2021 ASEE ANNUAL CONFERENCE
Virtual Meeting | July 26–29, 2021 | Pacific Daylight Time
Paper ID #32712

Student and Teacher Perceptions of a Classroom Response System: Demographic Comparisons in a First Semester Calculus Course

Dr. Patricia A. Ralston, University of Louisville

Dr. Patricia A. S. Ralston is Professor and Chair of the Department of Engineering Fundamentals at the University of Louisville. She received her B.S., MEng, and PhD degrees in chemical engineering from the University of Louisville. Dr. Ralston teaches undergraduate engineering mathematics and is currently involved in educational research on the effective use of technology in engineering education, the incorporation of critical thinking in undergraduate engineering education, and retention of engineering students. She leads a research group whose goal is to foster active interdisciplinary research which investigates learning and motivation and whose findings will inform the development of evidence-based interventions to promote retention and student success in engineering. Her fields of technical expertise include process modeling, simulation, and process control.

Dr. Campbell R. Bego, University of Louisville

An Assistant Professor with research in engineering education, Campbell R. Bego, PhD, PE, is interested in improving STEM student learning and gaining understanding of STEM-specific learning mechanisms through controlled implementations of evidence-based practices in the classroom. Dr. Bego has an undergraduate Mechanical Engineering degree from Columbia University, a Professional Engineering license in the state of NY, and a doctorate in Cognitive Science.

Student and Teacher Perceptions of a Classroom Response System in a First-Semester Calculus Course for Engineers

Abstract

With the goal to improve student engagement and learning, instructors implemented formative assessment in three classes of Calculus I for engineering students in fall 2019. Instructors used the Classroom Response System (CRS) Learning Catalytics to assign formative assessment questions during most lectures each week. At the end of the semester, students were given the classroom response system perceptions questionnaire (CRiSP) which assessed their perceptions of the usability of the CRS, and whether it was helpful in their engagement and learning. Survey results showed that all students, including minority and low-SES populations, perceived it to be very easy to use, and their impressions of its effect on their learning and engagement were slightly above average. Based on these initial results and further research into the effectiveness of formative feedback, we are looking to continue to fully incorporate formative feedback into the Engineering Analysis sequence of courses (Calculus I, II, and III for engineering students) using Learning Catalytics. We are also improving our implementation method above and beyond this first iteration.

Introduction

Despite a continued focus on course improvement initiated with the calculus reform movement 30 years ago [1], first year calculus courses continue to be challenging for STEM majors. Our engineering college continues to work diligently to improve retention of freshmen students, paying particular attention to at-risk students, and we have found that math performance is closely related to retention [2]. The calculus courses for engineering students are taught within our engineering college, so we can more easily implement evidence-based practices with potential to improve retention, of which there are several. Recently, we have focused on implementing formative assessment, which has been shown to improve retention of at-risk students and under-represented minorities [3], [4]. This paper presents and reflects on the implementation of formative assessment in our Calculus I course for engineering students.

Theoretical Framework: Formative assessment

Formative assessment is an instructional technique in which teachers quickly assess students during or outside class for minimal reward/penalty, with emphasis on improving learning. Formative assessment helps both students and teachers to identify knowledge gaps and misconceptions, and address them during the initial learning process when there are still many opportunities for correction [5]. Formative assessment allows instructors to create "moments of contingency" in the classroom when they can modify their instruction based on student understanding at the time. Teachers pose questions and then carefully observe student responses in order to determine whether feedback or additional instruction is needed to help students modify or correct their thinking [6].

The intention of formative assessment is to glimpse current levels of understanding during learning. This differs from the intention of the more commonly-used summative assessment, which is to determine whether students have successfully achieved understanding at the end of the learning period. Nicol and MacFarlane-Dick [6] posit that: "Practice in a classroom is *formative* to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited."

