

Mobile Applications to Measure Students' Engagement in Learning

John C Chen¹, David Janzen², Ryan Chang³, Karen Mcgaughey⁴ and Jim Widmann⁵

¹Mechanical Engineering, Computer Science and Software Engineering, California Polytechnic State University San Luis Obispo, CA, USA

²California Polytechnic State University San Luis Obispo, CA, USA

³Outreach University of Oregon Eugene, OR, USA

⁴Mechanical Engineering, Statistics California Polytechnic State University San Luis Obispo, CA, USA

⁵California Polytechnic State University San Luis Obispo, CA, USA

RESEARCH

Abstract

Evidence-based instruction or active learning is being more widely implemented in college teaching, and there is a need for instructors, evaluators and researchers to quantify their implementation in order to, for example, determine the efficacy of a new instructional technique. Here we introduce a new method for measuring students' level of engagement with their learning. The method relies on an established and research-based theoretical framework and is built in the form of a mobile application for the two most popular smartphone platforms. Five separate studies presented here establish the fidelity of the method, its ability to measure subtle variations among students within the same class, the students' patterns of learning during out-of-class study periods, and the versatility of the app to make different measurements of learning in different contexts, including an exploratory examination of the impact of the sudden shift to remote learning prompted by the coronavirus pandemic.

Keywords: learning engagement, evidencebased instructional practices, active learning, mobile applications

1 Introduction

Over the past several decades, there has been a shift in college teaching, especially in STEM disciplines, toward the use of evidence-based instructional practices (EBIP) [1–5], which are based on research demonstrating improved student performance when these practices are used. Many of these instructional techniques are aligned with the broad pedagogy of active learning [3], which has as its primary goal of increasing students' engagement in their learning. Prince [1] describes active learning as requiring “students to do meaningful learning activities and think about what they are doing” while engaging in the designed activities. Not all active learning is consistently effective [6], however, perhaps because of other factors such as the subject being incompatible with the technique, the instructor's lack of familiarity with the technique, or the lack of adherence to important aspects of the technique. Even with this caveat in mind, most educational researchers and those engaged with policy making [7–9] support the use of EBIP and active learning to improve student outcomes. With increased interest and implementation of EBIP and active learning, there is a need to measure the students' level of engagement with their learning in order to satisfy professional (e.g. teaching evaluation or improvement) or research needs. In this paper, we describe a smartphone-based method for this measurement, compare

OPEN ACCESS

Volume
13

Issue
1

Submitted 9 Sep 2022

Accepted 13 Dec 2021

Citation

Chen J.C., Janzen D., Chang R., Mcgaughey K. and Widmann J. “Mobile Applications to Measure Students' Engagement in Learning,” *Computers in Education Journal*, vol. 13, no. 1, 2022. doi:

its salient features to other existing methods, and demonstrate its abilities to gather information about how students engage with their learning in various engineering contexts.

2 Background and Context

Brief descriptions of existing methods for measuring learning engagement are provided below, as well as details of the development and implementation of our method. A common characteristic of all the methods is their reliance on measurements made during the learning activity or very soon thereafter. This element is critical since retrospective self-reports (i.e., delayed recall) are known to be highly inaccurate due to recall bias [10, 11]. This measurement characteristic is also superior to retrospective recalls because it occurs in real-time, or nearly so, to a specific event of interest and in the subject's natural ecology, which provides the data with context and ecological validity [12]. While the various measurement methods are suitable for a wide range of college subjects, this paper focuses on engineering studies due to the student populations being reported here. Also included is a preliminary examination of data collected during the time of the coronavirus pandemic to investigate the impacts on the students' patterns and habits of learning.

2.1 Existing methods for measuring learning engagement

Several methods for measuring learning engagement already exist in the literature, with several that have appeared in recently published literature and seem well suited for use in classes in which active learning or evidence-based practices are in use. While this brief review of other methods for measuring learning engagement is not meant to be exhaustive, it does present the most salient features of these methods and their advantages and drawbacks.

The Teaching Dimensions Observation Protocol (TDOP) was developed to examine, in a descriptive rather than evaluative way, behaviors and practices that are aligned with "interactive teaching" in a classroom [13, 14]. It is comprised of five categories that represent features of instruction, including teaching methods, pedagogical strategies, student-teacher interactions, cognitive engagement and instructional technology. Two criticisms of TDOP are its reliance on substantial judgement on the part of observers [15] and, as a result, its need for extensive training to reach acceptable levels of interrater reliability [14].

The PORTAAL (Practical Observation Rubric To Assess Active Learning) tool was designed based on a review of education research literature to identify best practices in active learning [16]. It includes 21 elements that have been shown to improve student learning outcomes. PORTAAL's creators claim that it is easy to learn, is validated, and has high interrater reliability. Its major drawback is that, because of so many elements being measured, it requires a video recording for observation and measurement. In addition, the protocol relies solely on observing the instructor, which may not always align with what students are doing.

The Classroom Observation Protocol for Undergraduate STEM (COPUS) [15] was developed to overcome several shortcomings of previous observation protocols and was specifically designed for the modern STEM classroom in which an instructor might be employing several forms of active learning activities. Its development evolved from the TDOP and, similar to that, COPUS relies on observing and categorizing what the students and instructor are doing in 2-min. intervals throughout a class meeting. The protocol categorizes these behaviors into 25 codes. Its creators claim that reliability is achieved after a 1.5-hour training period. Importantly, COPUS, as the authors acknowledged, could not judge the cognitive level of the participants since it relied solely on in-class observers for measurements.

2.2 New method for measuring learning engagement

As alluded to earlier, our method collects data from individual students rather than either observing the students and/or the instructor, or aggregating data across a cluster of students. This is

achieved by building our measurement in the form of a smartphone application (or app), called Actively Learning (ALApp). In this section we describe the theoretical framework on which our measurement method is based, as well as the architecture and technological resources supporting it.

2.2.1 Framework for measurements

Students learn engineering in a variety of contexts and through various activities. They experience various levels of active learning through attending lectures, completing homework assignments, preparing for class, studying for quizzes and examinations, and seeking additional help. To describe these experiences for a complete measure of each student's quality and quantity of learning engagement, we used a framework developed by Chi and coworkers: the interactive-constructive-active-passive (ICAP) differentiated learning activities [17, 18].

