ACTIVE LEARNING AND ACADEMIC PERFORMANCE: THE CASE OF REAL-TIME INTERACTIVE STUDENT POLLING

AMIR ARJOMANDI
University of Wollongong
amira@uow.edu.au

ALFREDO R. PALOYO
University of Wollongong
alfredo@paloyo.net

SANDY SUARDI
University of Wollongong
ssuardi@uow.edu.au

ABSTRACT

Active-learning pedagogical practices are encouraged in the education literature to improve academic outcomes. The research reported in this paper explored the effect of real-time interactive polling—a type of active learning—on three measures of academic outcome: final marks, student engagement, and failure probability. To provide credible impact estimates, a cluster randomized controlled trial in an introductory statistics class was conducted. With a little over 500 students, nine out of the 24 tutorial classes were conducted with real-time quizzes; the other 15 tutorial classes served as control classes. The results showed that this active-learning technique did not impact final marks and the probability of failing the subject. It did, however, increase tutorial attendance by about 24%, which corresponded to about 1.6 additional hours compared to the control classes. The findings support the literature on the benefits of classroom response systems that increase student engagement, student interest, and enjoyment from smaller group discussions.

Keywords: Active learning; academic performance; interactive teaching; student engagement

1. INTRODUCTION

Statistics for Business is a common subject first-year undergraduates are required to complete in many universities. In Australia, both lecturers and students of statistics find this subject generates some challenges in its teaching and learning, respectively (Arjomandi et al., 2021). Increasing student engagement has been identified in the literature as necessary for improving students' learning experiences in various subjects, including first-year statistics (Arjomandi et al., 2021; Korobova & Starobin, 2015; Thomas & Heath, 2014). At the same time, the advancement of digital teaching technology in an era when university students are seen as the "digital generation" is one of the pressures forcing instructors to examine their approaches regularly for modern relevance and to accommodate diverse learning needs (Buzzard et al., 2011; Jensen & Owen, 2003; Tait, 2009). Therefore, the use of teaching technologies is now more widespread than ever in educational settings (Mercader & Gairín, 2020) and various technologies are used to improve students' learning and performance. Reported in this paper is exploration of the benefits of *Learning Catalytics*—a proprietary classroom response system that facilitates students' learning outcomes in a first-year business statistics course through active, real-time student polling (Mazur et al., 2013). More specifically, students' learning experiences and outcomes using this interactive learning method, implemented via software to improve student learning in a first-year business statistics course were examined.

As an interactive student-response tool, *Learning Catalytics* facilitates active learning[1] by using students' smartphones, tablets, or laptops to engage them in interactive learning tasks. Initially developed at Harvard University by Mazur et al. (2013), *Learning Catalytics* was acquired by Pearson Education in 2013. *Learning Catalytics* allows the instructor to pose various open-ended and closed questions that support students to develop critical thinking skills. The platform monitors responses with real-time analytics to determine the concepts or ideas the students struggle to grasp. Equipped with this information, instructors can adjust their teaching strategy in real-time and attempt other ways of engaging students during class to assist them to overcome difficulties in understanding and applying certain concepts. While *Learning Catalytics* can manage student interactions by automatically grouping students for discussion, team-based learning[2], and peer-to-peer learning, in this experiment, the application of *Learning Catalytics* is limited to real-time interactive student polling at the start of tutorials.

Much of the published literature has asserted that interactive classrooms can enhance student understanding of concepts, retain engagement, and stimulate interest in the subject (Becker & Watts, 1995; Mazur, 1997; Crouch & Mazur, 2001; Durham et al., 2007; Salemi, 2009; Carter & Emerson, 2012). In recent years, given the pervasive use of digital learning and students' engagement with social media, a growing body of literature has called for a reevaluation of the traditional in-class, lecture-based course model to improve student achievement and the likelihood of success (McLaughlin et al., 2014). According to this thinking, flipped classrooms create more time for activities and interactions, which are critical for improving student learning (Caviglia-Harris, 2016; Lage & Platt, 2000; Lage et al., 2010). In addition, the use of other modes of learning involving classroom response systems with technology has been examined for its effectiveness as a teaching and learning toolkit that delivers positive learning outcomes. Such an approach is consistent with Mishra and Koehler's (2006) conceptual definition of integrating technology into pedagogy via TPACK (i.e., technological pedagogical content knowledge). The TPACK framework for teacher knowledge involves "a complex interaction among three bodies of knowledge: content, pedagogy, and technology. The interaction of these bodies of knowledge, both theoretically and in practice, produces the types of flexible knowledge needed to successfully integrate technology use into teaching" (Mishra & Koehler, 2006, p. 60). Advancements in technology have brought about a paradigm shift from TPACK to iPAC. iPAC is a framework that allows teachers and educators to define the use of mobile technology within their pedagogical framework and develop their unique learning paradigms. This new way of thinking has offered a fresh perspective on learning curricula and pedagogy, focusing on mobile devices and software applications (Burden & Kearney, 2018). Learners can interact with the learning material through mobile devices and utilize various applications for their learning. Burden and Kearney (2018) described the design and development of a mobile learning toolkit for educators to exploit the unique pedagogical capabilities of mobile technologies.

Central to *Learning Catalytics* is the notion of active learning—that is, students can construct their understanding of course material and engage in the learning process as they interact with their classmates. Independent of the disciplines, the benefits of active learning embodied in interactive learning are wide-ranging. Interactive learning, defined as is any learning activity that incorporates social interaction or networking and using interactive technology, has been associated with higher exam scores by an average of 6% in undergraduate science, technology, engineering, and mathematics (STEM) courses and lower failure rate by 36% (Freeman et al., 2014). Active-learning methods through interactive engagements have also been associated with enhanced problem-solving skills in undergraduate physics classes (Hake, 1998). Furthermore, interactive learning has been associated with higher engineering grades (Prince, 2004). In economics, students subjected to active learning obtained 2% and 3% higher marks than those without active learning (Emerson & Taylor, 2004). The cohort engaged in cooperative learning had final exam grades 4–6% higher than the other cohorts without cooperative learning (Yamarik, 2007). Similarly, the

---

[1] Active learning is any learning activity in which the student participates or interacts with the learning process instead of passively taking in the information.