Formative assessment can be done in many ways, but most methods incorporate instantaneous feedback, for which the learning benefits are well established in the literature [7]. Feedback is known to help students by revealing what they know and don't know, and allowing them time to correct any misconceptions that surface. In their meta-analysis of different types of feedback in the classroom, Hattie and Timperley [7] found larger benefits when students received information about the task and about how to perform it more effectively (formative feedback), and lesser benefits for interventions when only praise, or rewards, or punishment were given. The authors further noted that the feedback should be well-timed in the instruction cycle. More recently, researchers Wisniewski, Zierer, and Hattie [8] conducted a meta-analysis of various feedback interventions, and concluded that high-information feedback has a large effect. They suggested that students benefit greatly from feedback when it helps students identify their mistakes, understand why they made the mistakes, and then how to avoid the mistakes in the future, again all characteristics of formative feedback.

Formative assessment also requires students to be active during learning instead of passive. A large meta-analysis of undergraduate STEM courses showed active learning improved exam performance for STEM courses of all sizes [9]. Active learning techniques are known to increase student engagement (e.g., [10]), and are thus associated with positive academic outcomes, including persistence and achievement in school [11]. Active learning has also been reported to confer benefits to STEM students from disadvantaged backgrounds [12], and more recently for underrepresented minorities [13]. Theobald et al. [13] found that active learning reduced achievement gaps in examination scores by 33% and narrowed gaps in passing rates by 45%. Formative assessment engages students in active learning by requiring students to work through problems to answer the assessment questions.

In addition, formative assessment is a good way to provide "appropriate support for learning," which is a key support for at-risk learners [3]. Elliot and Gillen [3], in their thorough study of at-risk students using photo-elicitation to capture students' early impressions of college, found that at-risk students respond to the passion and care of a teacher who provides meaningful time-tables as well as activities that encourage class attendance and instill determination. Frequently-applied formative assessment activities do just that.

Formative assessment in mathematics courses

The benefits of formative assessment—helping students and teachers identify current misconceptions, providing feedback, engaging students in active learning, and increasing support for at-risk students—can be greatly beneficial to entry-level mathematics courses in engineering programs. In fact, some research has already shown that formative assessment benefits students in such calculus courses [14], [15]. Bode and colleagues [14] reported improved student perceptions of learning from the application of formative feedback in a first calculus course. Dibbs [15] observed increased long-term cognitive engagement (mental effort and thinking strategies) as well as behavioral engagement (actions and behaviors people take during learning) due to formative assessment. Dibbs' manuscript was interesting; she studied 8 students who were required to repeat calculus at the beginning of their STEM careers, and who experienced formative feedback the second time. Through interviews, Dibbs found that participants identified three primary reasons for their success the second time around: (1) processing their initial failure, (2) having a better instructor, and (3) participating in formative assessments. In her analysis, she found that formative assessment is what caused the cognitive and behavioral engagement benefits. These students later succeeded in STEM, with 6 eventually receiving a math major or minor.

In summary, effective use of formative assessment has the potential to high-information feedback that helps students remove misconceptions in a low-stress environment, prior to a high-stakes summative test. Early studies in calculus encourage us in the use of formative assessment in this domain.

Technology

Although formative assessment has been used for many years in many different courses with clicker technology, recent advancements in technology and device accessibility have made it easier to implement formative assessment in the classroom. Technologies now include phone and computer-based Classroom Response Systems (CRSs) such as TopHat™, Kahoot, and Mentimeter. Such technologies are promoted as opportunities for increasing engagement and offering instant, formative assessment [16], [17]. However, research on these technologies is limited. In one known study, McAlpin and coauthors [18] compared outcomes for students in one section who solved problems facilitated by TopHat™ to those in another section who solved problems using paper. The question-driven method facilitated by TopHat™ was associated with statistically significantly improved performance measures resulting in a 4.98% increase in final grades compared to the control section.

Recently, the Classroom Response System Perceptions (CRiSP) questionnaire (Richardson et al., 2015) was developed to assess student perceptions of a CRS. The survey consists of 26 items, and breaks down into three primary factors: usability, perceptions of engagement, and perceptions of learning. Use of this survey is limited in the literature, but promises to provide helpful feedback to instructors as they work to add useful formative assessment in their courses. It is important to know how new technologies are perceived by students. Students are more likely to be engaged and benefit from the technology if it is easier to use, and if they find value for learning and engagement through their participation. It is also critical that instructors determine whether the technology is a barrier to engagement or learning for at-risk students. For example, it is possible that less accessible technologies that are more difficult to use would be a bigger burden for low-SES students. The CRiSP survey allows instructors to measure student perceptions of their chosen CRS technology.