The ICAP framework classifies learning activities by observable, overt actions of the learner. A passive learning activity is one in which the learner essentially engages in no overt actions. Listening to a lecture, watching a video, and reading text are examples. By contrast, an active learning activity is characterized by overt actions that demonstrate paying attention. Examples include note taking or highlighting of text. (Note that at this point "active" learning has taken on a definition that is quite different than the general use of the term in education, which would classify note taking as a "passive" learning activity. The use of "active" learning here adheres to Chi's ICAP framework.) If the learner goes one step further and generates additional knowledge or information beyond that which is provided, she is engaging in constructive learning. Solving homework problems alone or resolving questions while reviewing notes alone are examples of this. The final category of interactive learning requires learners to interact with someone (e.g., a peer or expert) or something (e.g., a computer tutor) in order to build on the provided information. There must be an exchange of information between the members, such as defending one's responses, responding to questions, or correcting noted errors. The conventionally accepted active-learning techniques [1–4] would be classified as either constructive or interactive in the ICAP framework. Furthermore, based on the possible underlying cognitive mechanisms being activated by each kind of activity, the expected learning gains should increase in the order of passive < active < constructive < interactive; this is supported by the studies cited by Fonseca and Chi [18].

Our measurement method adopts the ICAP framework to measure the quality of active learning (or engagement level) experienced by study participants, with passive learning being lowest quality and interactive learning as the highest. The quantity of active learning is then simply the amount of total time students expend under each of the four ICAP categories for each course. A smartphone app to capture these data is desirable since it is convenient and familiar to students and facilitates data collection and storage. The app also sends reminders to the student after each scheduled class lecture or study session, as well as a few other times throughout the day to capture other learning experiences (e.g., study or homework time, or office-hour visit). The student would then record the quality and quantity of each learning experience within the app, which stores this data locally and uploads them automatically to a server whenever it connects to the Internet.

2.2.2 App development and architecture

A primary product of this project is the software application that was developed for collecting student data. This section describes ALApp's software architecture which involves the selection of software technologies and their organization. Software technologies evolve rapidly, with new technologies emerging frequently. As a result, selecting a robust, secure software architecture that will remain stable over time and can evolve to support unforeseen features is both challenging and essential.

a) Technology stack. [Figure 1](#) shows the organization of the primary software technologies used to create ALApp. Parse Platform is a mobile backend as a service that powers the ALApp. Parse Platform provides native iOS and Android Software Development Kits (SDKs) and a push notification service. Parse Platform was chosen over alternative services for having native SDKs, a

hosted cloud service, and a generous no-cost tier. The iOS version of ALApp uses the standard iOS SDK and Swift programming language. The Android version uses the standard Android SDK and Java. Developing native iOS and Android apps was chosen at the time over using a cross-platform tool, such as Xamarin or PhoneGap, based on lower perceived risk and the development team's existing expertise. Additionally, push notifications to non-native mobile web solutions had significant restrictions compared to native apps. Parse, Inc was acquired by Facebook in 2013 and eventually shut down but released to open-source as Parse Platform. We hosted our own Parse Platform instance on a Linode server. Parse Platform provides a web Dashboard for convenient administration of the Parse database and, despite the release of several new versions over the lifetime of this study, it has remained stable enough to support additional ALApp features. The remaining pieces of the architecture include a 'Class Scraper' script that we use at the start of each academic term to collect the course information (course name, instructor, days and times of class meeting) from the university's public database and to populate ALApp, and a database (MongoDB) that stores all data supporting and collected by ALApp.

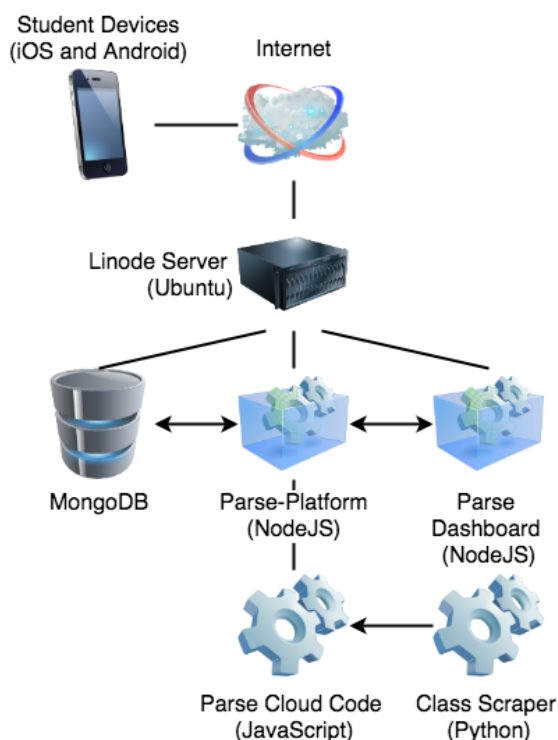


Figure 1. Actively Learning software architecture

b) Database schema. [Figure 2](#) shows the database schema implemented in Parse for the ALApp. Parse Platform uses MongoDB as its backend database. Arrows represent a Parse Platform reference from one table to another, called a Pointer. The Installation table keeps track of the Universally Unique Identifiers (UUIDs) of mobile devices required to send push notifications. The RegCodes table contains the list of approved codes to log students into the ALApp. The AvailableClasses table contains the list of classes from which students select their target courses. RegisteredClassTimes holds the list of classes students selected, and the association with a particular student. ICAPActivity holds the recorded student activity data.

c) Cloud code. ALApp uses Parse Cloud Code's Parse JavaScript SDK to interact with the database on the server side. Cloud Code powers the push notification service, email reminder system, and a Python-based web scraper to automatically populate class data for the AvailableClasses table at the start of each academic term. Cloud Code functions ensure push notifications are sent accurately and according to the specified schedule. Functions also monitor the database to send automatic email reminders to students when expected data entries are missed, and ingest the class data from the Python scraper.

