Paper ID #19780

What does an In-Class Meeting Entail? A Characterization and Assessment of Instructor Actions in an Active, Blended, and Collaborative Classroom

David Allen Evenhouse, Purdue University, West Lafayette

David Evenhouse is a Graduate Student and Research Assistant in the Purdue School of Engineering Education. He graduated from Calvin College in the Spring of 2015 with a B.S.E. concentrating in Mechanical Engineering. Experiences during his undergraduate years included a semester in Spain, taking classes at the Universidad de Oviedo and the Escuela Politécnica de Ingenieria de Gijón, as well as multiple internships in Manufacturing and Quality Engineering. His current work primarily investigates the effects of select emergent pedagogies upon student and instructor performance and experience at the collegiate level. Other interests include engineering ethics, engineering philosophy, and the intersecting concerns of engineering industry and higher academia.

Mr. Rohit R. Kandakatla, Purdue University, West Lafayette (College of Engineering)

Rohit Kandakatla is currently pursuing his Ph.D. in engineering education at Purdue University. He is the co-founder of Footsteps, a startup which working towards transforming engineering education in India. He completed B.Tech in Electronics and Communication Engineering from Manipal University and M.Tech in Embedded Systems from KG Reddy College of Engineering and Technology. He is currently serving as the President of Student Platform for Engineering Education Development (SPEED) and has conducted more than 30 workshops across Asia, Europe, and Latin America which were focused on engineering education development and action planning. He was the founder of Indian Student Forum (ISF) and has been awarded the IGIP SPEED Young Scientist Award for the year 2014 for his efforts. He has been instrumental in setting up Indo Universal collaboration for Engineering Education (IUCEE) students Chapters across 47 colleges across India impacting more than 5000 students since 2013 and has been awarded IUCEE Young leader for his contribution. He was elected as the Vice-President of the International Federation for Engineering Education Societies (IFEES) during World Engineering Education Forum which was held in Florence, Italy in 2015. His areas of interests include policy in higher education, integration of technology and entrepreneurship in engineering education, and service learning.

Mr. Nick A. Stites, Purdue University, West Lafayette (College of Engineering)

Nick A. Stites is pursuing a PhD in Engineering Education at Purdue University. His research interests include developing novel pedagogical methods to teach core engineering courses, leveraging technology to enhance learning experiences, and increasing the effectiveness of laboratory classes. Nick holds a BS and MS in Mechanical Engineering and has experience as an engineering practitioner and as an adjunct instructor at the community-college and research-university level.

Mr. Nimit Patel, McKinsey & Company

Nimit Patel is a Data Analyst with McKinsey & Company, at their Knowledge Center in Waltham. He completed his Master of Science in Industrial Engineering from Purdue University in December 2016. During this period, he also worked as a Data Analytics Research Assistant with the Engineering Education Department and as Vice President - Industry Committee, INFORMS at Purdue University. He completed his Bachelor of Technology in Mechanical Engineering from Sardar Vallabhbhai National Institute of Technology, India in May 2015.

Mr. Austin Zadoks, Purdue University, West Lafayette

Austin Zadoks is an Undergraduate in the School of Materials Engineering and a Research Assistant in the School of Engineering Education and Purdue University.

Ms. Taylor Prebel, Purdue University

Taylor Prebel is an undergraduate student in her senior year at Purdue University. Taylor is a research assistant in Dr. Jennifer DeBoer's lab within the School of Engineering Education.

Mr. Claudio Cesar Silva de Freitas, Purdue University

Claudio holds Bachelor's degree in Control Engineering at Higher Education Institute of Amazonia (2011), and he holds his Master's degree in Electrical Engineering from the State University of Campinas (2014) in Brazil. Currently, he is pursuing his Ph.D. at Purdue University in Engineering Education. He has experience as a visiting graduate researcher at the University of New Mexico (USA) and professional experience at K&A Wireless as a research associate in Albuquerque (USA). Additionally, he worked at Hitachi Automotive Systems America as an Intern in Research & Development in Detroit (USA) and Senior Product Engineer at Fiat Chrysler Automobiles in Brazil. Claudio's research focuses on understanding cross-cultural factors impacting student learning process as well as design curriculum based on engineering classes to empower engineering skills through hands-on activities and blended learning environment. He is also interested in design methods which are based on the technology tools usage to stimulate engineering thinking for diverse students.

Prof. Charles Morton Krousgrill, Purdue University, West Lafayette (College of Engineering)

Charles M. Krousgrill is a Professor in the School of Mechanical Engineering at Purdue University and is affiliated with the Ray W. Herrick Laboratories at the same institution. He received his B.S.M.E. from Purdue University and received his M.S. and Ph.D. degrees in Applied Mechanics from Caltech. Dr. Krousgrill's current research interests include the vibration, nonlinear dynamics, friction-induced oscillations, gear rattle vibrations, dynamics of clutch and brake systems and damage detection in rotor systems. Dr. Krousgrill is a member of the American Society for Engineering Education (ASEE). He has received the H.L. Solberg Teaching Award (Purdue ME) seven times, A.A. Potter Teaching Award (Purdue Engineering) three times, the Charles B. Murphy Teaching Award (Purdue University), Purdue's Help Students Learn Award, the Special Boilermaker Award (given here for contributions to undergraduate education) and is the 2011 recipient of the ASEE Mechanics Division's Archie Higdon Distinguished Educator Award.

Dr. Edward J. Berger, Purdue University, West Lafayette (College of Engineering)

Edward Berger is an Associate Professor of Engineering Education and Mechanical Engineering at Purdue University, joining Purdue in August 2014. He has been teaching mechanics for nearly 20 years, and has worked extensively on the integration and assessment of specific technology interventions in mechanics classes. He was one of the co-leaders in 2013-2014 of the ASEE Virtual Community of Practice (VCP) for mechanics educators across the country. His current research focuses on student problem-solving processes and use of worked examples, change models and evidence-based teaching practices in engineering curricula, and the role of non-cognitive and affective factors in student academic outcomes and overall success.