Current study

Instructors implemented formative assessment in three (of five) classes of Calculus I for engineering students in fall 2019 to improve student engagement and learning. The implementation of formative assessment was a deliberate modification of a primarily lecture-based Calculus I course, and the lead instructor of the course had 38 years of experience with primarily lecture-style instruction. The CRS was intentionally used during most lectures each week. At the end of the semester, students were given the CRiSP questionnaire which assessed their perceptions of the usability of the CRS, and whether it was helpful in their engagement and learning.

Our research questions included:

RQ1: Was the CRS easy for students to use?

RQ2: Did students perceive the CRS to be effective for their engagement and learning?

RQ3: Were there differences in perceptions for at-risk student groups? I.e., did any at-risk group perceive more/less engagement/learning than the majority?

Methods

This study was approved by our university's Institutional Review Board.

Participants

Participants included three classes of students who were enrolled in Calculus I for engineering students during fall 2019 at the University of Louisville's Speed School of Engineering and who took the CRiSP survey (*N* = 139). One student was removed from the dataset for only responding to 2 of the 26 survey questions. The sample was 28.8% female, 27.3% low SES, and 84.2% white.

Procedures

Calculus I met at least 4 times a week, for 50 minutes on MWF and 75 minutes on TTh. Every Tuesday was an exam day. Only students making below 70% on the Tuesday exam were required to attend on Thursdays. Throughout the semester, on every MWF class, formative assessment was used for a maximum of 10 minutes at the beginning of class, followed by lecture. Instructors prepared short and focused questions about important content and potential misconceptions. After most students submitted their responses, instructors released feedback and then worked each problem, highlighting elements that were confusing and making sure to relate the problem to appropriate concepts from the previous lectures. If many students were having trouble, instructors tailored their instruction to review the related concepts prior to moving forward with new material.

At the end of the semester, instructors gave the CRiSP survey to students during class. Demographic data was supplied from a longitudinal research database, and matched to survey data with coded Research IDs.

Materials

CRS. The CRS used in this study was Learning Catalytics, associated with Pearson's ® MyLabsPlus online portal. To use Learning Catalytics, the instructor starts session through a web browser, and students join on their own computers by clicking on a "join session" tab within the MyLabsPlus portal on the web. The instructor delivers the problem to the students by clicking on a "deliver" tab. Once this is done, the table of responses appears. Students select the appropriate answer, in the case of a multiple choice problem, they click on the response. As they begin entering answers, the instructor sees the response table update. The instructor can see how many students answer correctly, as well as types of incorrect answers entered. Figure 1 illustrates how a problem would appear to the instructor and to the students.

Survey. The survey included the 26 questions from the CRiSP survey, as well as 2 additional course-specific questions, and only 25 questions were used for this study. Survey responses were collected on a five-point Likert scale (from 1 - Strongly Disagree to 5 - Strongly Agree). Responses to the questions on each subscale were averaged to calculate a score for three subfactors:

- Usability (4 items, e.g., "For me, it was easy to use the Learning Catalytics voting system"; $\alpha = .806$);
- Impact on student engagement (9 items, e.g., "Learning Catalytics increased the frequency of my direct participation in the course"; $\alpha = .892$); and
- Impact on student learning (12 items, e.g., "Learning Catalytics allowed me to better understand key concepts"; $\alpha = .923$).

Low SES Status. Low SES status was assumed based on a binary (*Y/N*) Pell Eligibility variable from the longitudinal database, as supplied by the Institutional Research department. The Federal Pell Grant is awarded to undergraduate students who display exceptional financial need.

Gender. Gender was considered a binary variable (*M/F*) for this research, as recorded by Institutional Research and stored in the longitudinal database.