[2] Team-based learning is defined as "an active learning and small group instructional strategy that provides students with opportunities to apply conceptual knowledge through a sequence of activities that includes individual work, teamwork, and immediate feedback" (Parmelee et al., 2012, p. 275).

use of online videos in teaching economics has delivered 4–14% better grades compared to students in the traditional course on a set of common questions that had never been seen in the course previously (Caviglia-Harris, 2016). As a result of the interactive-learning approach, students were able to move up Bloom's taxonomy from comprehension to evaluation (Bloom, 1984) and make noticeable cognitive gains (Durham et al., 2007) and improvements in economic reasoning (Marburger, 2005).

Learning Catalytics is a type of classroom response system (CRS). More generally, CRSs come in various forms, such as audience response systems or clickers (Caldwell, 2007; Miller et al., 2003), voting machines (Reay et al., 2005), wireless keypad response systems (Burnstein & Lederman, 2003), classroom communication systems (Paschal, 2002), electronic response systems (Hall et al., 2002), and online student response systems (OSRSs; Balta & Guvercin, 2016; Muir et al., 2020). The typical CRS involves students using transmitters to send responses, receivers that collect these inputs, and computer software that interprets and aggregates these responses in real-time. More recently, students can access these systems using their own mobile devices with internet access. To access an OSRS, teachers and students log on to a website or app on their own devices (e.g., a computer, laptop, tablet, or smartphone) to interact and respond to questions in real-time via the internet. The instructor chooses to publicly or anonymously collect and display student input signals. The aggregated responses are almost always publicly displayed to inform instructors and learners of the overall distribution of selections in a classroom. Generally, the response items concern a true or false statement or selecting an option (or options) among given choices predetermined by the instructor. In addition, the instructor may poll the class as a formative assessment, with or without any student feedback, and decide the course of further instruction.

An analysis of 26 classroom network studies involving CRS use indicated good agreement regarding the benefits the systems conferred (Roschelle et al., 2004). They found greater student engagement, increased student understanding of complex subject matter, increased student attention, interest and enjoyment, heightened discussion and interactivity, increased student awareness of individual levels of comprehension, and increased teacher insight into student difficulties. Since CRS promotes cooperative learning environments in which students can work in groups before or after solving the poll questions, it promotes intellectually stimulating discussions, problem-solving, and team decision-making skills (Noel et al., 2015). Furthermore, as most systems immediately reveal the poll results, students can obtain instantaneous feedback that reinforces the focus on the content. Such formative assessment enables students to self-regulate their performance, consistent with student-centred learning (Salzer, 2018). Sun et al. (2014) argued that the anonymous response feature facilitates learners to answer questions honestly and engage in deeper discussions, leading to authentic learning, and meaning making. Balta and Guvercin (2016) found that the final grades of students enrolled in a statistics class who chose to engage with OSRS learning materials before their exam were higher than the grades achieved by students who did not engage. In contrast, a previous controlled trial using a sample of engineering students (Dabbour 2016) and a core statistics unit in psychology (Muir et al., 2020) reported mixed results, suggesting that implementing the OSRS increased student engagement but had no effect on student performance.

Be that as it may, these studies' rigour has been criticized, leading to the inability to draw firm conclusions about the technology's effectiveness. Two exceptions are Mestre et al. (1996) and Nicol and Boyle (2003), which compared two types of interactive instructions, one with and one without a CRS. Mestre et al. (1996) discovered that although using the CRS was not essential, displaying the results that summarized the students' responses generated by the software stimulated discussion. Students often addressed the majority selected answer option first. The learning environment was enhanced with visuals and activity—a break away from the usual lecture format—to the extent that students' attendance and attention improved (Dynan, 2014). Nicol and Boyle (2003) indicated that students preferred peer instructional segments (or small group discussions) that occur with the use of CRS over whole-class discussions as the latter became more confusing. In addition, students' responses were frequently counted toward course grades where partial credits were awarded for incorrect selections (Paschal 2002). Although such a reward system motivated participation, it is difficult to verify the benefits of CRS use associated with improved attendance and participation, especially if the course grade depends upon CRS input. For this reason, in this study, grades were not awarded for *Learning Catalytics* participation. Furthermore, Sun

(2014) showed that if in-class polling is not graded, students perceived it as a challenge rather than a threat, leading to effective preparation, review of course content and anxiety reduction. While Balta and Guvercin's (2016) result on final grade improvement is encouraging, a non-randomized, post-test design means their study could not confirm the beneficial effect of using OSRS. In other words, it is uncertain whether the difference in exam scores was due to underlying scholastic aptitude or motivation of the students who chose to engage with the OSRS.

The purpose of this study was to determine whether active learning arising from real-time student polling at the beginning of a tutorial class improves a student's final marks and/or diminishes the probability of failing the subject. In addition, the study investigated whether the use of *Learning Catalytics* in the form of real-time student polling increases student engagement through increasing attendance in tutorial classes.

2. INSTITUTIONAL AND SUBJECT DETAILS

The University of Wollongong (UOW) is located in the state of New South Wales (NSW) in Australia, about 80 km south of Sydney. There are about 27,000 on-shore students, with an additional 7,000 or so off-shore students spread over several affiliates overseas. In a federally funded student survey conducted around the time of the experiment,[3] UOW ranked highest overall and in NSW in the domains of "Learner Engagement" and "Learning Resources" among universities located in a similar geographic catchment area (i.e., in NSW and the Australian Capital Territory, where Canberra—the nation's capital—is located). At UOW, a subject is a course of study taken over one semester for a period of 13 weeks.