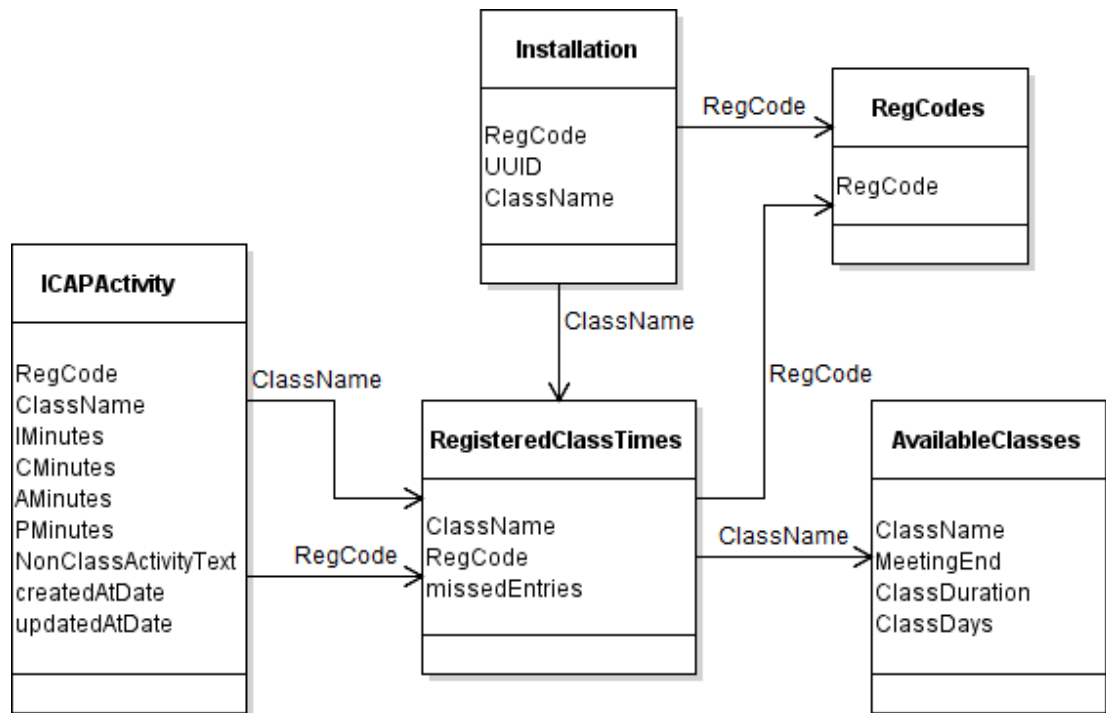


Figure 2. Actively Learning database schema

2.3 Summary of app and comparison with other methods

ALApp sends reminders (notifications) to students to record both the quantity (measured by time) and quality (based on ICAP scale) of learning engagement throughout a day in order to minimize errors due to memory recall. Notifications are sent immediately after a class meeting (Figure 3) and otherwise every three hours, from 10 am to 10 pm (Figure 4). Figure 5 shows the user interface for making a data entry in the app. Note that, from this screen, the user can tap on each of the I, C, A or P letters to pop-up a brief definition of that level of the scale as a reminder of the ICAP framework. The app is connected via the cellular network or Wi-Fi to a server that stores all of the data (and also displays prior data, which can be edited or deleted if necessary). What is stored on the server, therefore, is a database containing all users, the course or courses being tracked for each participant, and the ICAP data for both in- and out-of-class learning periods. The latter contains time spent under each of the I, C, A or P levels, the learning event (e.g., class or homework or office hour) being recorded, and the date and time of each set of ICAP entries. The database can be exported from the server and imported into common software for analysis. ALApp differs from other methods to measure student engagement in learning in three important ways: (1) It collects data from the students' viewpoint instead of observing what the instructor does; (2) it measures learning during and outside of class meetings; and (3) data are collected from each student rather than aggregated across all students or a cluster of students in a class. A comparison of the most salient features among the various measurements methods is presented in Table 1 .

3 Study Methods

All five studies reported here took place at a large, western-U.S., state-supported university. The participants were a convenience sample of compensated volunteers drawn from particular courses that were each the focus of the study, although the particular subject of each course was not salient to the study. Participation in the study did not have any effect on each participant's grade.

Study 1 comprised 14 students taking an introductory thermodynamics course during the same academic term. The participants were a mix of engineering majors and year-of-study. The par-

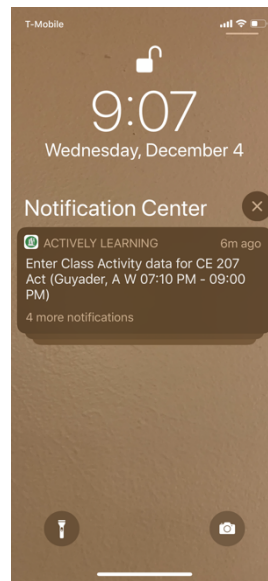


Figure 3. Screenshot from an iOS-based smartphone of post-class notifications

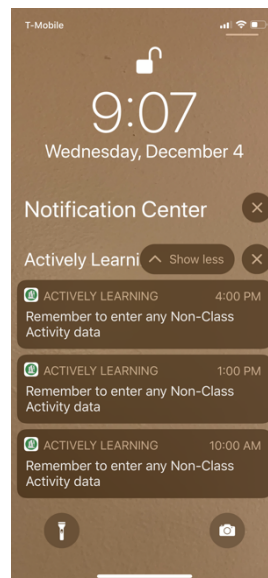


Figure 4. Screenshot from an iOS-based smartphone of non-class notifications

Participants were trained on the ICAP framework and use of the ALApp during a ~50 min. training session conducted in person approximately one week prior to the start of the study. Study 2 involved 42 mechanical engineering students who were at approximately the same time point in their academic careers, which was the start of their second year of studies. The participants were recruited from students taking the first mechanics course (engineering statics) within a sequence of five mechanics courses in the curriculum. The participants were trained on the ICAP framework and use of the ALApp through two on-line modules created by the investigators and hosted on the university's learning management system (Moodle). The training was estimated to take approximately 45 min. Three participants from the Study 2 sample were selected at random and their learning patterns examined in detail for Study 3. Study 4 comprised 29 students in the Software Engineering Capstone course, which had a total enrollment of 68 students. Finally, Study 5 took place during the 2020 coronavirus pandemic (while the previous four studies took place prior to it) and compared the study patterns of students prior to and during the forced shift to remote (online) learning.