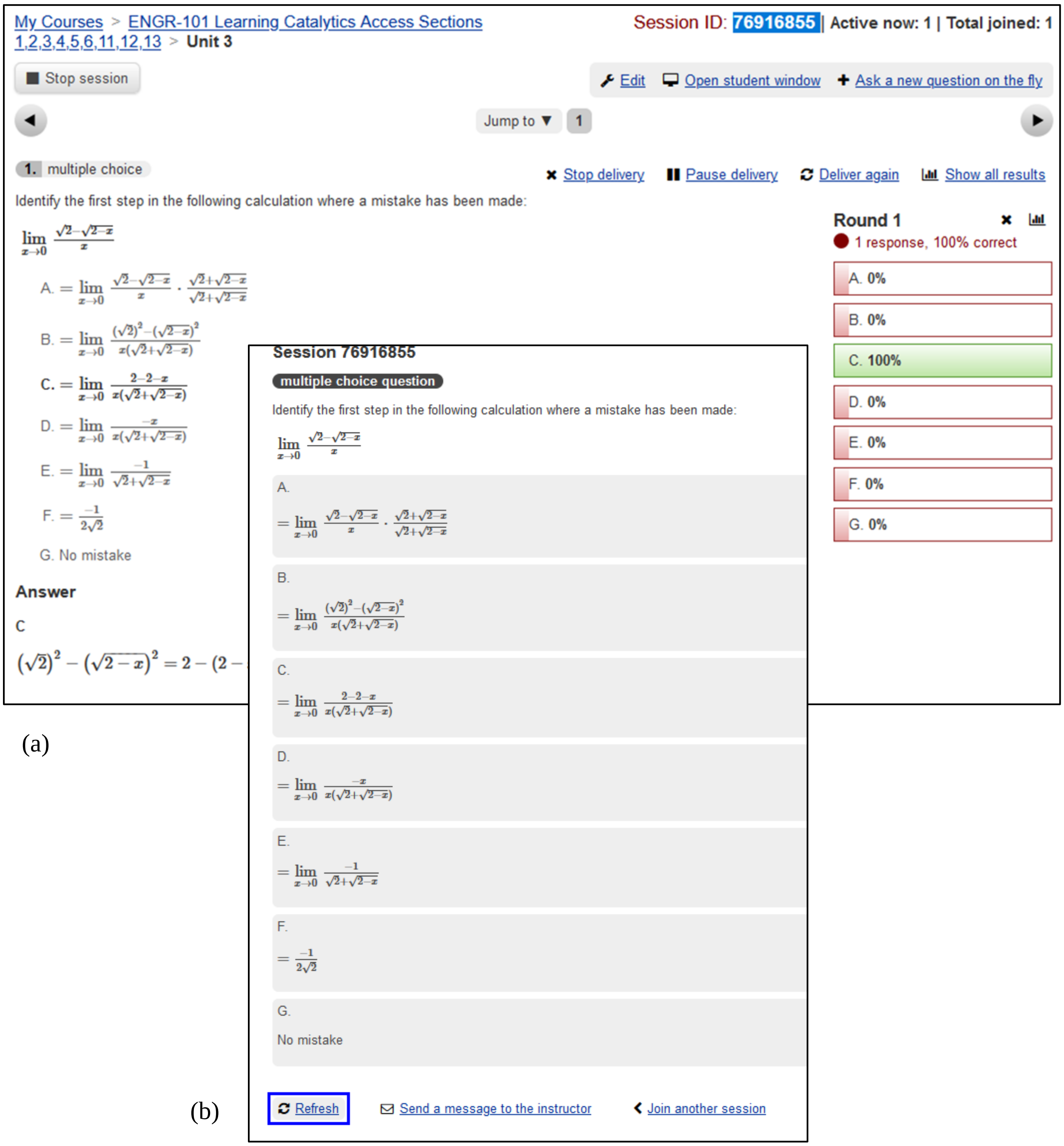

Figure 1*:* A representative question used for formative assessment at the beginning of class in Calculus I for engineering students as seen by (a) the instructor, and (b) the student.

Data analysis

We first used descriptive statistics to evaluate learner perceptions of usability, engagement, and learning using responses to the CRiSP questionnaire. We then compared perceptions across genders and socio-economic status using statistical analyses of variance (ANOVAs). Although it is possible to test mean differences with *t*-tests, ANOVAs are more robust to normality violations such as kurtosis and skew.

Results

Descriptive statistics for usability, engagement, and learning

Overall, students found the CRS to be easy to use ($M = 4.13$, $SD = .65$). The distribution of responses was negatively skewed, with most students reporting high usability, and only 6 students reporting usability below neutral (3). Students perceived the CRS to positively impact both their engagement ($M = 3.32$, $SD = .75$) and learning ($M = 3.41$, $SD = .75$). Descriptive statistics are illustrated in Figure 2.