Prof. Jeffrey F. Rhoads, Purdue University, West Lafayette (College of Engineering)

Jeffrey F. Rhoads is a Professor in the School of Mechanical Engineering at Purdue University and is affiliated with both the Birck Nanotechnology Center and Ray W. Herrick Laboratories at the same institution. He received his B.S., M.S., and Ph.D. degrees, each in mechanical engineering, from Michigan State University in 2002, 2004, and 2007, respectively. Dr. Rhoads' current research interests include the predictive design, analysis, and implementation of resonant micro/nanoelectromechanical systems (MEMS/NEMS) for use in chemical and biological sensing, electromechanical signal processing, and computing; the dynamics of parametrically-excited systems and coupled oscillators; the thermomechanics of energetic materials; additive manufacturing; and mechanics education. Dr. Rhoads is a Member of the American Society for Engineering Education (ASEE) and a Fellow of the American Society of Mechanical Engineers (ASME), where he serves on the Design Engineering Division's Technical Committee on Vibration and Sound. Dr. Rhoads is a recipient of numerous research and teaching awards, including the National Science Foundation's Faculty Early Career Development (CAREER) Award; the Purdue University School of Mechanical Engineering's Harry L. Solberg Best Teacher Award (twice), Robert W. Fox Outstanding Instructor Award, and B.F.S. Schaefer Outstanding Young Faculty Scholar Award; the

ASEE Mechanics Division's Ferdinand P. Beer and E. Russell Johnston, Jr. Outstanding New Mechanics Educator Award; and the ASME C. D. Mote Jr., Early Career Award. In 2014, Dr. Rhoads was included in ASEE Prism Magazine's 20 Under 40.

Prof. Jennifer DeBoer, Purdue University, West Lafayette (College of Engineering)

Jennifer DeBoer is currently Assistant Professor of Engineering Education at Purdue University. Her research focuses on international education systems, individual and social development, technology use and STEM learning, and educational environments for diverse learners.

What does an In-Class Meeting Entail? A Characterization and Assessment of Instructor Actions in an Active, Blended, and Collaborative Classroom

Abstract

Although STEM education researchers recognize the need to incorporate a variety of in-class instructional approaches in undergraduate classrooms, few empirical benchmarks exist for the proportion of time instructors dedicate to each approach or activity. Over the past few years, our team has made a concerted effort to implement and disseminate an innovative, undergraduate mechanics learning environment known as Freeform; a pedagogical system integrating active, blended, and collaborative (ABC) instructional elements. Our work has been complicated by the fact that very few previous studies describe, in sufficient detail, what a typical ABC classroom experience looks like from the instructor's perspective. As a result, adopters of ABC approaches such as Freeform do not have a template describing what activities are typically involved in the day-to-day use of an ABC system. To address this knowledge gap, and to inform future implementations of the Freeform environment, this paper defines a pedagogical benchmark quantifying what happens during a typical Freeform class session. This study focuses specifically on the actions of the instructor in order to answer the question: as part of the Freeform environment, what specific actions do experienced instructors take during in-person class meetings?

Since their inception, Freeform dynamics courses have seen a drastic drop in the rate at which students are earning a D grade, failing, or withdrawing from the course (the so-called DFW rate). On-going work examines the actions and behaviors of students and faculty, in addition to a variety of other variables, as a way of understanding the drastic improvement in DFW rate. For this study, each relevant in-class meeting (i.e., not including cancelled classes, those involving exams, etc.) taught by two experienced Freeform instructors was video recorded over the course of the Spring 2016 semester and subsequently analyzed with respect to instructor actions. Continuous video coding analysis was used to capture how much time these two instructors dedicated to various instructional activities such as assessments, traditional lecturing, demonstrations, and writing notes or examples in real-time. The analysis provides a clearer picture of how and when these two veteran instructors employed active, blended, and collaborative approaches in their classrooms.

The implications of the analysis are two-fold. First, we strive to improve Freeform instruction at our institution by providing instructors with an opportunity to reflect on their instructional practices in the context of rigorously-derived, quantitative summaries of real-time teaching actions. Second, we establish a benchmark characterization of ABC instructional elements in engineering mechanics, and discuss its potential implications for undergraduate STEM education at large. Through the evidence developed in this study about specific instructor actions in Freeform classrooms we expect to inform and encourage the implementation of ABC pedagogical practices by other faculty in other courses and at other institutions, as well as to provide an assessment framework suitable for the analysis of STEM in-class instructional practices.

Introduction

On-going calls for a transformation of engineering education[1] recommend a pedagogical overhaul. Engineering educators must transition from a teacher-centered to a student-centered learning environment incorporating the use of active learning techniques in the classroom. Educational research also suggests that students benefit from blended courses which mix online and in-class resources[2], and courses that facilitate collaborative learning[3]. While Active, Blended, and Collaborative (ABC) approaches demonstrate a positive impact on student performance, few educators have intentionally integrated all three into the design of a single course.

Freeform is an innovative, ABC learning environment that was first applied to sophomore-level dynamics classes in the School of Mechanical Engineering at Purdue University[4]. It employs a student-centered approach incorporating active and collaborative strategies with blended resources to enhance instruction in both conceptual knowledge and problem solving skills. Since the introduction of the Freeform environment to these dynamics courses, the rate at which students receive failing grades or withdraw from the course (the so-called DFW rate) has declined dramatically[5]. This success has given rise to various research projects centered around understanding, improving, and disseminating the Freeform environment. In continuing this work, our research team has begun to bring the Freeform environment to other educational instititions, but its implementation has proven challenging due, in part, to the lack of literature on what ABC classrooms should look like in practice.

To address this gap in our understanding of Freeform's ABC environment, this paper presents an initial video coding analysis of the Freeform classroom, characterizing the pedagogical practices used by two experienced dynamics instructors who originally helped to develop Freeform. This analysis categorizes the Freeform classroom based on the time allotted by the instructors to various learning activities, with a focus on identifying and quantifying ABC instructional practices and laying the groundwork for future study. The results of this paper have assisted us in professional development within the Freeform environment, and serve as our first glimpse into what an in-class meeting actually entails within this innovative ABC framework.

Background

Active, blended, and collaborative learning

In an Active learning environment, instructors intentionally engage students in ways which require action of the part of the students themselves, with the goal of improving student learning outcomes. A meta-analysis of 225 studies reported that students who were taught in an active learning environment had their average examination scores improve by 6% over those students in traditional classrooms. Likewise, students in a traditional classrooms were 1.5 times more likely to fail compared to those in classes which employed active learning[6]. Active learning helps students better retain the concepts learned in class, develops thinking skills, and motivates them for further classwork[7].

Blended learning integrates the use of online, digital media both inside and outside the classroom. Students in blended environments have performed demonstrably better as compared those who learned the same materials in instructional evironments lacking online elements[2]. Additonally, the integration of multimedia into lectures has been shown to improve students' attentiveness[8], as well as their engagement and participation[9].