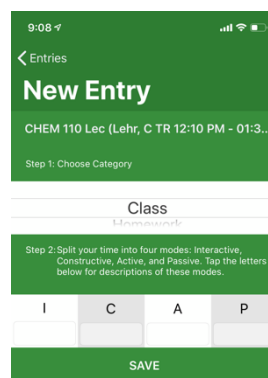


Figure 5. Screenshot from an iOS-based smartphone of the user interface for data entry. Android-based smartphones are similar.

Table 1. Comparison of various methods of measurement for learning engagement

Measurement method	Method's features				
	Data recorder	Data source	Data type	Training required	In-/out-of-class data
Teaching Dimensions Observation Protocol (TDOP)	External observer	Instructor and students	Qualitative	Extensive	In-class
Practical Observation Rubric To Assess Active Learning (PORTAAL)	External observer or instructor	Instructor	Quantitative	4-5 hours	In-class
Classroom Observation Protocol for Undergraduate STEM (COPUS)	External observer	Instructor and students	Quantitative	1.5 hours	In-class
Actively Learning app (ALApp)	Student	Student	Quantitative	~1 hour	In-class and out-of-class

4 Results and Discussion

Results from five studies are presented and discussed below. Study 1 and Study 2 have been presented in a prior conference [19] but are summarized here to provide context and validity for the remaining three studies that are the foci of this paper to demonstrate the types of data gathered and provide insights into students' engagement with learning.

4.1 Study 1

The primary purpose of Study 1 was to validate the fidelity of the data recorded by the students through ALApp. Fourteen students taking an introductory thermodynamics course from one of two instructors were the participants. Instructor A relied almost exclusively on lecturing during classes while Instructor B used an active learning pedagogy that requires students to do individual

work before each class meeting and, during class, to work in long-term groups to solve problems or complete quizzes. Instructor B used brief lectures (< 10 min.) to set the context for each day's activity.

The findings from the study [19] showed the difference in pedagogy between the two instructors was clearly seen in the data: Students in the lecture class recorded nearly all Active and Passive engagement, while the active learning class recorded a mix of all four engagement levels, with a majority being at the Constructive or Interactive levels. The data also showed that variations existed between participants within each class, and these individual variations were confirmed by two investigators who attended a randomly selected class to make direct observations of students (recall that the ICAP framework relies on overt, observable actions). The agreements between the investigators and between each investigator and the participant were very good and, importantly, the data showed that variations in level of engagement did indeed exist between students within the same class. This finding points to the importance of tracking student engagement individually as opposed to an average across all or a cluster of students.

4.2 Study 2

The study period was the final four weeks of the 10-week quarter during which the students were taking the first of a sequence of mechanics course (engineering statics). The students were enrolled in one of 14 possible sections of the course, taught by six different instructors. Most instructors relied on traditional lecturing, but one instructor used an informal active learning method in which a topic was briefly introduced and the students were formed into ad hoc teams of two or three people to work through problems as the instructor roamed the class to observe and assist. The objectives of Study 2 were to confirm the variations in levels of engagement among students within the same class and to examine the students' learning habits outside of class.

The data [19] confirmed again that, even within the same class and regardless of the pedagogy used, students cognitively experience each class differently as exhibited by their reported ICAP time distributions. These relatively smaller variations between students, however, did not mask the demonstrable difference between the instructors' pedagogical style (i.e., more active learning class vs. a purely lecture class). The students' out-of-class study habits were surprisingly varied, as measured by the number of out-of-class study events, with some averaging just less than once per week of such study events to over 10 times per week. The vast majority of these events were for homework but significant numbers were also recorded for office-hour visits, group study, or reviewing of notes. The variations in the students' frequency of out-of-class entries did not, however, result in a large variation in the total amount of time spent in out-of-class studies. This finding was independent of the instructor (and therefore the instructional mode) and likely demonstrated the difference between students' study strategies.

4.3 Study 3

The objective of this study is to examine more closely the learning habits of three participants, randomly drawn from Study 2, as they progress through the mechanics sequence within the curriculum. While these students are not necessarily representative of the entire study population, nor is this study trying to draw conclusions about particular learning habits or patterns, this initial examination provides a first glimpse of how students navigate a complex curriculum while learning increasingly challenging content under various instructional methods. The sole criterion used to select these three participants for comparison is that all three completed each course in the mechanics sequence at the intended time designated by the curriculum, and therefore all three completed each course concurrently. The majority of the 42 participants from Study 2 met this criterion and these three were randomly selected. All three students are high achieving, with current overall grade point averages above 3.65 (out of 4.00).

[Table 2](#) examines the quality of the in-class learning experiences of the participants and shows the percent of total class meeting times for each course which was spent at the Interactive +

Constructive levels of cognitive engagement. Similar to Study 1, the pedagogical style of various instructors can easily be discerned from the data. For example, all three students had the same instructor for ME 211 and it is clear that they were actively engaged for much of these class meetings. A similar conclusion can be drawn for ME 212, where two different instructors were involved.

Interestingly, the data seem to suggest that Student B was able to self-motivate and cognitively engage at the Interactive and Constructive levels during class, regardless of the instructor's teaching style. This can be seen when comparing the data from Table II for CE 204, CE 207 Tutorial, CE 207 and ME 326 Tutorial. Student B always reported a high level of cognitive engagement while one or both of the other students, who had the same instructor, did not. This suggests, perhaps, that some students are able to motivate themselves to engage with the class regardless of the pedagogical style of the instructor. This finding again highlights the importance of measuring learning engagement for individual students instead of an aggregate of them.