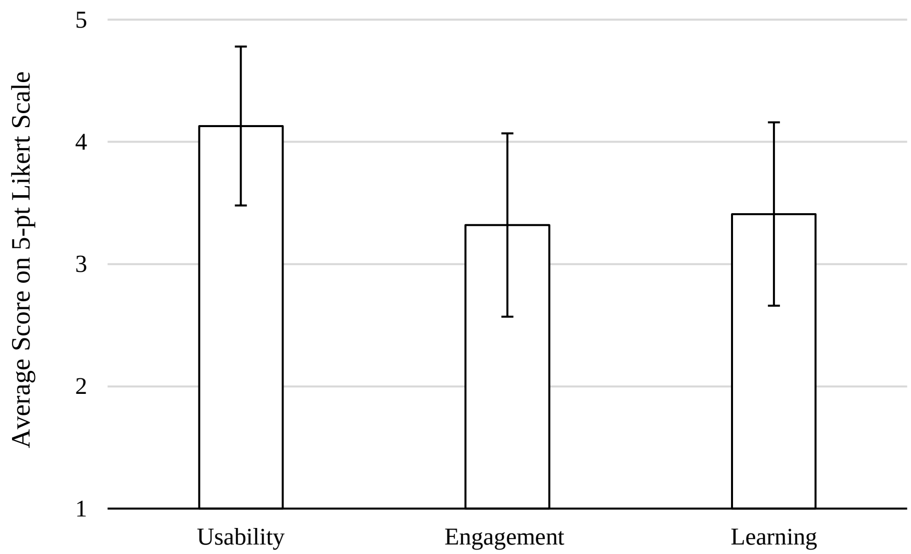

Figure 2: Descriptive statistics for CRiSP survey subscales for all participants ($N = 139$). Error bars = ± SD.

Considering at-risk students

We ran two repeated measures ANOVAs with CRiSP subscale (*usability*, *engagement*, and *learning*) as a within-subjects factor. The first assessed low SES (*Y/N*) as a between-subjects factor, and the second assessed gender (*M/F*). We could not run both of these factors in the same analysis because there were only 9 females in the low-SES category, making any three-way interactions invalid. In both ANOVAs, the assumption for sphericity was violated in the within-subjects factor, so conservative Greenhouse Geisser values are reported here. Results were similar using the less conservative Huynh-Feldt values.

Results indicated that there was not a significant difference between low SES student perceptions ($M = 3.53$, $SE = .10$) and not-low SES student perceptions ($M = 3.66$, $SE = .06$), $F(1, 137) = 1.21$, $p = .273$. There was a significant main effect of CRiSP subscale, $F(1.48, 137) = 97.82$, $p < .001$,

$\eta_p^2 = .417$. Usability (M = 4.12, SE = .06, CI 95% [4.00, 4.25]) was rated significantly higher than engagement (M = 3.28, SE = .07, CI 95% [3.14, 3.42]) and learning (M = 3.37, SE = .07, CI 95% [3.23, 3.51]), which were not significantly different from each other. The interaction between low SES and CRiSP subscale was not significant (p = .393).

There was not a significant difference between genders, although the p-value was low, $F(1, 137) = 3.58$, $p = .061$, $\eta_p^2 = .025$. Females' perceptions of the CRS (M = 3.77, SE = .09) trended higher than males' (M = 3.56, SE = .06). Again, there was a significant main effect of CRiSP subscale, $F(1.48, 137) = 91.91$, $p < .001$, $\eta_p^2 = .40$, with usability (M = 4.17, SE = .06, CI 95% [4.05, 4.29]) rated significantly higher than engagement (M = 3.37, SE = .07, CI 95% [3.23, 3.51]) and learning (M = 3.46, SE = .07, CI 95% [3.32, 3.59]). The interaction between gender and CRiSP subscale was not significant (p = .921).