Student active learning was tested using a classroom response system in an introductory statistics class. The subject was Statistics for Business, typically taken in the first year for students in the Faculty of Business and Law.[4] It was a compulsory subject for all students of the faculty. The student enrollment was about 500 per semester. There were 13 teaching weeks, with a one-week mid-semester break. There was also a one-week study break at the end of the teaching weeks before a two-week final examination period commenced. Typically, the failure rate was about 40%.

The subject coverage included four weeks of probability theory, four weeks of statistical inference, and five weeks of analysis of variance and regression analysis. Students were also introduced to the use of Microsoft Excel and its various basic data functions. The learning outcomes included explaining and demonstrating basic concepts in probability and statistics, demonstrating statistical techniques used in Business, applying statistical techniques to business decision-making, using and interpreting the output from statistical software packages, and evaluating the role of statistics in empirical research and business practices.

Subject delivery was comprised of three components. First, a two-hour face-to-face lecture was delivered in a large lecture hall every week. Second, smaller "tutorial classes" of about 15 to 25 students met for an hour over the same week. Non-permanent academic staff members usually led the tutorial activities. Third, the subject was supported by weekly one-hour supplementary instruction (SI) in the form of weekly peer-assisted study sessions (PASS), but participation in PASS is voluntary. These were small-group classes led by a peer leader, who was a student who successfully completed the subject in a previous semester and had undergone a training program in preparation for leading the PASS classes.[5] There were 11 tutorial classes and 10 PASS/SI sessions held over the semester.

There were four assessment tasks, with the following corresponding weights: three in-class quizzes (15%), a practical exam (15%), a mid-session exam (20%), and a final exam (50%). The in-class quizzes and the practical exam were conducted during the tutorial classes. To be clear, in-class quizzes are not the

---

[3] Quality Indicators for Learning and Teaching (QILT), 2018 Student Experience Survey National Report: https://www.qilt.edu.au/docs/default-source/default-document-library/2018-ses-national-report.pdf.
[4] The subject was available as an elective subject for other students, and some students enrolled in the Faculty of Business and Law take it late in their degrees. All students enrolled in the Statistics of Business were onshore.
[5] PASS at UOW was evaluated in Paloyo et al. (2016). The results did not show any impact on final marks, although PASS may help first semester students more than others.

review questions assigned for that tutorial class. These review questions form part of the "*Learning Catalytics* questions (poll)" and are formative assessment tasks, while the quizzes are summative assessment tasks. Lecture slides and other class materials are provided in advance via the electronic learning platform (Moodle). The lectures are also recorded and are made available to students immediately via a video platform (Echo360). Students were free to watch the video at any point during the semester, and there was no limit to the number of times they could watch the recordings.

3. EXPERIMENTAL DESIGN

The study included implementation of a cluster randomized controlled trial to estimate the impact of real-time student polling. Ethics approval was granted to run the experiment from the Social Science Human Research Ethics Committee of the University of Wollongong (No. HE16/020). The trial was used to evaluate the impact of *Learning Catalytics* on academic outcomes for an undergraduate Business statistics class. This study compared students' academic performance in a control group characterized by a "traditional" undergraduate course, including chalk-and-talk lectures mixed with traditional tutorial classes. The treatment group was complemented by using *Learning Catalytics* in tutorial classes over 11 weeks. A set of questions relating to the materials covered in the tutorial was prepared in each tutorial class.

Students in the treatment and control groups were required to attempt their tutorial questions before class. Only students in the treatment groups were supplemented by real-time student polling at the beginning of the tutorial class. The polling comprised multiple-choice questions that reviewed the subject material covered in the previous week's lecture. The questions were presented using the *Learning Catalytics* application owned by Pearson Education, which students could access and answer these questions using their electronic devices. As the responses were collated in real time, the tutor was informed of the weaknesses in certain concepts that students encountered, which would be examined in that week's tutorial.

This approach was built on the strength of *Learning Catalytics*, which provided a real-time assessment of students' understanding of the materials before the start of the tutorial. By clearly identifying students' lack of understanding of certain statistical concepts, the tutor could probe and further inquire about these concepts during the tutorial when answering the compulsory questions that related conceptually to the *Learning Catalytics* questions. Often this process involved peer explanation or reconciliation, where students were encouraged to discuss among themselves to arrive at the correct answer collaboratively. Through this interactive learning process, both the instructor and students could engage more profoundly on concepts and highlight possible impediments in understanding and applying the concepts. Some of the concepts examined in the real-time polling overlapped with the assigned tasks given in that tutorial. Accordingly, real-time polling and peer engagement provided instant feedback that might have facilitated learning. Still, it could also help instructors teach better by augmenting their teaching focus and improving learning outcomes (King & Sen, 2013).

The experiment took place in Semester 2 of 2017. The Semester 2 class was chosen because it is typically the largest class in the year relative to the Semester 1 and the Summer 3 (that is Summer Session) classes. Since it was a compulsory subject for all school of Business students, it drew from a variety of majors, including from the two schools within the faculty at that time: the School of Accounting, Economics, and Finance; and the School of Management, Operations, and Marketing. The non-permanent tutors varied every year. In this instance, there were five tutors.

All students were provided with a participant information sheet at the start of the session, indicating the experiment's background, purpose, and confidentiality provisions. Participation in the study was emphasized as entirely voluntary, and students had the option to exit from the study at any point during the semester. All students provided their consent, and none withdrew from the study before completion.

The treatment involved using real-time student polling at the beginning of the tutorial class to review the subject material taught during the previous week's lecture. The polling took the form of about 10 multiple-choice questions that the students could answer using their electronic devices (e.g., laptop, phone, or desktop computer if they were in one of the computer labs). It was conducted using the *Learning Catalytics* application owned by Pearson Education. The 10 multiple-choice questions comprised factual

recall, conceptual understanding, simple application of procedures, and problem-solving. Some of the questions were designed with common incorrect answers as distractor options, while specific answers were provided to point out misconceptions. The polling took 5 to 10 minutes at the start of a 60-minute tutorial.