Table 2. Percent of total class meetings spent at the I+C. levels

	Student A	Student B	Student C
ME 211 ^b	45.0 ^{a,1}	59.2 ¹	49.0 ¹
CE 204 Tutorial	61.3	77.3	23.8
CE 204	2.9 ²	43.7 ²	0.0 ²
ME 212	43.9 ³	66.0 ³	79.6
CE 207 Tutorial	36.6 ⁴	71.6 ⁴	32.8
CE 207	11.1 ⁵	69.0 ⁵	0.0
ME 326 Tutorial	24.2 ⁶	51.4 ⁶	97.5
ME 326	10.0 ⁷	58.6	12.8 ⁷

a. Identical numerical superscripts denote the same instructor for that course

b. ME 211=statics; CE 204=mechanics of materials I; ME 212=dynamics; CE 207=mechanics of materials II; ME 326=intermediate dynamics

The learning patterns and habits of these three students during out-of-class times are examined in Table 3. For this comparison, only the lecture portion of each of the five courses are included (tutorials were excluded). The values shown in Table III represent the averages per student for all five mechanics courses.

As expected, doing homework was the most common out-of-class activity for all three students. These students, to varying degrees, also attended office hours, reviewed their textbook, practiced with additional problems and attended study groups. Generally, all three students used a variety of study strategies (Table 3, first row), but the frequency of these uses varied widely (second row). While Students B and C averaged approximately 30 out-of-class study entries per course, Student A averaged 56.4 entries. Student A also averaged the most time spent per course in out-of-class studying, while the other two students had similar study times (third row). Interestingly, of the total out-of-class study times, the average time per course devoted to completing homework was roughly the same for all three students, ranging from 1471 to 1700 min. per course in a 10-week term (fourth row). This finding suggests that Student A spends a majority share (51.8%) of out-of-class studying time on activities that are not mandatory. Such a finding might suggest, perhaps, better self-regulatory behavior for Student A in comparison to another student who spends the vast majority of out-of-class time on completing homework only. It should be pointed out that Students B and C, while spending roughly one-third of their out-of-class time on non-mandatory studies, are also academically very strong, which suggests that their approach, perhaps, may be sufficient and more efficient than Student A.

4.4 Study 4

The purpose of this study is to demonstrate the versatility of the ALApp for conducting different studies with other objectives. One such possibility is the effort and teamwork that team members

Table 3. Out-of-class study habits and patterns^c

	Student A	Student B	Student C
Number of types of study strategies	4.8	4.4	2.8
Number of entries for all strategies	56.4	30.3	27.6
Total time of all entries (min.)	3190	2527	2341
Total time of all entries spent on homework (min.)	1538	1700	1471
Percent of time not doing homework	51.8%	32.7%	37.2%

c. Values shown represent averages for all five mechanics courses

contribute while working on a group project. This study was conducted in a single-term, senior capstone project in software engineering, in which students form teams and design a software solution for an actual client. In such a learning environment, the ICAP framework of learning engagement would not be relevant but what is valuable instead is a measurement of the students' efforts toward each project and the mode of work involved. Twenty-nine of 68 students in the course were compensated volunteers for this study. Participation in the study had no effect on the students' grades and the course instructor was blinded as to which and how many students were participating.

For this study the participants were presented not with recording ICAP times but with four different modes of work on their project: TPIR – Team, Partial team, Individual, and Remote (online or at distance, and regardless of whether it was work done alone or with the entire or partial team). Similar to how the app is used in the previous studies, students are prompted by notifications immediately after a class meeting or at fixed intervals otherwise to enter the amount of time they have worked on their project in each of the four possible modes. The modification of the ALApp to accommodate this study was simple and merely involved flagging the participants in this study differently than the other participants in the server's database (two additional studies using the ICAP framework were simultaneously being undertaken). This flag triggered the ALApp to fetch the TPIR categories and present these to the participants instead of the ICAP categories. The notifications did not need to be modified for use with this study.

Figure 6 presents the time distributions of all 29 participants for the 10-week project. Each participant's total time spent on the project is shown as a stacked column that includes time working remotely, individually, as a partial team, or with the entire team. The number above each column is the total number of entries by that participant in ALApp. Each separated cluster in Figure 6 denotes a separate project team; no team had all members participate in this study.

The variation in effort (as measured by total time spent) among the participants is striking, varying from 12.8 to 100.8 hours over the 10-week project. Although there was some relationship of this time between team members of the same project (e.g., members 1, 2, and 3; total times of 70.0, 76.7, and 87.3 hours, respectively), there was also evidence of large disparities within a team (e.g., members 11-15; total times of 69.2, 66.3, 32.3, 24.7, and 79.0 hours). Finally, the differences in how each team member worked on the project is also apparent in the data. For example, members 1-3 worked almost exclusively either individually or with the whole team, while members 26 and 27 worked roughly equally independently, in a partial team, or with the whole team. These differences may reflect the preferences of team members or could have been necessitated by the nature of the project, though we cannot make this distinction from the data.

Figure 6 also shows the different patterns or habits by which the participants worked on the project, as demonstrated by the number of entries that each participant made through the app. There was a large variation in this measure, from 8 total entries (or an average of less than one entry per week) to 59 (nearly six entries per week). While there is some relationship between the number of entries and the total time devoted to the project, there were also exceptions. For example, participants 11 and 12, who were members of the same team, made 32 and 18 total app entries. Their total time spent on the project, however, were similar at 69.2 and 66.3 hours, demonstrating the different preferences of frequent/shorter work periods versus less frequent/longer work periods, or perhaps such differences signal the different requirements of

each team member to complete their respective tasks for the project.

While ALApp can make measurements of effort toward a project, it cannot, by itself, speak to the efficiency of this work, the value of the effort, or the effects of working individually vs. collaboratively and remotely vs. face-to-face. To do that, measures of quality, such as individual or team project scores or peer evaluations of team members need to also be considered. In a future publication we will explore how measurements made through ALApp are correlated with or support such measures of performance, both individually and as a team, and what they tell us about the efficiency of the participants in achieving their individual performance.