Discussion

Students perceived Learning Catalytics to be easy to use, and to have a positive influence on their engagement and learning. The average score of 4.13 out of 5 on the useability scale indicates that students found the CRS very easy to use, which is key to determining whether this technology is appropriate for this classroom. Importantly, students of all demographics felt the same about usability; females and low SES students reported equally high usability as males and not-low SES students. That our minority students found this as easy to use as all other students was a great result, letting us know we are not creating an added frustration with the technology for our at-risk students.

Perceptions of engagement and learning were just above the mean (3.32 and 3.41). It is possible that the improvements in engagement and learning due to formative assessment were actually only moderate. The intention was to engage students and help them identify potential misconceptions ahead of exam time. But we, the instructors, did not spend much time emphasizing the importance of the activity during class, which is key communication with students necessary for buy-in. We also did not count participation as part of the grade. While we occasionally discussed with students our reason for using Learning Catalytics, we perhaps did not reinforce this frequently enough in class to actually get them engaged and participating.

Another possibility is that the amount of formative assessment was not enough to dramatically improve engagement and learning. This was an initial implementation of formative assessment in this course, and instructors were new to both formative feedback and the Learning Catalytics platform. Although there is advice on best practices for implementing formative assessment using a CRS [19], much depends on the course, the students, and the teachers. Well-crafted problems take time to develop, especially if the goal is to elicit common student misconceptions. While we feel we did a good job crafting many questions tailored to our course, it is entirely probable that the use of one question simply wasn't *enough* formative feedback to make a difference to the students, or enough of a change from the traditional class to be perceived as engaging.

Related to this being an initial implementation, another possibility is that we did not allow enough time in class for the majority of students to respond to the questions. A challenge for new implementors is how much time to allow for answers. If you wait too long, you are wasting good students' time, and if you go too fast, students just don't bother, or type in anything and wait for you to solve the problem for them. Most always, we did the problem at the start of class, frequently having it open and available a few minutes before the start of class. Early arrivers were often finished before the start of class. Since we had not completely re-designed the course to allow too much time for this activity, we may have frustrated students who arrived just on time, or a little late to class, depriving them of full participation in the activity which would affect both their perception of engagement and learning from Learning Catalytics.

There are ways to make the activity more engaging by using a “team round” or “deliver a second time” option. Delivering a question a second time is very effective especially when many students get the answer wrong the first time. It is very useful to ask the question again and tell students to talk to one or more classmates. However, both of these options are most effective in an active learning classroom where students can easily interact with others. We were confined to a traditional lecture auditorium and could not use the “team round” option at all. While we tried the “deliver a second time” option, it had limited usefulness because some students were isolated, and others could only consult with a student to their left or right.

As we view these results and decide on future improvements, we must acknowledge that perceptions of learning are notoriously hard to interpret. There is an entire body of research on “judgement of learning” that studies assessments students make about their learning. Our results are from the students’ perspective, rather than true measures of learning or engagement. It is common for students to think something is helpful if it is easy, and think it is unhelpful for learning if they find it difficult, even if it is good for learning. Recently, Carpenter et al. [20] discuss what they call “illusions of learning” as follows:

> Students’ judgments of their own learning are often misled by intuitive yet false ideas about how people learn. In educational settings, learning experiences that minimize effort and increase the appearance of fluency, engagement, and enthusiasm often inflate students’ estimates of their own learning, but do not always enhance their actual learning. [20]

A related concept is that of “desirable difficulty” [21], which states that creating deliberate challenges can enhance learning. It is possible that students who couldn’t easily answer the questions didn’t like these problems because they posed a challenge they ordinarily didn’t have during class, even though they knew it didn’t count against their grade. The difficulty may make them perceive that formative assessment *didn’t* help with learning when it actually did. Yet another possibility is that some students dislike any activity that requires them to do something besides be passive in class. They may perceive it as taking time away from seeing the teachers solve problems, which they may have grown to prefer.

Despite potential initial limitations in the implementation, student opinion engagement and learning with formative assessment using Learning Catalytics was not negative. Overall, our expectation is that students will recognize the benefits of formative assessment as we integrate it and use it more frequently in future courses. But for now, even moderate acceptance is a success.