Studies have also suggested that collaborative learning demonstrates both cognitive and non-cognitive benefits for students[10]. Students engaged in collaborative learning environments demonstrated better knowledge retention, greater persistence, and improved class attendance when compared to students in more traditional lecturing environments[11,12,13]. Students were more motivated when working in groups compared to working alone[14], and groupwork enabled them to recognize gaps in their knowledge, to synthesize and communicate ideas more efficiently, and to advance their conceptual understanding[15].

Understanding and evaluating classroom instruction

Classroom observations reveal the detailed structure and subtle nuances of teaching practice[16], and can elicit insight from the behaviors of instructors and students within the classroom[17]. Observation has been used extensively to assess the quality of instruction[18] as well as to catalogue overt teaching and learning behaviors demonstrated in the classroom[19]. Classroom Observation Protocols (COP's) focus more specifically on these instructional practices[20] and catalogue specific actions taken by the instructor at specific times. Feedback based upon classroom observations can inform improvements to teaching practice when provided to instructors[21]. Classroom observations can also be triangulated with other data such as student grades or pre/post test surveys to identify specific teaching practices that lead to improved student course outcomes[22]. Study into the impact and use of pedagogical methods help practitioners to effectively utilize instructional practices in their own classrooms.

Methods

Data collection and sampling

Video data was collected from two Freeform classrooms, each taught by a veteran instructor during the Spring semester of 2016. It was originally intended that all regularly scheduled, relevant class meetings (not including exams and review sessions) of the two sections would be recorded over the course of the full semester. Unfortunately, some complications arose. A number of videos were cut off a few minutes early by the recorders, artificially limiting our ability to observe activities at the end of class. Also, some of these videos became corrupted or otherwise inaccessible, resulting in a total of 72 videos, approximately 54 hours of video data, for subsequent analysis. Both participating instructors were involved in the creation of the Freeform environment and had taught the course on multiple previous occasions. Likewise, both had demonstrated similar success within the Freeform environment, seeing the drastic drop in DFW rate mentioned earlier, as well as receiving overwhelmingly positive reviews from the students in their courses, as well as their fellow faculty members.

With this in mind, the choice to record the instructional activities of these two professors was intentional. We wish to compare between the two instructors, analyzing similarities and differences to inform our understanding, and future implementations of, the Freeform learning environment. Although analyzing lectures by instructors who are less experienced with Freeform will contribute to our understanding of what variations Freeform can take on in practice, the focus of this initial study is to look at Freeform according to the actions of two original Freeform developers, situated within its original context. In this way, this work characterizes an index case (the first or primary instance) of Freeform's application, with which we hope to inform future comparative analyses and possible pedagogical recommendations.

Development of a coding scheme

There exist several classroom observation protocols addressing individual aspects of ABC classroom instruction, but none provided us with the comprehensive accounting of ABC learning activities that we required in this study. For this reason, we decided to develop our own observation protocol to act as a video coding scheme, tailored to the needs of our research project and the software package available to us.

To code this data for subsequent analysis, our research team first reviewed existing classroom observation protocols which seek to characterize the pedagogical practices of instructors. We largely limited our search to literature dealing with higher education and engineering, though some documents from beyond this body of literature were also considered. Some popular tools such as the Reformed Teaching Observation Protocol (RTOP)[23] proved to be too broad, with their openness to description and interpretation making their use rather infeasible. Considering the scale of the data we had on hand, we needed a more concise and efficient measurement tool. Other more scoped protocols, such as the Classroom Observation Protocol for Undergraduate STEM (COPUS)[24] and the protocol for analyzing student active learning[25] provided a more directly applicable means of coding points of interest, but failed to encompass every aspect of the Freeform environment that we wished to capture. Because of this, we decided to develop our own coding scheme which would align the class environment, the research questions, and the data analysis methods with one another, creating a tool that could be uniquely valid and reliable for the evaluation of prospective ABC learning environments. For more on this development process, please refer to the authors' companion publication[26].

Data analysis

The core of our coding scheme is a list of nine instructional activities which, for the purpose of this study, are assumed to be both mutually exclusive and all-inclusive (due to the catch-all "Other" term). This means that every moment in a given class period will be coded as one, and only one, of these instructional activities. These activities are further clarified using two other sets of codes describing which instructional practices (A, B, C, or P to indicate Passive instruction) are being employed by the instructor, and how many (None, Some, or All) of the students the instructor is intending to draw into Active learning. These codes and their relationships to one another are further defined in Table 1 below. In the Characterization and Degree of Engagement columns, we also lay out what relationships, if any, have been built into the coding scheme in order to facilitate greater reliability between the coders. The word "Forced" indicates codes that are forcibly activated when the given event is selected. Likewise, "Excluded" indicates codes that are forced off when the event is selected. "Optional" codes may be activated as needed. The codes are arranged in order from "least potential for active learning" to "most potential for active learning" with the exception of the Other/Administrative code, which was appended to the coding scheme to capture activities that are not necessarily instructive in nature.

Table 1: Coding scheme definitions and relations

Event	Description	Characterization	Degree of Engagement
Conceptual Talking / Lecturing	The instructor is talking directly to the students; a monologue or purely didactic form of instruction. The content is purely conceptual; theoretical knowledge is delivered to the students.	Forced: P Optional: B Excluded: A, C	Forced: None Excluded: Some, All
Problem Solving Talking / Lecturing	The instructor is talking directly to the students; a monologue or purely didactic form of instruction. The content is generally a verbal discussion of a problem-solving activity, the reading out of the problem statement, etc.	Forced: P Optional: B Excluded: A, C	Forced: None Excluded: Some, All
Conceptual Real Time Writing	The instructor is explaining some concept (e.g. Free body diagrams, equation derivations, etc.) by writing on the board.	Forced: P Optional: B Excluded: A, C	Forced: None Excluded: Some, All
Problem Solving Real Time Writing	The instructor is solving some example problem on the board, demonstrating the application of equations, or enumerating a problem-solving process.	Forced: P Optional: B Excluded: A, C	Forced: None Excluded: Some, All
Questions	This categorization includes both when the students ask a question of the instructor, and when the instructor asks a question of the students. In the second case, this specifically refers to instances where the instructor is not expecting, nor requiring, every student to respond.	Forced: A Optional: B, C Excluded: P	Forced: Some Excluded: None, All
Graded Assessment	This categorization includes instances where the instructor asks a question or series of questions of the students that all the students are expected to answer for a grade. For example, quizzes,	Forced: A Optional: B, C Excluded: P	Forced: All Excluded: None, Some

	exams, extra credit in-class problems, etc.		
Ungraded Assessment	This categorization includes where the instructor asks a question or series of questions of the students that all the students are expected to answer but their responses are not graded. For example, feedback forms, problems or examples given to solve in class.	Forced: A Optional: B, C Excluded: P	Forced: All Excluded: None, Some
Demonstration	This categorization includes any kind of demonstration that uses some accessory, digital resource, or real-world object, and is intended to ease the understanding or visualization of a phenomenon or a concept. This includes the use of videos and simulations.	Optional: A, B, C, P	Optional: None, Some, All
Other	This categorization includes any other events which may transpire that do not fit the above categories. For example, administrative work, logistics, waiting when the instructor is late to class, etc.	Optional: A, B, C, P	Optional: None, Some, All