Students were given *Learning Catalytics* questions at the start of the class. Responses were collected in real-time, and the tutor could view the distribution of correct and incorrect answers almost instantaneously. This information revealed to the tutor which parts of the previous week's lecture he or she should concentrate on during that session's tutorial class on top of the usually assigned tasks. The tutor could identify possible impediments in understanding and applying the concepts that students would further encounter in that tutorial class, plan their tutorial class by allocating more time to questions with complex concepts and elicit greater student participation. Additionally, the tutor could use polling results to prompt student discussion, encouraging them to explain and reconcile concepts with their peers in order to promote cognitive engagement. This process took place when they had addressed the compulsory tutorial questions. Specifically, the tutor would first address the compulsory tutorial question in class before referring to the related *Learning Catalytics* question and its polled answers. If the polled answers were skewed away from the correct answer, the tutor motivated students to identify the mistakes and clarify misconceptions. The *Learning Catalytics* question was used as a conceptual clarification and verification mechanism through students' participation in explaining their conceptual understanding.[6]

There were 509 enrolled students with 20–25 students per tutorial class, resulting in 24 tutorial classes. Nine of the 24 classes were randomly assigned to be part of the treatment group. To ensure randomization, a number was drawn from a uniform distribution per tutorial class. If the number drawn was below 0.5, the tutorial class was assigned to the control group; otherwise, it was assigned to the treatment group. The tutors were given written instructions to ensure consistency in the delivery of the tutorials across the 24 classes. The format of the delivery differed for the subset of nine treatment classes in which the real-time polling using the *Learning Catalytics* took place. The remaining classes served as the control group. Students in the treatment and control groups were required to attempt tutorial questions before class. Students were provided with various opportunities to be actively engaged in the classroom environment for all classes. These activities included applied exercises, real-world case studies, and videos. Classroom discussions were facilitated to increase student engagement. Students were given ten minutes to discuss each question in small groups and decide whether the question needed to be covered on the board by their tutor. This was followed by a review of the pertinent concept for an additional five minutes. The review time in the treatment groups included a discussion of the *Learning Catalytics* question that related to the concept covered in the compulsory question. In addition to the main questions addressed in all tutorials, students were given additional questions, which were a variation of the compulsory questions with which to practice.

Instructions for tutors when delivering discussion-based tutorials. Students in both the treatment and control groups had ten minutes to discuss each question, followed by a five-minute review of the relevant content. Students were required to share their prepared solutions with their group members, and tutors needed to walk around each group, listen, and briefly join the conversation.

If multiple groups were struggling with the same problem, tutors needed to present on the board. There were three to four compulsory questions covered in each tutorial, and tutors could explain the question on the board to all students in the class if necessary. Upon completing the mandatory questions, tutors could cover one or two of the additional questions, which were a variant of the compulsory questions, based on students' decisions in their group discussions if they had extra time. The additional questions were in the tutorial materials but were not the *Learning Catalytics* questions.

---

[6] To ensure uniformity and consistency in the standard of *Learning Catalytics* delivery, tutors were trained by the head tutor and Pearson on the use of the *Learning Catalytics*. There were five tutors, of whom two tutors conducted tutorials with and without the use of *Learning Catalytics*. The tutors, on average, had taught this subject for at least three semesters, thus mitigating confounding factors like teaching skills and experience that could influence teaching delivery and student learning performance. The teaching was limited to two tutors for the nine *Learning Catalytics* classes to limit variation in the skills of implementation of the *Learning Catalytics* amongst a greater number of tutors as a possible impediment to the uniformity of the "treatment" (i.e., the tutorial delivery with the use of *Learning Catalytics).*

In treatment classes, *Learning Catalytics* polling questions were used in the first five to ten minutes to assess students' understanding of the concept before moving to the compulsory tutorial questions. Tutors in treatment classes were not to discuss the answers to the polling questions in the first five minutes.

Below is a summary of steps that were performed by tutors in treatment classes using the *Learning Catalytics* polling questions in the first tutorial:

1. *Go through Learning Catalytics polling questions (5–10 mins)*

   Review content, if necessary, then move to group work.

2. *Chapter group work (10–15 mins)*

   Students share solutions to their work. Walk around, listen, and briefly join the conversation in each group. Shift to presenting on the board if multiple groups struggle with the same problem. The tutor refers to the *Learning Catalytics* question's polling answers and explains how this question relates to the mandatory question, and further reinforces conceptual understanding.

3. *Repeat Step 2 for the other mandatory questions from other chapters. Discuss the answer to the Learning Catalytics questions and relate them to these required questions. (30 mins)*

   For the control tutorial classes, Step 1 is excluded. The control tutorial proceeds in the order following Steps 2 and 3. Each tutorial had 3 to 5 mandatory questions, depending on the complexity of the questions. The tutor needed to adjust the small group discussion time and the review time to ensure that all questions were covered in the tutorial.

4. ESTIMATION STRATEGY

The baseline estimating equation was as follows:

$$(1) \qquad y_{ij} = \alpha + \tau d_{ij} + \varepsilon_{ij}$$

where y_{ij} is an outcome variable of interest for student i in tutorial class j, d_{ij} is a treatment indicator for being in a class with *Learning Catalytics*, and ε_{ij} is an idiosyncratic error term (i.e., specific to student i in tutorial class j). The parameters α and τ were estimated using the ordinary least squares (OLS) method. τ represented the causal impact of real-time student polling on the outcome. In this paper, the impacts on final marks, which ranged from 0 to 100, were examined. The number of tutorial classes attended ranged from 0 to 11.

Equation (1) was augmented with a vector of control variables, **x**, with an associated vector of coefficients, $\boldsymbol{\beta}$:

$$(2) \qquad y_{ij} = \alpha + \tau d_{ij} + \boldsymbol{\beta}' \boldsymbol{x}_{ij} + \varepsilon_{ij}$$

The control variables included the student's age, number of previously completed subjects, tutorial class size, and indicators for being an international student, being a Dean's Scholar,[7] and having previously failed the subject. For inference, heteroskedasticity- and cluster-robust standard errors were reported below.