Teacher perspectives

It is also important to characterize teacher perceptions as well, to understand the other perspective on the benefits of formative assessment. The first-author and lead instructor of the course has reflected at length on the use of formative assessment in her classroom. Her primary thoughts are as follows:

> As an instructor with over 35 years of teaching experience, this was really eye-opening to me. Being able to view the answers as they came in on the teacher dashboard revealed what students were doing in “real-time”. Seeing how long it took most students to process information and work a problem was often longer than I thought it should be. Things that I was sure were “crystal clear” were often not, even among many high performing students. This demonstrated to me that removing misconceptions and working more interactively with students was necessary for most students. “Telling and demonstrating” are not enough, as I had previously thought. All classes that I teach have had a day set aside for questions on the unit material prior to the unit test. However, students benefit from interaction as they initially engage with the material. And, they didn’t always ask questions despite my asking them every day “are there any questions”. Seeing

> misconceptions exposed in real time, because of answers selected in several problems we specifically designed to bring these out, gave us an opportunity to reinforce the correct concepts. We found that extremely valuable. We always discussed the problem and tried to elicit questions especially when many students selected a wrong answer.
>
> A frustration of mine has been the appeal for students of peer-assisted learning sessions, which I often felt were sessions designed just to "figure out how to bypass learning and get a good grade on the next test." However, I recognize the need for more direct engagement with the teacher as the material is covered. I think these sessions offer an information opportunity for questions, and through effective use of formative assessment, we as instructors can replicate this.
>
> I especially am encouraged to re-design my courses to allow more time for this type of activity, as it brings out questions students often won't ask, questions that arise while first engaging with difficult material. Usually, these are very good questions that guide all students.

If other engineering mathematics instructors hear these opinions, more may be likely to attempt formative assessment in their classroom. This has the potential to benefit many first-year engineering students, especially those at-risk.

Limitations & Future Work

In retrospect, this study was limited due to the somewhat restricted implementation of the formative feedback, as explained above. As we continue to implement formative feedback with Learning Catalytics, we will emphasize to the students its purpose and ultimate value to them by assigning participation credit. We also think we should allow more time for students to solve the problems initially, and perhaps ask the questions a second time after talking to others. We particularly look forward to our lecture hall being renovated into a spaced where students can swivel side to side, to improve student-to-student interactions and working together. Since this survey was given with the instructors' first exposure to classroom response systems, it was not an optimal implementation. More integrated and frequent use will be key to both students' appreciation of the formative feedback and the efficacy of it for both learning and engagement.

In fact, we have already begun incorporating formative feedback more regularly throughout our math course sequence. We have changed our implementation by allowing for more time for more problems delivered via Learning Catalytics. We would like to administer the CRiSP survey again, and we hypothesize that students in our current courses will perceive higher engagement and learning than those who experienced this first implementation.

Conclusions

This paper presented results of implementing formative feedback in a first course in Calculus I for engineers using Learning Catalytics. Survey results showed the students perceived it to be easy to use, and their impressions of its effect on their learning and engagement were slightly above average. Based on these initial results and further research into the effectiveness of formative feedback, we are looking to continue to fully incorporate formative feedback into the Engineering Analysis sequence of courses (Calculus I, II, and III for engineering students) using Learning Catalytics. Some thoughtful re-design to give more time for this type of activity is expected to improve both students' perceptions of its effectiveness on their learning and engagement, and students' actual learning and engagement.

References

[1] L. A. Steen, Ed., *Calculus for a new century: a pump, not a filter. Papers from a national colloquium (MAA Notes No. 8, October 28-29, 1987)*. Washington, DC: Mathematical Association of America, 1988.

[2] C. R. Bego, J. L. Hieb, and P. A. S. Ralston, "Barriers and bottlenecks in engineering mathematics: How performance throughout a math sequence affects retention and persistence to graduation," in *October 2019 IEEE Frontiers in Education Conference (FIE)*, 2019.

[3] D. L. Elliot and A. Gillen, "Images and stories: Through the eyes of at-risk college learners," *Int. J. Qual. Stud. Educ.*, vol. 26, no. 7, pp. 912–931, 2013, doi: 10.1080/09518398.2012.693217.

[4] R. Stiggins and J. Chappuis, "Using student-involved classroom assessment to close achievement gaps," *Theory Pract.*, vol. 44, no. 1, pp. 11–18, 2005, doi: 10.1207/s15430421tip4401_3.