A coding window reflecting this scheme was built in StudioCode, a video-analysis software package used historically in sports analysis, but which is seeing increasing application in education research. It uses a visual interface to facilitate video coding in real-time, preserving data as instances of time which are marked and saved on a timeline. It also allows for numerical computation in a worksheet-style statistical window. This enables our team to code class video continuously, capturing every second of instructional time available for analysis. Continuous coding sets our work apart from other prominent observation protocols, such as COPUS, which call on the observer to code instruction in two minute increments. In practice, coders would go through video at 1.5x1 to 2x1 speed, waiting three seconds between overt changes in instructional activity before switching between codes. Selecting a new code would automatically switch off the previously active code, resulting in no time lost between coded instances. All of this combines to produce a coding method that is both time efficient, and accurate to an extent that is not typically possible in classroom observation work. Figure 1 shows a simplified representation of our video coding scheme, with arrows indicating activation relationships between codes. More on StudioCode and how its features were considered in the development of our coding scheme is included in our companion publication[26].

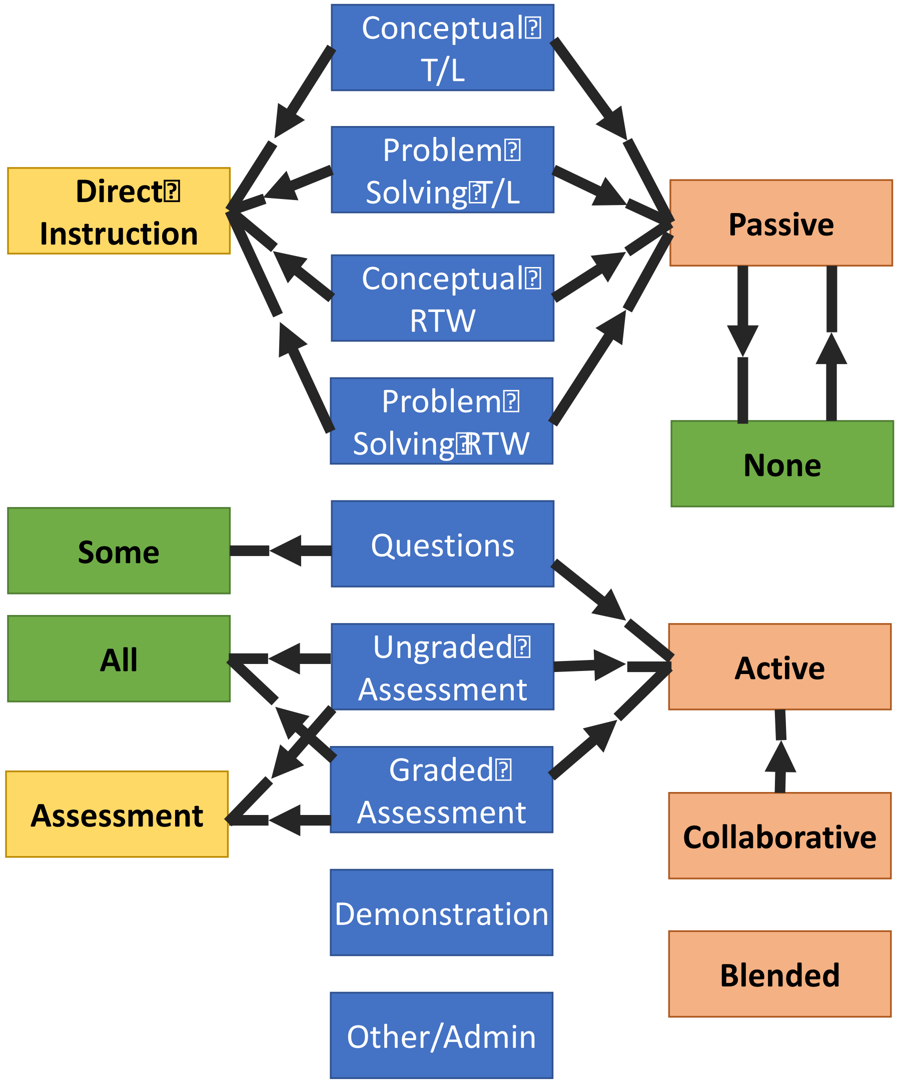

Figure 1: A representation of the coding scheme for Freeform video analysis. Arrows indicate a forced activation (where activating one code automatically activates another). Activation and deactivation relationships were included in the coding structure to increase coding reliability.

As the video analysis was limited to the overt actions of the instructor, instructional activities were characterized as Active, Blended, and Collaborative based upon what was directly facilitated, not by what the students did in response. Thus, instances where the instructor required some sort of active response from the students were coded as Active. Likewise, instances when the instructor made direct use of, or reference to, the online resources for the course were coded as Blended. Finally, instances when the instructor directly facilitated collaboration between students were coded as Collaborative.

Three researchers took part in the video coding, and Cohen's Kappa was employed at four regular intervals in the coding process to monitor interrater reliability (IRR) while controlling for chance agreement[27, 28]. Kappa values consistently fell just above 0.6, indicating moderate to substantial agreement based on the Landis and Koch Kappa benchmark, a level of reliability we deemed appropriate for our work[29]. After coding, time-duration data describing each class meeting were exported from StudioCode for analysis. The analysis resulted in an empirically-generated numerical description of how classes were conducted in the Freeform environment during the Spring of 2016.