To improve the balance of covariates, the probability of being in the treatment group (i.e., the propensity score) was predicted as a function of the observable characteristics. The sample was then trimmed to keep observations only within the common support—that is, the range in which the propensity scores for the treatment and control groups overlap.

The probability model for being in the treatment group was the following probit model:

$$\Pr[d_{ij} = 1 | \mathbf{x}] = \Phi(\boldsymbol{\gamma}' \mathbf{x}_{ij}),$$

---

[7] The Dean's Scholars Program was for high-achieving students at the faculty. To be eligible, students needed to satisfy a number of academic requirements. Successful applicants were provided with a book allowance, special mentoring with academic staff members, and extended library borrowing rights, among others.

where the (vectors of) variables were as previously defined, and where Φ is the cumulative distribution function of the standard Normal distribution. The vector of parameters $\boldsymbol{\gamma}$ was estimated using maximum likelihood. These coefficients were used to calculate the predicted probabilities. The probit model was typically used to estimate a regression model with a binary outcome. Of interest was the impact of real-time polling on the failure probability. To estimate this, the right-hand side of Equations (1) and (2) were taken the linear index in a probit model to account for the binary nature of the dependent variable. Next, the average marginal effects of the treatment on the probability of failing were determined. Finally, standard errors were calculated using the delta method.

In Table 1, presented are the difference in means between the treatment and control groups over the following variables: age, number of completed subjects, tutorial class size, and indicators for being an international student, being a Dean's Scholar, and having failed the subject previously. Based on observable characteristics, our randomized treatment assignment balanced all but one observable characteristic, which is the indicator for being an international student. There were slightly more international students in the control group (8.6 percentage points).

Table 1. Balance test of baseline covariates

Baseline Covariate	Control (1)	Treatment (2)	Difference (3)
Age (years)	20.133 (1.864)	19.974 (1.751)	-0.159 (0.333)
International student	0.231	0.145	-0.086 (0.014)
Dean's Scholar	0.041	0.078	0.037 (0.102)
Number of completed subjects	6.177 (4.124)	6.207 (3.051)	0.030 (0.925)
Failed subject previously	0.415	0.373	-0.041 (0.352)
Tutorial class size	21.367 (2.365)	21.705 (2.213)	0.338 (0.104)
Observations	316	193	509

Notes. Standard deviations are enclosed in parentheses in Columns (1) and (2) for continuous variables. Exact *p*-values are in parentheses in Column (3).

The kernel density estimates of the propensity score (predicted probabilities) are presented in Figure 1. There were seven control observations with no matching treatment observation based on the propensity score; there was one treatment observation with no comparable control unit. These observations were deleted from the estimation sample to obtain the common support. Using a Kolmogorov–Smirnov test, there was no evidence suggesting that the two distributions are different.
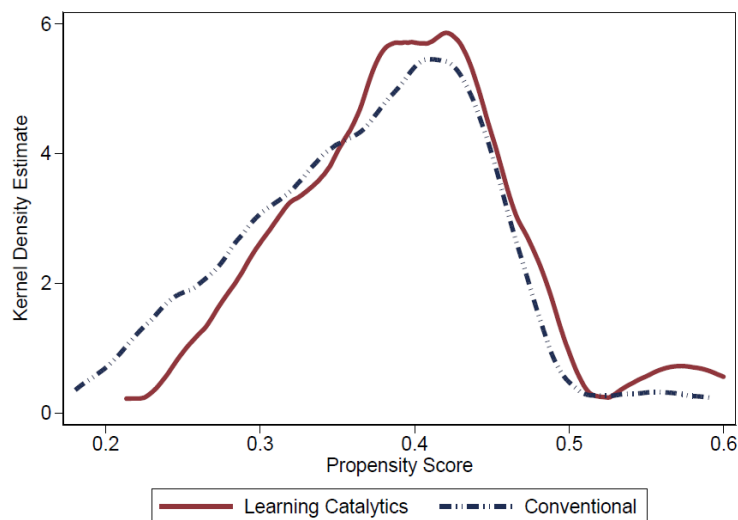
Figure 1. Kernal density estimates of the propensity score by treatment status.

Using observations within the common support, balance was achieved between the treatment and control groups over all covariates (Table 2). The average age was about 20 years old, 19% were international students, 5% were Dean's Scholars, 40% failed the subject previously, and the average tutorial class size was 22 students. On average, students had already completed six subjects previously. The estimates presented below were based on observations within the common support. Including the deleted observations did not have a material impact on the magnitude and significance of the results.

Table 2. Balance test of baseline covariates (common support)

Baseline Covariate	Control (1)	Treatment (2)	Difference (3)
Age (years)	20.120 (1.871)	19.979 (1.754)	-0.141 (0.396)
International student	0.214	0.146	-0.068 (0.051)
Dean's Scholar	0.042	0.073	0.037 (0.162)
Number of completed subjects	6.204 (4.101)	6.208 (3.059)	0.004 (0.989)
Failed subject previously	0.408	0.375	-0.033 (0.465)
Tutorial class size	21.456 (2.302)	21.693 (2.212)	0.236 (0.253)
Observations	309	192	501

Notes. Standard deviations are enclosed in parentheses in Columns (1) and (2) for continuous variables. Exact *p*-values are in parentheses in Column (3).

5. RESULTS

The estimates of Equation (1)—linear regressions with only the treatment indicator and no covariates—appear in Columns (1) to (3) of Table 3. All estimations were made with *Stata/SE 15.1*. The outcome variables were final marks, standardized final marks (i.e., [a student's final mark—the mean final mark] divided by the standard deviation of the final mark), and the number of tutorial classes attended. The table also shows the average marginal effect of the treatment on failure probability based on a probit model in Column (4). Two sets of standard errors were calculated: heteroskedasticity-robust standard errors are in parentheses, and cluster-robust standard errors are in brackets. The cluster was defined as the tutorial class to allow for intra-class correlation.