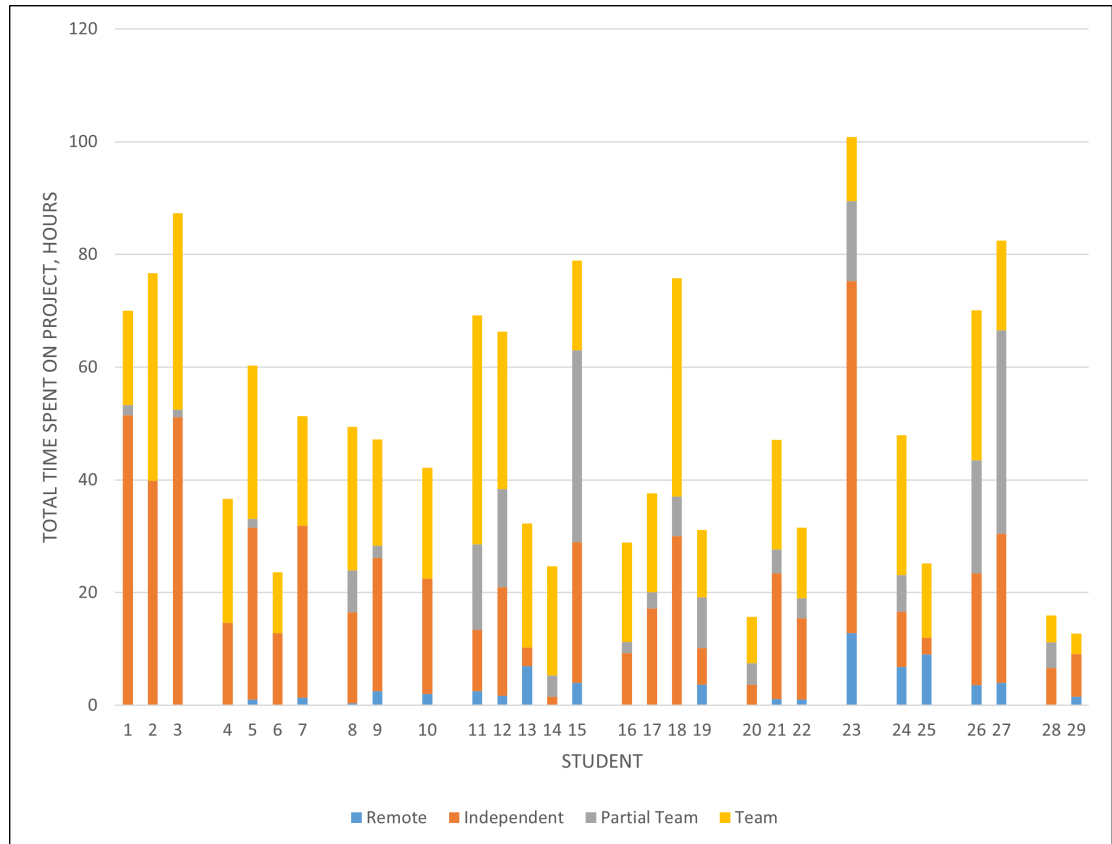


Figure 6. Time distributions among all 29 study participants for a senior software-engineering capstone project. Each separated cluster represents members of the same team. The number above each participant is the total number of entries by that participant during the 10-week-long project.

4.5 Study 5

On March 14, 2020, just prior to final examinations for the winter term (the university operates on a three-term, September to June, academic year), the university shifted all further instructional activities to remote, online learning due to the coronavirus pandemic. This resulted in the spring term being completely online with little time for students and faculty to prepare for the transition. The multitude of changes and adaptations presented an opportunity to use ALApp to examine how the pandemic affected the students' patterns of learning. We did this by looking at the same students' study patterns before and during the pandemic term. (We also recruited and trained a new cohort of students at the start of the shift to online instruction with the goal of looking at different students' study patterns in the same course before and during the pandemic. We plan to describe the full research findings in a separate publication in the future since the focus on this paper is on the capabilities of the ALApp as a research tool.)

In the academic term prior to the pandemic, 16 mechanical engineering students formed the final cohort that had been tracked for five prior consecutive terms through the ALApp to learn about their study patterns and habits in the first two years of engineering studies. Here we compare their study patterns immediately before and during the shift to online learning. Figure 7 and Figure 8 shows for each student their in-class distribution of learning engagement at the Interactive plus Constructive (I+C) levels during the two comparison terms. These two levels of engagement are highlighted since they are the most cognitively engaging [17, 18] and likely the most challenging to achieve in an online format. This analysis includes only lecture courses (as opposed to laboratories or tutorials) in science, math and engineering courses, and the number of courses ranged from two to five depending on each student's schedule.

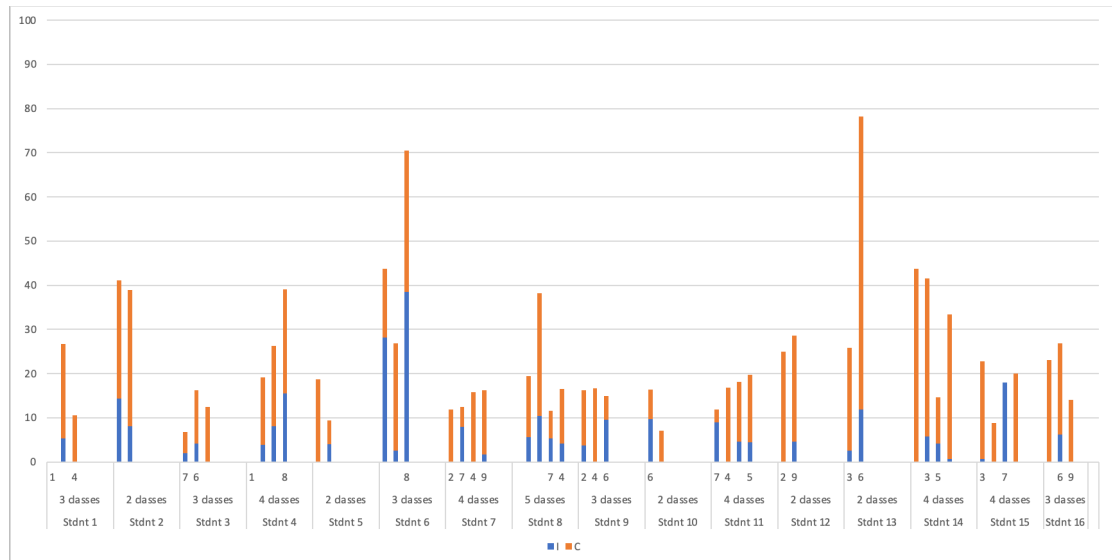


Figure 7. Per-class average levels of engagement at the I+C levels, expressed as a percentage of the average for each class meeting (i.e., $I+C+A+P = 100\%$ for every class) to account for the variation in class duration. Data derived from the pre-pandemic, Winter '20 term. For each of the 16 students tracked, the number of classes for that student is shown above the student identifier (e.g., "Stdnt 1" had three classes in the Winter '20 term and five classes in the Spring '20 term). For the classes in each term, same numbers or letters along the x-axis indicate the identical class during that term.