Benchmarking and statistical analysis

Time-duration data for this analysis was processed to find a benchmark for what types of instructional activities occur in a typical Freeform classroom when a developer of Freeform is the instructor. Benchmarking is a standard practice in both engineering industry, and engineering education, with averages being the most broadly employed method for the creation of benchmark data[30, 31].

Descriptive statistics such as variance and standard deviation were calculated assuming they represented a full population, rather than a sampled subset. This is because the courses analyzed represented the full set of classes for an entire semester, encompassing one full implementation of Freeform.

Results

Desctiptive statistics

To better understand the characteristics of our data, we analyzed the descriptive statistics of the proportions of class time spent on each of the nine pedagogical categories. By looking at how these proportions changed across multiple classes, we gain insights into how consistent instruction is within the Freeform environment for the two veteran instructors. Table 2 lists the descriptive statistics for the nine coded instructional activities, and Figures 2-5 illustrate examples of the types of distributions encountered.

Most of the instructional activities demonstrated minimum proportions that run up against the lower bound of the domain, in this case 0% of class time. This implies that the type of instruction used on a given day depends on the content for that day, and no single activity universally dominates class time. Some activities, such as graded assessment and demonstrations, tended to be absent from class periods except for a few limited instances of extended use, as exemplified in Figure 2. The large bars on the left side of the histograms denote the large numer of class periods without graded assessments.

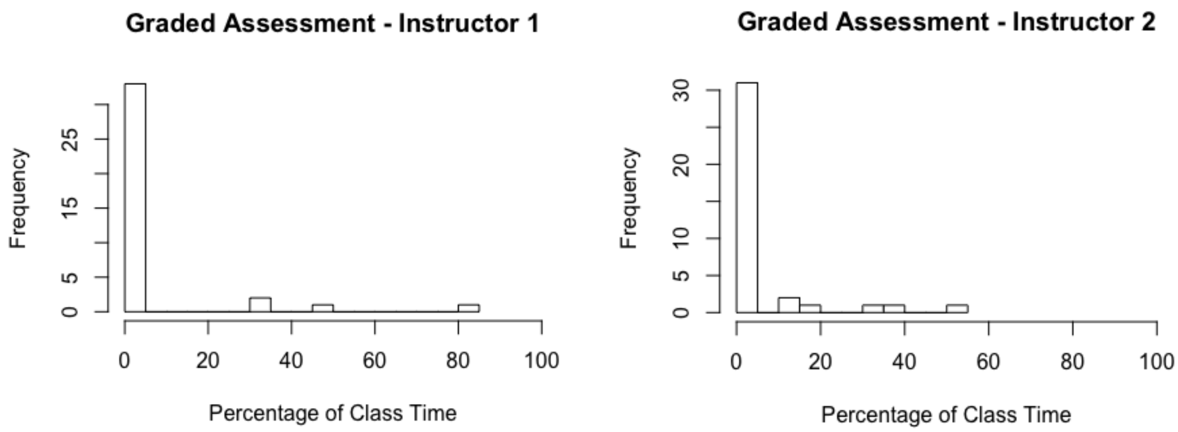

Figure 2: Histograms displaying the instructional use of Graded Assessment over the course of the semester. The heavy skew to the right indicates that most days did not include this activity, inflating the standard deviation and limiting analysis.

Other activities, particularly those used more often in class, displayed much more normal distributions. Activites used to teach problem solving knowledge and skills, for instance, were present in almost every class period. Problem Solving RTW consumed the largest proportion of class time on average, and its distribution of proportions, Figure 3, demonstrated far less skewness.

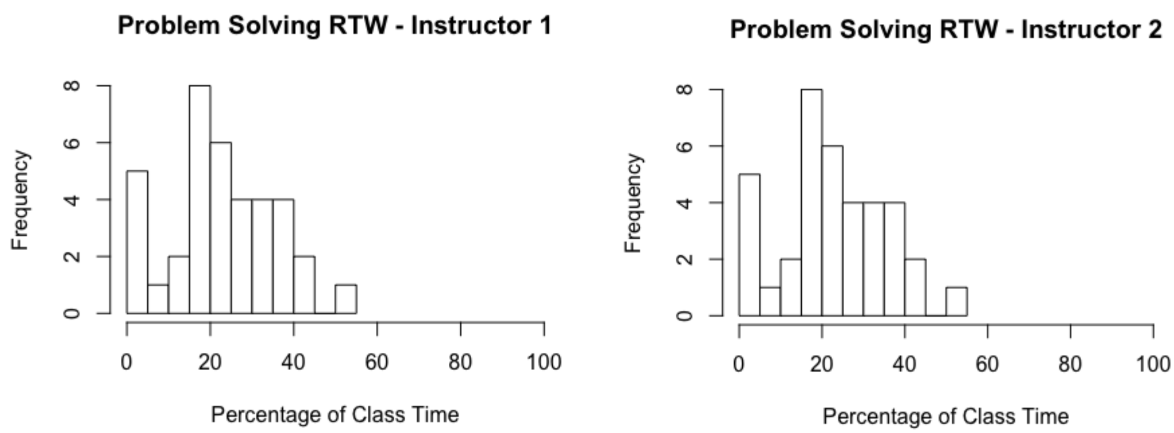

Figure 3: Histograms displaying the instructional use of Real Time Writing to teach students Problem Solving knowledge and skills. Note that the data spreads out over the distribution, rather than clustering at the y-axis.

Finally, some activities were used often, but in a more limited capacity. Questions, for instance, were employed in almost every class period and often accounted for about 20% of the class time. Other activities, such as those related to instruction on conceptual topics as well as administrative and other tasks, were used as required. This was shown in the case of Conceptual T/L in Figure 5. These distributions contrast with those in Figure 4 both due to the concentration of classes near zero, as well as the broader range of the distribution overall.