[5] D. Nicol and D. MacFarlane-Dick, "Formative assessment and selfregulated learning: A model and seven principles of good feedback practice," *Stud. High. Educ.*, vol. 31, no. 2, pp. 199–218, 2006, doi: 10.1080/03075070600572090.

[6] P. Black and D. Wiliam, "Developing the theory of formative assessment," *Educ. Assessment, Eval. Account.*, vol. 21, no. 1, pp. 5–31, 2009, doi: 10.1007/s11092-008-9068-5.

[7] J. Hattie and H. Timperley, "The power of feedback," *Rev. Educ. Res.*, vol. 77, no. 1, pp. 81–112, 2007, doi: 10.3102/003465430298487.

[8] B. Wisniewski, K. Zierer, and J. Hattie, "The Power of Feedback Revisited: A Meta-Analysis of Educational Feedback Research," *Front. Psychol.*, vol. 10, p. 3087, 2020, doi: 10.3389/fpsyg.2019.03087.

[9] S. Freeman *et al.*, "Active learning increases student performance in science, engineering, and mathematics," in *Proceedings of the National Academy of Sciences*, 2014, pp. 8410–8415, doi: 10.1073/pnas.1319030111.

[10] M. T. H. Chi and R. Wylie, "The ICAP framework: Linking cognitive engagement to active learning outcomes," *Educ. Psychol.*, vol. 49, no. 4, pp. 219–243, 2014, doi: 10.1080/00461520.2014.965823.

[11] J. A. Fredricks, P. C. Blumenfeld, and A. H. Paris, "School engagement: Potential of the concept, state of the evidence," *Rev. Educ. Res.*, vol. 74, no. 1, pp. 59–109, 2004, doi: 10.3102/00346543074001059.

[12] D. C. Haak, J. HilleRisLambers, E. Pitre, and S. Freeman, "Increased structure and active learning reduce the achievement gap in introductory biology," *Science (80-.).*, vol. 332, pp. 1213–1216, 2011, doi: 10.1126/science.1204820.

[13] E. J. Theobald *et al.*, "Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 117, no. 12, pp. 6476–6483, 2020, doi: 10.1073/pnas.1916903117.

[14] M. Bode, D. Drane, Y. B. D. Kolikant, and M. Schuller, "A clicker approach to teaching calculus," *Not. Am. Math. Soc.*, vol. 56, no. 2, pp. 253–256, 2009.

[15] R. Dibbs, "Forged in failure: engagement patterns for successful students repeating calculus," *Educ. Stud. Math.*, vol. 101, pp. 35–50, 2019, doi: 10.1007/s10649-019-9877-0.

[16] C. Fies and J. Marshall, "Classroom response systems: A review of the literature," *J. Sci. Educ. Technol.*, vol. 15, no. 1, 2006, doi: 10.1007/s10956-006-0360-1.

[17] R. H. Kay and A. LeSage, "Examining the benefits and challenges of using audience response systems: A review of the literature," *Comput. Educ.*, vol. 53, pp. 819–827, 2009, doi: 10.1016/j.compedu.2009.05.001.

[18] E. McAlpin, D. Shilane, and S. Kalaycioglu, "A pedagogically effective use of an audience response system to increase outcomes in mathematics," *J. Comput. Math. Sci. Teach.*, vol. 37, no. 4, pp. 355–386, 2018.

[19] J. E. Caldwell, “Clickers in the large classroom: Current research and best-practice tips,” *CBE Life Sci. Educ.*, vol. 6, pp. 9–20, 2007, doi: 10.1187/cbe.06-12-0205.

[20] S. K. Carpenter, A. E. Witherby, and S. K. Tauber, “On Students’ (Mis)judgments of Learning and Teaching Effectiveness,” *J. Appl. Res. Mem. Cogn.*, vol. 9, no. 2, pp. 137–151, 2020, doi: 10.1016/j.jarmac.2019.12.009.

[21] E. L. Bjork and R. A. Bjork, “Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning,” in *Psychology and the real world: Essays illustrating fundamental contributions to society*, 2nd Editio., M. A. Gernbacher and J. Pomerantz, Eds. New York, NY: Worth, 2014, pp. 59–68.