9

Table 3. Treatment effect estimates (no controls, common support)

	Final Mark (1)	Final Mark (Std.) (2)	Tutorial Attendance (3)	Failure Probability (4)
Learning Catalytics	1.428 (0.536) [0.496]	0.056 (0.536) [0.496]	1.616 (0.000) [0.002]	0.008 (0.846) [0.817]
Mean (Control Group)	54.566	-0.016	6.644	0.330
Observations	501	501	501	501

Notes. Exact *p*-values are in parentheses (robust) and brackets (clustered). The marginal effect based on a probit model for failure probability is reported in Column (4).

Active learning that took the form of real-time student polling at the beginning of a tutorial class did not influence final marks and the probability of failing the subject. It did, however, increase student engagement as measured by the number of tutorial classes attended. With a mean attendance of 6.6 in the control group, the estimated impact of 1.6 corresponded to a 24% increase; as a reminder, the total number of tutorial classes for the semester was 11.

In Table 4, the analogous results with control variables are presented. The addition of covariates did not meaningfully change the estimated results from Table 3. The estimated treatment effect was indistinguishable from zero for final marks and failure probability. The magnitude of the impact on tutorial attendance was slightly lower (1.492 with control variables against 1.616 without control variables), corresponding to an increase of about 22%.

Figure 2 shows the frequency histogram of tutorial attendance by treatment status over the common support. Tutorial attendance stands out because it was the only outcome variable that appeared to be affected by real-time student polling. The graph shows that students randomly assigned to the *Learning Catalytics* tutorial classes attended more than those assigned to the control group. In particular, many more treated students attended 10 or 11 tutorial classes. Conversely, students in the control classes were mainly within the 0 to 5 range. As in the estimation results presented in Table 4, this was interpreted as evidence of increased student engagement in tutorial attendance. Still, there were several ways to operationalize student engagement (class participation, use of the e-learning platform, etc.). This form of student engagement, however, did not result in better marks.

Table 4. Treatment effect estimates (with controls, common support)

	Final Mark (1)	Final Mark (Std.) (2)	Tutorial Attendance (3)	Failure Probability (4)
Learning Catalytics	0.651 (0.619) [0.550]	0.026 (0.619) [0.550]	1.492 (0.000) [0.002]	0.057 (0.145) [0.144]
Mean (Control Group)	54.566	-0.016	6.644	0.330
Observations	501	501	501	196

Notes. Control variables included age, number of previously completed subjects, tutorial class size, and indicators for being an international student, being a Dean's Scholar, and having previously failed the subject. Exact *p*-values are in parentheses (robust) and brackets (clustered). The marginal effect based on a probit model for failure probability is reported in Column (4). Indicators for being a Dean's scholar and having previously failed the subject are perfect predictors in the probit model, so some observations were disregarded.
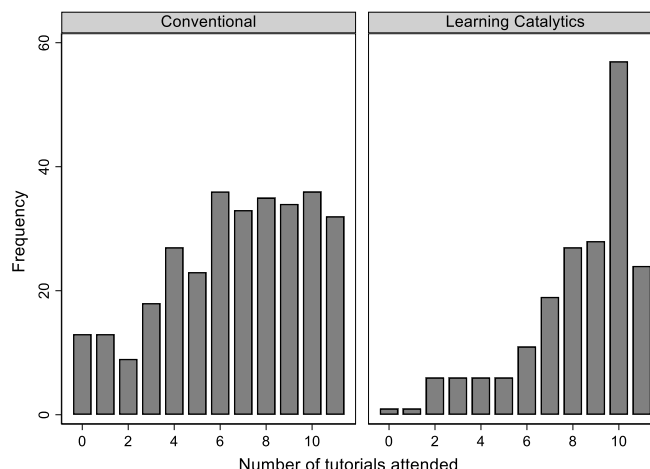
Figure 2. Frequency histogram of tutorial attendance by treatment status (common support)

6. DISCUSSION

In this paper, the impact of active learning was evaluated—in particular, real-time interactive student polling—on the following academic outcomes: final marks, student engagement measured as tutorial attendance, and the probability of failing a subject. A cluster randomized controlled trial with the software *Learning Catalytics* in a first-year undergraduate statistics class at a regional university in Australia was conducted to estimate the treatment effect. The results showed that interactive quizzes at the start of a tutorial class did not impact marks and failure probability. It did, however, increase tutorial attendance by about 24%.Past studies focusing on classroom response systems document increased student participation as their responses are counted toward course grades (Burnstein & Lederman, 2003). Our experiment did not offer a reward for participation; hence the marked improvement in tutorial attendance could reflect the benefits of classroom response systems associated with perceptions that tutorials are more interactive, engaging, and enjoyable. Although the classroom response systems can provide greater awareness to both students and instructors of the condition of the students' understanding, which would lead to more responsive instruction and better understanding, the results did not show that this translated to higher marks and lower failure probability. Although the increased student participation did not translate into improved outcomes in this experiment, there may be other benefits that student engagement with the subject may generate. For example, increased socialization with classmates may contribute to good mental health and an ability to adjust to university life. Dawson et al. (2014) and Paloyo et al. (2016) indicated that the interaction amongst peers in peer-assisted learning schemes provides an enhanced feeling of support, promotes social relationships and personal wellbeing, and reduces anxiety. It could also increase the probability of students completing the subject irrespective of the mark. Unfortunately, these hypotheses could not be tested due to the lack of data. Still, future studies looking at the impacts of active learning techniques should examine whether other benefits or costs materialize beyond the final grade. The experimental results could yield a lower bound for the treatment effect for several reasons. First, this was the first time these tutors were conducting this pedagogical technique, and there is reason to believe that they may improve over time. Time constraints also could have hindered an instructor from making quality adjustments to the content when incorporating online polling, limiting the treatment effect. While the tutors are trained in using *Learning Catalytics*, they may have become more familiar with the software and utilize its features extensively. Stowell (2015) identifies the need for teachers to adjust their teaching styles using CRS and find more effective ways of delivering misunderstood topics. Second, the treatment was not intense. In particular, the students were not asked to discuss their answers during polling. Not all the features available in *Learning Catalytics* were used, such as the peer learning feature of grouping students and encouraging discussion. The tutors encouraged students to discuss among themselves when reviewing the