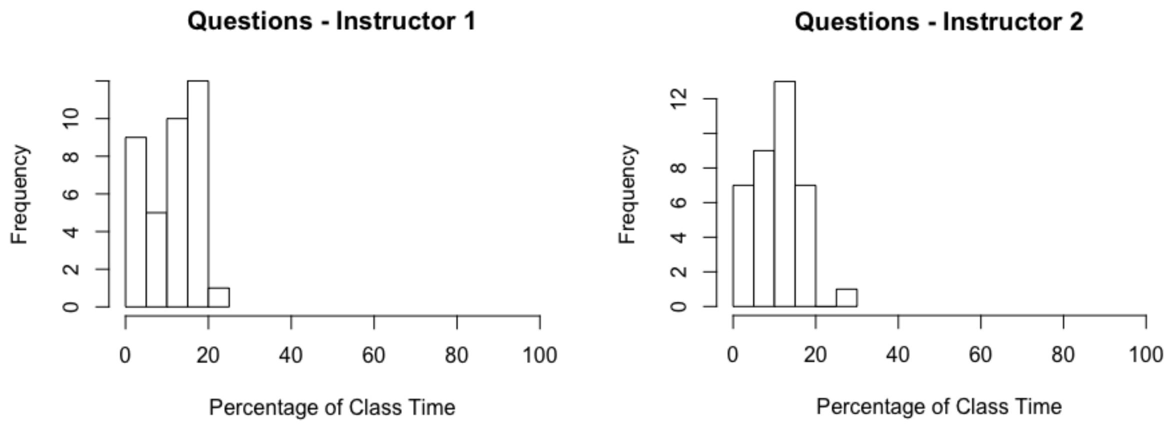

Figure 4: Histograms showing the instructional use of Questions. Questions tended to be present for a consistent period of time during each class.

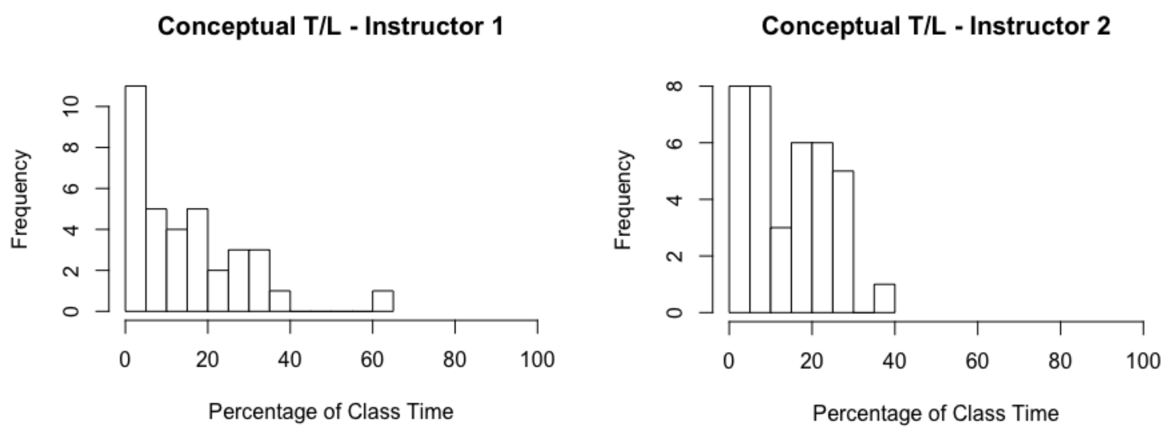

Figure 5: Histograms showing the instructional use of Talking and Lecturing to communicate Conceptual knowledge. Lecturing on concepts tended to be used for small periods of time (such as when deriving an equation or explaining a theorem).

Test of independence between the instructors

We employed the chi-squared test of independence to test for dependency between the instructor and the instructional activities. The total number of seconds dedicated by each instructor, to each activity, over the course of the semester constituted the unit of analysis. Our null hypothesis of independence was rejected ($p < 0.001$), but the effect size of the dependency between instructor and instructional activity was small (w index = 0.189) according to Cohen's published guidelines[32]. The rejection of the null hypothesis indicates that each instructor displayed their own instructional style over the course of the semester, but the small effect size indicates that these differences in instruction were minimal overall.

Test of independence for each instructional activity

To further investigate how the two instructors differed in the time they allotted to each instructional activity, we conducted a series of post-hoc, pairwise chi-squared tests, examining the time associated with each instructional activity against the time allotted to all of the other categories[33]. The null hypothesis of indepence was rejected with a *p*-value of less than 0.001 for every instructional activity except Demonstration. This was not unexpected given the large sample size of our analysis (Instructor 1 had 92,709 observations, or seconds of coded video, and Instructor 2 had 103,038 observations) and the known sensitivity of the chi-squared test to sample size[33]. However, the effect sizes, see Table 2, associated with each of these tests were once again small. In fact, most of the categories had effect sizes smaller than the 0.10 minimum threshold that Cohen set for *small* effects[32], therefore rendering the majority of differences in time spent on a given instructional activity *very* small when comparing the two veteran instructors. Though the instructors differed, these differences were slight when aggregating instructional time over an entire semester.

Table 2: Instructor usage of coded activities with statistical strength of instructional differences

Instructional Activity	Instructor 1 Average Time (SD)	Instructor 2 Average Time (SD)	P-Value	Effect Size
Conceptual T/L	15.8% (2.8%)	14.1% (1.8%)	**<0.001**	0.023
Problem Solving T/L	11.8% (1.3%)	13.0% (1.2%)	**<0.001**	0.017
Conceptual RTW	1.6% (4.6%)	4.0% (3.2%)	**<0.001**	0.072
Problem Solving RTW	24.9% (1.6%)	37.0% (1%)	**<0.001**	**0.130**
Questions	12.3% (1.5%)	11.3% (1.4%)	**<0.001**	0.015
Ungraded Assessment	11.5% (4%)	6.9% (5.8%)	**<0.001**	0.080
Graded Assessment	6.0% (9.6%)	4.2% (7.4%)	**<0.001**	0.041
Demonstration	2.2% (8.4%)	2.3% (4%)	0.148	0.003
Other/Admin	13.9% (1.9%)	7.2% (2.5%)	**<0.001**	**0.111**

The largest difference in instructional style is evidenced in the time used for Problem Solving RTW activities. This difference seems to show the slightly contrasting instructional styles demonstrated by these two professors. We can observe that Instructor 2 tended to be more didactic in their teaching, relying less on collaborative learning activities than Instructor 1. Instructor 1, who spent considerably less time in the RTW and Lecturing activities, also appears to spend more time on activities coded as intentionally facilitating Active learning such as Assessments and Questions. While the Freeform instructors employed both Passive and Active learning strategies in the classroom, they did so to differing degrees.

Creating a benchmark for instructional activities in Freeform

While the differences in instructional style revealed in this data certainly complicate and enrich our understanding of the Freeform learning environment, they are also not unexpected. Teaching style would naturally differ between instructors, and even between class periods. To this extent, the question then becomes not one of "can there be variation within the Freeform environment" but rather of how much variation can be expected, and what instructional activites would we expect to be present in such an environment. Because the differences in instructional activities over the course of a semester between the two veteran instructors in the study were small (or very small), we averaged the proportion of time spent on each instructional activity across the two instructors. This set of average proportions of class time acts as a set of loose benchmarks, seen in Table 3, that describe the instructional content of a typical Freeform classroom for these veteran instructors. The minimum, maximum, and standard deviation for each category also illustrate how these proportions can stand to change throughout the semester.

