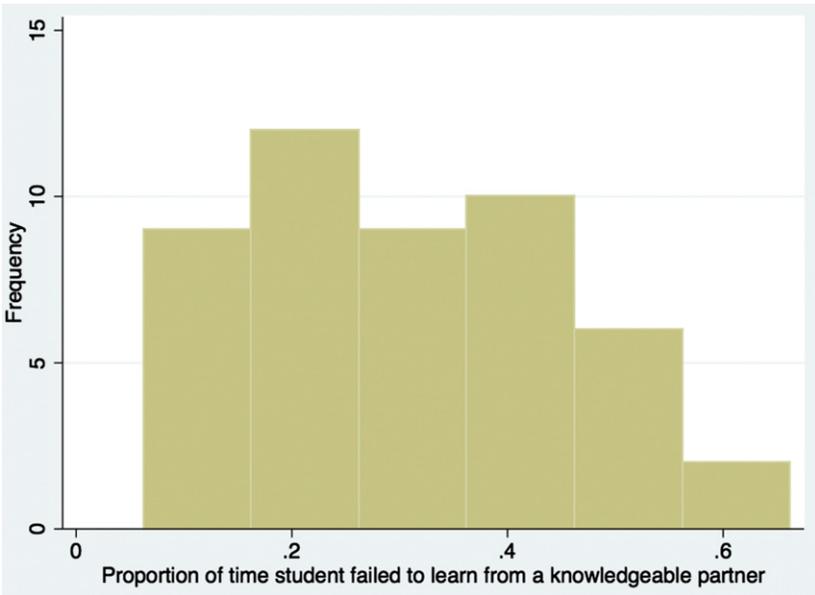


Fig. 6. Distribution of the proportions of the time that students failed to learn from a knowledgeable partner in our pilot course.

Natural Teachers

Some students, after responding correctly the first time, can fail to convince their discussion partner of their answer, while others are natural teachers who are consistently able to explain their answer cogently. It may be possible to identify these natural teachers through the Peer Instruction data: in our pilot class, we looked at all Peer Instruction sessions where a student has a discussion partner who answered incorrectly before discussion, and we measure the proportion of those occasions when the partner then answered correctly after discussion. Fig. 7 shows the distribution of this proportion across all students in the course. Interestingly, the distribution appears to be roughly normal, but with a fairly wide spread: on average, students' partners improve about half the time, but some students are able to get their partners to improve nearly 85% of the time and others are able to get their partners to improve only about 25% of the time. This is interesting as a phenomenon for further research, but could also be relevant as a measurement tool for a particular instructor: identifying the "natural

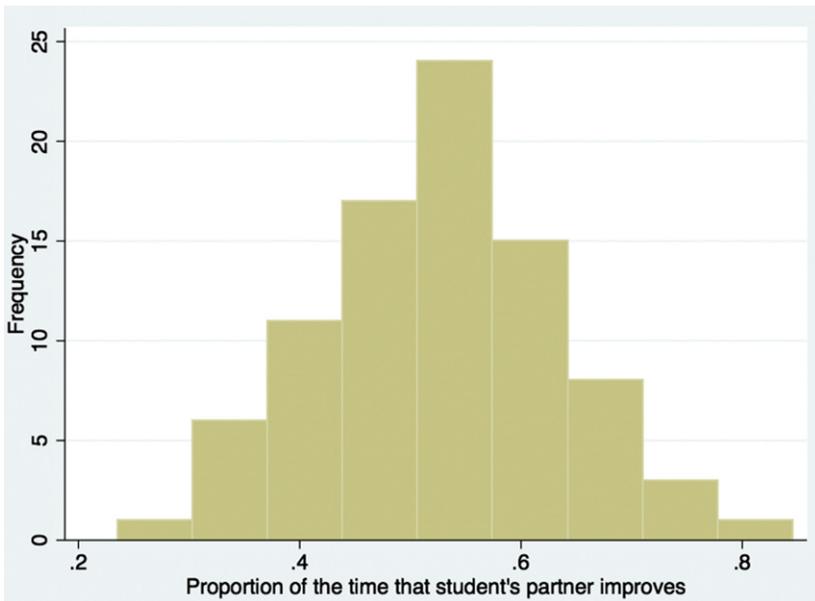


Fig. 7. Distribution of the proportions of the time that a student's partner improved when comparing their post-discussion response to their pre-discussion response.

teachers" could be a valuable device in constructing persistent groups for projects, helping students to form study or review groups, or even for advising students on potential career ideas.