Learning Catalytics questions, but the *Learning Catalytics* feature to pair students who got the answer wrong with another who got it right was not used. Studies have shown that student engagement with the course content increases in classes incorporating electronic feedback devices with discussions (Bode et al., 2009; Lasry, 2008). Moreover, the studies showed that students who participate interactively in classes learn the material better, retain concepts longer, and apply them more effectively than students who do not. Still, it is possible that student polling could have no impact on academic performance. Finally, our findings also carried important implications for the teaching of statistics. Statistics is commonly viewed as a dull and daunting subject. Most students study statistics unwillingly (Gordon, 2004). Their course evaluations of statistics units tended to indicate poor engagement (Lavidas et al., 2020). Not surprisingly, students' tutorial attendance was usually low. Preis et al. (2011) showed that online polling promotes enjoyable learning, enhancing intrinsic motivation. If the use of *Learning Catalytics* in a first-year undergraduate statistics subject could lead to significant improvement in tutorial attendance, this could impact students' willingness to persist in their learning and application of skills, and possibly a change in their attitudes and beliefs about statistics.

ACKNOWLEDGEMENTS

REFERENCES

Arjomandi, A., Seufert, J., O'Brien, M., & Rossetto, L.C. (2021). An empirical study of business student engagement with active teaching strategies: A comparison of first year and senior students. *American Business Review*, *24*(1), 67–89.

Balta, N., & Guvercin, S. (2016). Increasing undergraduate students' exam performances in statistics course using software Socrative. *The Turkish Journal of Educational Technology*, *15*, 314–321.

Becker, W. E., & Watts, M. (1995). Teaching tools: Teaching methods in undergraduate economics. *Economic Inquiry*, *33,* 692–700.

Bloom, B. (1984). *Taxonomy of educational objectives, Book 1: Cognitive domain* (2nd ed.). Addison Wesley Publishing.

Bode, M., Drane, D., Kolikant, Y. B.-D., & Schuller, M. (2009). A clicker approach to teaching calculus. *Notices of the American Mathematical Society, 56*, 253–256.

Burden, K., & Kearney, M. (2018). Designing an educator toolkit for the mobile learning age. *International Journal of Mobile and Blended Learning, 10,* 88–99.

Burnstein, R. A., & Lederman, L. M. (2003). Comparison of different commercial wireless keypad systems. *The Physics Teacher, 41*, 272–275. https://doi.org/10.1119/1.1571261

Buzzard, C., Crittenden, V. L., Crittenden, W. F., & McCarty, P. (2011). The use of digital technologies in the classroom: A teaching and learning perspective. *Journal of Marketing Education*, *33*, 131–139.

Caldwell, J. E. (2007). Clickers in the large classroom: Current research and best practice tips. *CBE- Life Sciences Education, 6*(1), 9–20. https://doi.org/10.1187/cbe.06-12-0205

Carter, L. K., & Emerson, T. L. (2012). In-class vs. online experiments: is there a difference? *The Journal of Economic Education*, *43,* 4–18.

Caviglia-Harris, J. (2016). Flipping the undergraduate economics classroom: using online videos to enhance teaching and learning. *Southern Economic Journal*, *83,* 321–331.

Crouch, C. H., & Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American Journal of Physics, 69*, 970–977

Dabbour, E. (2016). Quantifying the effects of using online student response systems in an engineering ethics course. *Journal of Professional Issues in Engineering Education and Practice*, *142*, 04015010, https://doi.org/10.1061/(ASCE)EI.1943-5541.0000260

Dawson, P., van der Meer, J., Skalicky, J., & Cowley, K. (2014). On the effectiveness of supplemental instruction: A Systematic review of supplemental instruction and peer-assisted study sessions literature between 2001 and 2010. *Review of Educational Research, 84,* 609–639

Durham, Y., McKinnon, T., & Schulman, C. (2007). Classroom experiments: Not just fun and games. *Economic Inquiry, 45,* 162–178.

Dynan, R. (2014). Teaching the hierarchy of effects using live audience polling in the classroom. In *Proceedings of the Marketing Management Association, 2014 Fall Educators Conference*, *San Antonio, TX* (pp. 208–213).

Emerson, T. L. N., & Taylor, B. A. (2004). Comparing student achievement across experimental and lecture-oriented sections of a principles of microeconomics course. *Southern Economic Journal, 70,* 672–693.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. In B. Alberts (Ed.), *Proceedings of the National Academy of Sciences, 111*(23)*,* 8410–8415. https://doi.org/10.1073/pnas.1319030111

Gordon, S. (2004). Understanding students' experiences of statistics in a service course. *Statistics Education Research Journal*, *3*(1), 40–59.

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.

Hall, S. R., Waitz, I., Brodeur, D. R., Soderholm, D. H., & Nasr, R. (2002). Adoption of active learning in a lecture-based engineering class. In *Leading a revolution in engineering and computer science education.* Proceedings of the 32nd ASEE/IEEE Frontiers in Education Conference, Boston, MA (pp. T2A9–T2A15).

Jensen, E. J., & Owen, A. L. (2003). Appealing to good students in Introductory Economics. *Journal of Economic Education*, *34*(3), 299–325. https://doi.org/10.1080/00220480309595225

King, G., & Sen, M. (2013). How social science research can improve teaching. *PS: Political Science and Politics*, *46*(3)*,* 621–629.

Korobova, N., & Starobin, S. S. (2015). A comparative study of student engagement, satisfaction, and academic success among international and American students. *Journal of International Students*, *5*(1), 72–85.

Lage, M. J., & Platt, G. (2000). The internet and the inverted classroom. *Journal of Economic Education*, *31*(1)*,* 11.