Table 3: Benchmark values for use of instructional activities in the Freeform classroom

Instructional Activities	Average Time (SD)	Min | Max
Conceptual T/L	14.9% (11.5%)	0.0% | 61%
Problem Solving T/L	12.4% (5.9%)	0.0% | 30%
Conceptual RTW	2.8% (3.7%)	0.0% | 15%
Problem Solving RTW	31.3% (14.5%)	0.0% | 60%
Questions	11.8% (5.6%)	0.0% |28%
Ungraded Assessment	9.1% (13.8%)	0.0% | 42%
Graded Assessment	5.1% (14.8%)	0.0% | 85%
Demonstration	2.2% (4.8%)	0.0% | 31%
Other/Admin	10.4% (8.4%)	0.0% | 45%

Again, the bulk of the time in class (34.1%) was spent in Conceptual or Problem Solving Real Time Writing activities, denoting teaching which involves the professor writing on the board, or on a projected image, in real time during instruction. Another large portion of the class period (27.3%) was taken up by lecturing (Conceptual or Problem Solving T/L), with the professor verbally walking students through the lesson. This being said, perhaps the most surprising result of this analysis was the 10.4% of class time (on average) devoted to Other and Administrative activities. From Table 2, we know that Instructor 1 spent considerably more time on administrative tasks than Instructor 2. However, it could be argued that the administrative time of both was too high. Using the combined average from Table 3, instructors spent about five minutes of each class period taking care of administrative and other tasks. While this may not seem like much when considering one instance, over the course of a full semester this added up to about 3.5 hours worth of class time spent on administrative concerns. This is especially concerning in the case of instructor 1, who spent almost twice as much time on administrative tasks as compared to instructor 2.

Assessments and questions, which all act to generate feedback for the instructors, on average encompassed another 26% of the class period. These activities, combined with a very limited number of active demonstrations, comprised the portion of the class period coded as directly facilitating Active learning on the part of the students (a total of 26.8%). Of these activities, just under half (or about 12.8% of the total class period) were also Collaborative in nature. Finally, only 2% of a given class period involved any kind of blended activity.

Limitations of the analysis

First, but perhaps most importantly, the proportions of time spent on some of these activities may be skewed due to the recording methods used during data collection. As mentioned previously, the starting and final few minutes of some class periods were cut off, artificially reducing the amount of time we would likely code as Ungraded Assessment, Graded Assessment, and Other/Admin activities which were present at the start and end of class. Future work will have to address this missing data, and future recorders will have to be warned not to cut off video at the end of class, even if they perceive there to be nothing interesting going on.

The difference between Active facilitation on the part of the instructor, and active learning on the part of the students, also merits further discussion. As mentioned previously, these videos captured the actions of the instructors, not the students. We could not confirm whether or not students were actually engaging in their learning actively; we could only describe the overt actions of the instructor. Thus, while we only felt comfortable coding 26.8% of instruction as directly facilitating active learning, there is arguably much more of the instruction in the course that could be characterized as such. For example, the meta-analysis of active learning studies conducted by Freeman et al.[6] includes anything beyond traditional lecturing as an active learning oriented instructional practice. If we also coded everything beyond lecturing and administrative activities as Active, we would characterize an average of 62.3% of class time as containing active learning.

Limitations extend to the use of Blended and Collaborative learning activities as well. Over the semester 12.8% of total class time was coded as collaborative, which comprises about half of the time that had been coded as active overall. Though most of the collaborative learning that occurs in the classroom could be captured using this framework, it cannot capture collaboration that occurs in other academic spaces. One key resource provided to students in the Freeform environment is the course blog, an online information hub which not only provides students with individual learning resources, but also with a forum to interact online. Threaded discussions are created for each homework question in the course, and the students are encouraged to collaborate with one another throughout the semester. Likewise, the other blended resources provided by the learning environment simply cannot be captured using classroom observations. Resources that help with homework or with exam studies simply are not present, leading to a very low percentage of class time being devoted to blended activities. For example, only 2% of class time observed over the course of the semester was coded as incorporating blended aspects of the course. However, as much of this time took the form of reminders to students regarding the resources available to them, this limited amount of time could still have proven important to the student experience.

Discussion

Considering the care we took to build reliability and validity[25] into our coding scheme, as discussed in our companion publication[26], we feel confident that the data that we have is representative of what occurs during a typical semester’s worth of in-class meetings within the Freeform environment. As we see in the results, two highly successful Freeform instructors can employ different instructional strategies. Considering the differences in instructional style represented here, the team must now begin to dig deeper as we seek to characterize what Freeform is, how it can vary across instructors and contexts, and what can be considered essential to its faithful implementation. Answering these questions would help to put an empirical bound around how we define implementation fidelity in the Freeform environment, and enable greater transparency between the research team and future Freeform instructors. Literature on instructional change indicates that professors naturally try to integrate newly adopted pedagogical innovations with their own personal instructional style[34]. The ability to talk transparently with adopting instructors about what aspects of Freeform are essential to implementation fidelity, or what instructional elements have been useful in the past, could streamline future implementations of this innovative curricular framework.

The generalized view of Freeform according to the benchmarked actions of two experienced instructors has already proven useful. For example, in reviewing our findings with the participating instructors, they appeared shocked to find that just over 10% of their average class period was taken up by administrative tasks. Each instructor took this as an area for immediate improvement, working to streamline logistics at the start of each class period. This information also played a role in our adoption of a third-party software package for submitting, grading, and returning homework assignments online. Digitizing this process relieved instructors from having to return homework in-class, cutting out several minutes worth of in-class administrative work per week.

Future work

More closely examining how these instructors employed each of the coded activities could serve to better inform our understanding of instruction in the Freeform environment. The nature of the dataset that we are working with allows us to examine not only how much of each activity the instructors employed but also where, when, and for how long each of these activities were present. Figure 6 shows a timeline representation of the data taken from a single coded class period. Names of codes are on the left side of the timeline, and the blocks in each row represent instances of time during which the given code was selected. For example, looking at the rows titled "Questions" in our timelines, we can see that over the course of the class period, the instructor engaged students in a large number of short questions. These questions may have acted to break up the RTW and Lecture style instruction, presumably drawing students into the content and promoting active engagement over the course of the class period.