A Limitation of the Pilot Course Data

We have observed that students sometimes fail to form the groups they are directed to form. In our pilot course, which used an early version of the platform, students were required to enter their seat number without the aid of a graphical seating map (Fig. 2), and students often did not enter the correct seat. As a result, students were sometimes assigned to groups with students that they were not actually sitting next to. Other times, students spoke with other adjacent students that were formally in another group, or students spoke with a member of the teaching staff. Therefore, the grouping data in our pilot course is extremely noisy, and the particular results we obtained above using this grouping data are intended to be illustrative of the

techniques that are possible rather than a conclusion about the particular students in our pilot course.

CONCLUSIONS

This chapter introduces a new CRS that facilitates learner engagement in ways not possible with clickers or non-technology CRSs. We described the pedagogical basis for Learning Catalytics, particularly its ability to engage students in a diverse array of tasks, well beyond multiple-choice and traditional open-ended questions. We also described a pilot study where we posed research questions and answers about how Learning Catalytics might both help students learn and help instructors measure learner engagement. Future work will include collecting more data in different classrooms and disciplines to help us both further respond to these questions and pose others.

Implications for Practice

The measures highlighted in this chapter can provide faculty with extensive feedback on students' varying levels of engagement and also cue faculty on how to intervene. For example, Learning Catalytics can identify students who appear to have less metacognitive skill than others. Equipped with this information, faculty could develop metacognitive interventions, such as effortful retrieval practice, to help build this capacity – a capacity that is the mark of a master learner. Second, faculty can use Learning Catalytics to conduct fine-grained analysis on student learning and engagement. By making use of the formative assessment data provided by open-ended questions (e.g., the graphs students draw, the text they compose, or the passages they highlight), faculty can understand student learning at a deeper level and in ways that can enrich instruction.

As the use of older CRSs such as clickers continues to grow, so too will the awareness of their limitations. Learning Catalytics is already poised to offer a more robust, research-based instructional technology for driving learning and learner engagement in ways that no other technology currently can. The future of higher education will be shaped by faculty's willingness to turn their back on the time-honored lecture method and turn toward cutting-edge pedagogies and instructional technologies that arm them with strong and detailed data on students' levels of understanding and engagement, freeing

instructors from ever having to experience the deadening silence that so often accompanies the question, Any questions? Anybody?

ACKNOWLEDGMENTS

We would like to recognize the significant work of several individuals who have contributed to the development of Learning Catalytics and the framework for this chapter: Learning Catalytics co-founder Gary King, members of the education subgroup of the Mazur Group at Harvard during the time of the pilot study (Ives Araujo, Jason Dowd, James Fraser, Jungsook Lee, Kelly Miller, Alvaro Neves, and Laura Tucker), and Alyssa Reinhart of the University of Texas at Austin.

REFERENCES

- Ambrose, S., Bridges, M. W., Dipietro, M., Lovett, M. C., Norman, M. K., & Mayer, R. E. (2010). *How learning works: Seven research-based principles for smart teaching* San Francisco, CA: Jossey-Bass.
- Bjork, E. L., & Bjork, R. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society*(pp. 56–64). New York, NY: Worth Publishers.
- Bonwell, C. C., & Elison, J. A. (1991). *Active learning: Creating excitement in the classroom*. ERIC Digests. Retrieved from <http://www.oid.ucla.edu/about/units/tatp/old/lounge/pedagogy/downloads/active-learning-eric.pdf>
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school* (expanded edition) Washington, DC: National Academy Press.
- Bruff, D. (2009). *Teaching with classroom response systems: Creating active learning environments* San Francisco, CA: Jossey-Bass.
- Bruner, J. S. (1974). *Toward a theory of instruction*. Cambridge, MA: Harvard University Press.
- Crouch, C. H., & Mazur, E. (2001). Peer Instruction: Ten years of experience and results. *American Journal of Physics*, 69, 970–977. doi: 10.1119/1.1374249.
- Duchastel, P. C. (1981). Retention of prose following testing with different types of test. *Contemporary Educational Psychology*, 6, 217–226.
- Epstein, M. L., Lazarus, A., Calvano, T., Matthews, K., Hendel, R., Epstein, B. B., & Brosvic, G. M. (2002). Immediate feedback assessment technique promotes learning and corrects inaccurate first responses. *Psychological Record*, 52(2), 187–201.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring. *American Psychologist*, 34(10), 906–911.
- Glover, J. A. (1989). The “testing” phenomenon: Not gone but nearly forgotten. *Journal of Educational Psychology*, 81, 392–399.

- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66, 64–74.
- Haladyna, T. M. (2004). *Developing and validating multiple-choice test items*. Mahwah, NJ: Lawrence Erlbaum.
- Halpern, D. F., & Hakel, M. D. (2003). Applying the science of learning at the university and beyond: Teaching for long-term retention and transfer. *Change*, 35(4), 36–41.
- Henderson, C., & Dancy, M. (2009). The impact of physics education research on the teaching of introductory quantitative physics in the United States. *Physical Review: Special Topics, Physics Education Research*, 5, 020107-1–020107-9.
- Heron, P., Shaffer, P. S., & McDermott, L. C. (n.d.). *Identifying and addressing student conceptual difficulties: An example from introductory physics*. Washington, DC: National Academy Press. Retrieved from http://www7.nationalacademies.org/bose/Heron_CommissionedPaper.pdf
- Kang, S. H., McDermott, K. B., & Roediger, H. L. (2007). Test format and corrective feedback modulate the effect of testing on memory retention. *European Journal of Cognitive Psychology*, 19, 528–558.
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77, 1121–1134. doi: 10.1037/0022-3514.77.6.1121
- Lasry, N. (2008). Clickers or flashcards: Is there really a difference? *Physics Teacher*, 46, 242–244.
- Lasry, N., Charles, E., Whittaker, C., & Lautman, M. (2009). When talking is better than staying quiet. *AIP Conference Proceedings*, 1179, 181–184.
- Martinez, M. E. (1999). Cognition and the question of test item format. *Educational Psychologist*, 34(4), 207–218.
- Mazur, E. (1997). *Peer instruction: A user's manual*. Upper Saddle River, NJ: Prentice Hall.
- Mazur, E. (2009). Farewell, lecture? *Science*, 323(5910), 50–51.
- McDaniel, M. A., Derbish, J. L., & Morrisette, N. (2007). Testing the testing effect in the classroom. *European Journal of Cognitive Psychology*, 19, 494–513.
- Miller, M. (2011). What college teachers should know about memory: A perspective from cognitive psychology. *College Teaching*, 59(3), 117–122.
- Palinscar, A., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, 1, 117–175.
- Piaget, J. (1985). *The equilibration of cognitive structures: The central problem of intellectual development*. Chicago, IL: University of Chicago Press.
- Piaget, J., Green, D. R., Ford, M. P., & Flamer, G. B. (1971). The theory of stages in cognitive development. In D. R. Green, M. P. Ford & G. B. Flamer (Eds.), *Measurement and Piaget* (pp. 1–11). New York, NY: McGraw-Hill.
- Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66(2), 211–227.
- Rieber, R. W. (Ed.). (1998). *The collected works of L. S. Vygotsky: (Vol. 5. Child Psychology)*. New York, NY: Plenum.
- Schoenfeld, A. H. (1987). What's all the fuss about metacognition. In A. H. Schoenfeld (Ed.), *Cognitive science and mathematics education* (pp. 189–215). Hillsdale, NJ: Lawrence Erlbaum.
- Schoenfeld, A. H. (2010). *How we think: A theory of goal-oriented decision making and its educational applications. Studies in mathematical thinking and learning series*. New York, NY: Routledge.

- Smith, M. K., Wood, W. B., Adams, W. K., Wieman, C., Knight, J. K., Guild, N., & Su, T. T. (2009). Why peer discussion improves student performance on in-class concept questions. *Science*, 323(5910), 122–124. doi: 10.1126/science.1165919
- TERC. (n.d.). *Wind on Oahu*. McDougal Littell. Retrieved from http://www.classzone.com/books/earth_science/terc/content/investigations/es1806/es1806page01.cfm
- Watkins, J. E. (2010, May). Examining issues of underrepresented minority students in introductory physics Cambridge, MA: Harvard University.
- Wiggins, G. P., & McTighe, J. (2005). *Understanding by design* Alexandria, VA: ASCD.
- Zull, J. (2006). Key aspects of how the brain learns. *New Directions for Adult and Continuing Education*, 110, 3–9.