Several features are prominent from a visual inspection of Figure 7 and Figure 8. First, the frequency of engagement at the Interactive level was much higher in the pre-pandemic term than the pandemic term. This is not surprising given the difficulty of having students interact with one another during online classes, although it is worth noting that several classes achieved this to a significant extent on average (Figure 8). Second, it is mildly surprising that during the pandemic term, many classes still engaged students at the I+C levels when it might be expected that instructors would rely on pure lecturing in online classes, which would have led to mostly Passive and Active (i.e., listening and notetaking) levels of engagement. Still, however, the frequency of classes achieving I or C levels of engagement was substantially lower, as shown in Table 4. Of the total classes tracked by the 16 students during each term, the percent of class meetings during which no engagement at the I+C, I or C levels were all substantially higher in the pandemic term compared to the pre-pandemic term. Surprisingly, however, the average percent of class time spent at the A+P levels was nearly identical between the two terms (last row of Table IV), suggesting that those pandemic-term classes which engaged students at the I or C levels did so at substantially higher amounts compared to the pre-pandemic term. This further suggests two possible reasons for this finding: (1) these instructors invested great efforts to design class meetings that engaged students in substantial amounts of learning that required the constructing of new knowledge, or (2) the students adapted during this period of challenging learning.

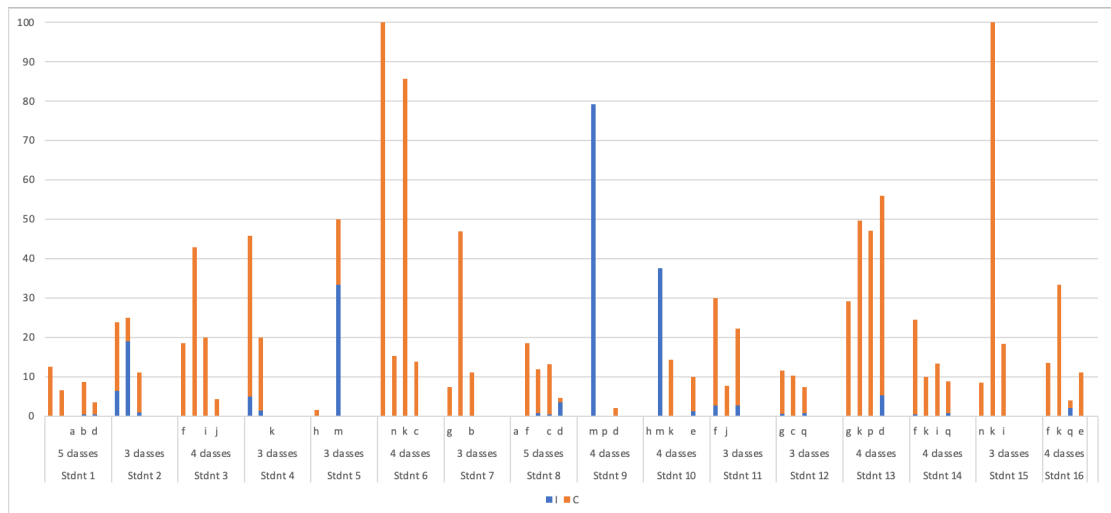


Figure 8. Per-class average levels of engagement at the I+C levels, expressed as a percentage of the average for each class meeting (i.e., I+C+A+P = 100% for every class) to account for the variation in class duration. Data derived from the pandemic, Spring '20 term. For each of the 16 students tracked, the number of classes for that student is shown above the student identifier (e.g., “Stdnt 1” had three classes in the Winter '20 term and five classes in the Spring '20 term). For the classes in each term, same numbers or letters along the x-axis indicate the identical class during that term.

Table 4. Variation in engagement levels during the two comparison terms

	Winter '20pre-pandemic	Spring '20pandemic
Number of classes tracked	50	59
Percent of class meetings with no I+C engagement	6.0	11.9
Percent of class meetings with no I engagement	34.0	62.7
Percent of class meetings with no C engagement	8.0	15.3
Average percent of class time spent at A+P levels	38.9	39.1

For the same set of courses represented by the data of [Figure 7](#) and [Figure 8](#), [Table 5](#) examines the students' use of out-of-class learning times, averaged over all courses for the 16 students during the two comparison terms. The results show that the students, on average, used somewhat less variety in their study strategies (i.e., fewer types of entries), lower frequency of studying (i.e., fewer entries), and lower total time spent per class outside of class meetings. Note that the results from both terms show high variations (as shown by the large standard deviations), demonstrating the widely varying ways that students studied for each class. In addition, the average percent of out-of-class time spent on homework decreased slightly during the pandemic term which, when combined with the lower total time devoted to each class, means that time devoted to homework was substantially reduced. What is not clear, and perhaps more important, is whether these changes occurred due to lower levels of motivation or engagement with the classes because of the shift to online learning, or the courses were fundamentally changed to require less engagement because of the circumstances brought on by the pandemic, or perhaps by stress-induced distractions caused by the pandemic.

Table 5. Out-of-class study habits and patterns^d

	Winter '20pre-pandemic Average (Std Dev)	Spring '20pandemic Average (Std Dev)
Number of types of study strategies	2.9 (1.2)	2.4 (1.2)
Number of entries	17.7 (12.0)	15 (14.5)
Total time of all entries	1767 min. (1225)	1488 min. (1093)
Percent of time spent on homework	69.4% (28.3)	67.6% (30.4)

d. Values shown represent averages for all 16 students across all courses

5 Summary and Conclusions

We introduced here a new method for measuring the level of student engagement with their learning. This method was developed within an engineering-learning context but we believe it is applicable to most college-level disciplines. Furthermore, it is suitable with nearly all pedagogies currently in use in higher education. The method, called ALApp, is built in the form of mobile applications for the smartphone and is based on a well-researched educational framework designed to evaluate student engagement.