Lage, M. J., Platt, G.J., & Treglia, M. (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *The Journal of Economic Education*, *31*(1)*,* 30–43.

Lasry, N. (2008). Clickers or flashcards: Is there really a difference? *The Physics Teacher, 46*, 242–244.

Lavidas, K., Barkatsas, T., Manesis, D., & Gialamas, V. (2020). A structural equation model investigating the impact of tertiary students' attitudes toward statistics, perceived competence at mathematics, and engagement on statistics performance. *Statistics Education Research Journal, 19*(2), 27–41. https://doi.org/10.52041/serj.v19i2.108

Marburger, D. R. (2005). Comparing student performance using cooperative learning. *International Review of Economics Education*, *4*(1)*,* 46–57. https://doi.org/10.1016/S1477-3880(15)30138-9

Mazur, E. (1997). *Peer instruction: A user's manual series in educational innovation*. Prentice Hall.

Mazur, E., King, G., & Lukoff, L. C. (2013). *Learning Catalytics*. Pearson. [software] https://www.pearson.com/us/higher-education/products-services-teaching/learning-engagement-tools/learning-catalytics.html

McLaughlin, J. E., Roth, M. T., Glatt, D. M., Gharkholonarehe, N., Davidson, C. A., Griffin, L. M., Esserman, D. A., & Mumper, R. J. (2014). The flipped classroom: A course redesign to foster learning and engagement in a health professions school. *Academic Medicine*, *89*(2)*,* 236–243.

Mercader, C., & Gairín, J. (2020). University teachers' perception of barriers to the use of digital technologies: The importance of the academic discipline. *International Journal of Educational Technology in Higher Education*, *17*, Article 4.

Mestre, J. P., Gerace, W. J., Dufresne, R. J., & Leonard, W. J. (1996). Promoting active learning in large classes using a classroom communication system. *The Changing Role of Physics Departments in Modern Universities.* Proceedings of the International Conference on Undergraduate Physics Education, College Park, MD. https://doi.org/10.1063/1.53119

Mishra, P., & Koehler, M. J. (2006). What is technological pedagogical content knowledge? *Contemporary Issues in Technology and Teacher Education, 9*(1)*,* 60–70.

Muir, S., Tirlea, L., Elphinstone, B., & Huynh, M. (2020). Promoting classroom engagement through the use of an online student response system: A mixed methods analysis. *Journal of Statistics Education, 28*(1), 25–31. https://doi.org/10.1080/10691898.2020.1730733

Nicol, D. J., & Boyle, J. T. (2003). Peer instruction versus class-wide discussion in large classes: A comparison of two interaction methods in the wired classroom. *Studies in Higher Education, 28*(4), 458–473. http://dx.doi.org/10.1080/0307507032000122297

Noel, D., Stover, S., & McNutt, M. (2015). Student perceptions of engagement using mobile-based polling as an audience response system: Implications for leadership studies. *Journal of Leadership Education, 14*(3), 53–65. https://doi.org/1012806/V14/I3/R4

Paloyo, A. R., Rogan, S., & Siminski, P. (2016). The effect of supplemental instruction on academic performance: an encouragement design experiment. *Economics of Education Review, 55,* 57–69.

Parmelee, D., Michaelsen, L. K., Cook, S., & Hudes, P. D. (2012). Team-based learning: A practical guide. *AMEE guide, No. 65. Medical Teacher*, *34*(5), 275–287. https://doi.org/10.3109/0142159X.2012.651179

Paschal, C. B. (2002). Formative assessment in physiology teaching using a wireless classroom communication system. *Advances in Physiology Education, 26*, 299–308.

Preis, M. W., Kellar, G. M., & Crosby, E. (2011). Student acceptance of clickers in large introductory business classes. *American Journal of Business Education, 4*(5), 1–14. https://doi.org/10.19030/ajbe.v4i5.4219

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education*, *93,* 223–231. https://doi.org/10.1002/j.2168-9830.2004.tb00809.x

Reay, N. W., Bao, L., Pengfei, L., Warnakulasooriya, R., & Baugh, G. (2005). Toward an effective use of voting machines in physics lectures. *American Journal of Physics*, 73, 554–558.

Roschelle, J., Penuel, W. R., & Abrahamson, A. L. (2004). *Classroom response and communications systems: Research review and theory.* Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.

Salemi, M. K. (2009). Clickenomics: Using a classroom response system to increase student engagement in a large-enrollment principles of economics course. *Journal of Economic Education, 40,* 385–404.

Salzer, R. (2018). Smartphones as audience response systems for lectures and seminars. *Analytical and Bioanalytical Chemistry, 410*, 1609–1613.

Stowell, J. R. (2015). Use of clickers vs. mobile devices for classroom polling. *Computers and Education, 82*, 329–334. https://doi.org/10.1016/j.compedu.2014.12.008

Sun, J. C., Martinez, B., & Seli, H. (2014). Just-in-time or plenty-of-time teaching? Different electronic feedback devices and their effect on student engagement. *Education Technology and Society, 17,* 234–244.

Sun, J. C. (2014). Influence of polling technologies on student engagement: An analysis of student motivation, academic performance, and brainwave data. *Computers and Education, 72*, 80–89.

Tait, K. (2009). Understanding tertiary student learning: Are they independent thinkers or simply consumers and reactors? *International Journal of Teaching and Learning in Higher Education*, *21*(10), 97–107.

Thomas, L., & Heath, J. (2014). Institutional wide implementation of key advice for socially inclusive teaching in higher education: A practice report. *International Journal of the First Year in Higher Education*, *5*(1), 125–133. https://doi.org/10.5204/intjfyhe.v5i1.206

Yamarik, S. (2007). Does cooperative learning improve student learning outcomes? *The Journal of Economic Education, 38*(3)*,* 259–277. http://dx.doi.org/10.3200/JECE.38.3.259-277

ALFREDO R. PALOYO
School of Business
University of Wollongong
Northfields Avenue
Keiraville, New South Wales 2500
Australia