ALApp shares many common features with other modern methods of measurement of student engagement. These include data recording at or near the time of each learning event to eliminate recall bias, quantitative measures of both the quality and quantity of student engagement, accommodations for active learning or evidence-based instructional practices, and a reasonable level of training by the user for accurate measurements. ALApp differs from the other measurement methods in three ways: (1) Measurements are made by individual students rather than relying on observing the instructor or representative students; (2) measurements made at such a student level captures differences between students instead of averaging over a cluster of observed students; and (3) both in- and out-of-class learning are measured.

We described two studies that support the measurement accuracy of ALApp and its ability, through examination of the data, to discern both the type of pedagogy in use and the sometimes subtle differences between students in the same class. A third study demonstrated how ALApp is able to reveal the learning patterns and habits of engineering students in a single or a sequence of courses, and how these patterns shed light on the complex ways that students approach learning. A fourth study demonstrated the versatility of the ALApp and, rather than measuring student engagement at a cognitive level, it was adapted to measure individual students' contributions to a group project. The final study, which took place during the academic term that was forced by the coronavirus pandemic into an online mode, revealed the subtle yet substantial differences in students' learning patterns as a result.

This project had its start in 2014, when we decided to build a tool for gathering student learning engagement data during both in- and out-of-class times. Based on the features and functions that we required, and to address our concerns with ease-of-use and security around data and user-privacy, we believed that building a native mobile application was the best solution. In fact, we saw no other option. Admittedly, we could have gone in many directions with the app's design and architecture, but with our team's background and skillset, ALApp was created. Today, there are options beyond a native app for accomplishing the original goals of this project, but we still believe that a native app is the optimal solution.

While we do not anticipate that ALApp would be suitable for instructor use for assigning grades or participation points since it could be easily manipulated by the participant, we do foresee its use as a research tool for educational research, as an evaluative tool for measuring classroom practices and or instructional efficacy, and perhaps even as a student tool to evaluate a course and to provide course reviews to prospective students.

Acknowledgment

We gratefully acknowledge the funding support from the National Science Foundation. This material is based upon work supported by the National Science Foundation under Grants No. (1431879, 1561623). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. We would also like to recognize the valuable contributions of the many students who worked on this project in its earliest stages, including the app development and research support.

References

- [1] M. Prince, "Does active learning work? A review of the research," *Journal of Engineering Education*, vol. 93, no. 3, pp. 223–231, 2004.
- [2] M. Prince and R. M. Felder, "Inductive teaching and learning methods: definitions, comparisons, and research bases," *Journal of Engineering Education*, vol. 95, no. 2, pp. 123–138, 2006.
- [3] R. M. Felder and R. Brent, "Active Learning: An Introduction," *ASQ Higher Education Brief*, vol. 2, no. 4, 2009.
- [4] M. J. Drinkwater, K. E. Matthews, and J. Seiler, 2017.
- [5] Stains, "Anatomy of STEM teaching in North American universities," *Science*, vol. 359, pp. 1468–1470, 2018.
- [6] T. M. Andrews, M. J. Leonard, C. A. Colgrove, and S. T. Kalinowski, "Active learning not associated with student learning in a random sample of college biology courses," *CBE life sciences education*, vol. 10, no. 4, pp. 394–405, 2011.
- [7] "Discipline-Based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering," S. Singer, N. Nielsen, , and H. Schweingruber, Eds. Washington, DC: The National Academies Press, 2012. [Online]. Available: http://www.nap.edu/catalog.php?record_id=13362
- [8] N. R. Council, "The Mathematical Sciences in 2025," and others, Ed. National Academies Press, 2013.
- [9] C. B. of the Mathematical Sciences., "Active Learning in Post-Secondary Mathematics Education," 2016.
- [10] N. Bradburn, L. Rips, and S. Shevell, "Answering autobiographical questions: the impact of memory and inference on surveys," *Science*, vol. 236, pp. 157–161, 1987.
- [11] R. Tourangeau, "Remembering what happened: memory errors and survey reports," in *The Science of Self-Report: Implications for Research and Practice* and A. S. et al., Eds., 2000, pp. 29–48.
- [12] S. Shiffman, A. A. Stone, and M. R. Hufford, "Ecological momentary assessment," *Ann. Rev. Clin. Psychol*, vol. 4, pp. 1–32, 2008.
- [13] M. T. Hora and J. J. Ferrare, "Instructional systems of practice: A multidimensional analysis of math and science undergraduate course planning and classroom teaching," *Journal of the Learning Sciences*, vol. 22, no. 2, pp. 212–257, 2013.
- [14] M. T. Hora and B. A. Couch, "Assessing the presence of interactive teaching in postsecondary STEM classrooms," 2013. [Online]. Available: <http://www.wcer.wisc.edu/publications/workingPapers/papers.php>

- [15] M. K. Smith, F. Jones, S. L. Gilbert, and C. E. Wieman, "The Classroom Observation Protocol for Undergraduate STEM (COPUS): a New Instrument to Characterize University STEM Classroom Practices," *CBE-Life Sciences Education*, vol. 12, no. 4, pp. 618–627, 2013.
- [16] S. L. Eddy, M. Converse, and M. P. Wenderoth, "PORTAAL: A classroom observation tool assessing evidence-based teaching practices for active learning in large science, technology, engineering and mathematics classes," *CBE - Life Sciences Education*, vol. 14, pp. 1–16, 2015.
- [17] M. T. H. Chi, "Active-Constructive-Interactive: A conceptual framework for differentiating learning activities," *Topics in Cognitive Science*, vol. 1, pp. 73–105, 2009.
- [18] B. A. Fonseca and M. T. H. Chi, "Instruction based on self-explanation," in *Handbook of Research on Learning and Instruction*, R. Mayer and P. Alexander, Eds. Routledge, 2011.
- [19] D. S. Janzen, R. Chang, and J. Chen, "Development of mobile applications to study engineering students' patterns of learning," *IEEE*, 2019.