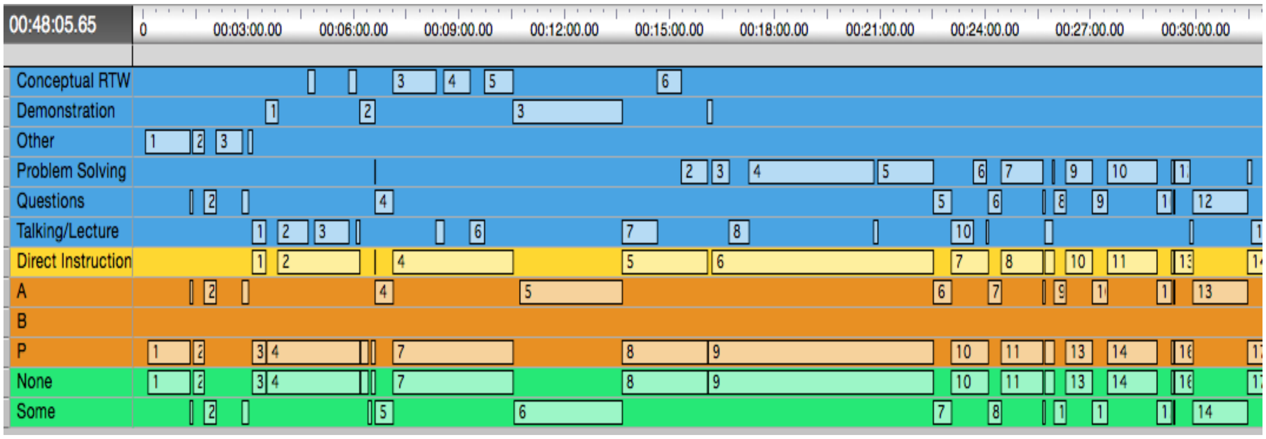

Figure 6: Example timeline representing the data generated from a single coded video. In StudioCode, the rows are color-coded to facilitate ease of interpretation.

Our future approach to the analysis of this unique data set will drastically expand upon the work that has been done here. Observing how these two instructors employ short, pointed questions to break up periods of extended direct instruction informs us of one possibility for future analysis. A statistical evaluation that not only takes into account the time-duration of coded instances, but also their frequency and number, could drastically change how we see and interpret the benchmark values that we have generated thus far. Likewise, evaluating timelines in order to identify different types of class periods (such as class meetings that are dominated by demonstrations, or graded assessments) and examining them each in turn would provide us with a much clearer picture of how instructors utilize these activities than can be garnered from a semester-wide average. The way in which an activity is employed could be more informative to us regarding implementation fidelity than simply comparing averages to benchmark values. Future work will take better advantage of the timelines generated by StudioCode to describe instructional style on a longitudinal basis, rather than in bulk.

Additionally, combining classroom observation data with data taken from other sources such as surveys and blog analytics could begin to paint a picture of the full ABC learning experience cultivated by the Freeform environment. This kind of supplementary data is important; class instruction is only one small part of what an ABC learning environment offers students who engage with it. Countering the limitations of the video coding data through triangulation with other types of data and analysis will be a critical next step as we continue to characterize the Freeform environment and inform our broader understanding of ABC instruction.

Conclusion

After analyzing a full semester's worth of video data, we have produced an empirical description of what the Freeform learning environment looks like when taught by two highly successful instructors. Though differences in instructional style were present, much of the variation we saw within this ABC framework was very small, and the benchmark values laid out here have already proven helpful in facilitating professional development and informing our understanding of the environment as a whole.

There is still much work that can be done. The descriptive benchmark established by this study could be further informed by a more intentional, longitudinal analysis of when, where, and why these two instructors employed the instructional activities evidenced here. Likewise, combining this analysis with other data on student actions and activities can serve to provide us with a more holistic perspective on what this innovative ABC learning environment looks like in practice. To the best of our present knowledge, this study represents the first time that a full semester's worth of continuously coded, time-duration data has been used to empirically describe the instructional practices of engineering faculty. We are excited about the potential uses for this method in our work, as well as the opportunities it may afford in other applications for Engineering Education and beyond.

Acknowledgments

The analysis detailed in the paper is based upon work supported by the National Science Foundation, Division of Undergraduate Education under Grant No. 1525671. 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.

References

1. Jamieson, L. H. & Lohmann, J. R. Creating a culture for scholarly and systematic innovation in engineering education: Ensuring US engineering has the right people with the right talent for a global society. *American Society for Engineering Education.* Washington, DC, (2009).
2. Means, B. Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies. *U.S. Department of Education, Office of Planning, Evaluation, and Policy Development.* (2009).
3. Armstrong, N., Chang, S. M. & Brickman, M. Cooperative learning in industrial-sized biology classes. *CBE-Life Sci. Educ.* 6, **2**, 163–171 (2007).
4. Rhoads, J. F., Nauman, E., Holloway, B. & Krousgrill, C. M. The Purdue Mechanics Freeform Classroom: A new approach to engineering mechanics education. *121st ASEE Annual Conference and Exposition.* (2014).
5. DeBoer, J. et al. Work in progress: Rigorously assessing the anecdotal evidence of increased student persistence in an active, blended, and collaborative mechanical engineering environment. *123rd ASEE Annual Conference and Exposition*. (2016).
6. Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. Active learning increases student performance in science, engineering, and mathematics. *Proc. Natl. Acad. Sci. U. S. A.*, 111, **23,** 8410–8415 (2014).
7. Bonwell, C. C., Eison J. A., Active learning: Creating excitement in the classroom. *1991 ASHE-ERIC Higher Education Reports.* Washington D.C., (1991).
8. Berk, R. A. Multimedia teaching with video clips: TV, movies, YouTube, and mtvU in the college classroom. Int. J. Technol. Teach. Learn., 5, **1**, 1–21 (2009).
9. Meyers, S. A. Using transformative pedagogy when teaching online. *Coll. Teach.* 56**, 4,** 219–224 (2008).
10. Dillenbourg, P., Collaborative learning: Cognitive and computational approaches. *Advances in Learning and Instruction Series.*, New York, NY, (1999).

11. Armbruster, P., Patel, M., Johnson, E. & Weiss, M. Active learning and student-centered pedagogy improve student attitudes and performance in introductory biology. *CBE-Life Sci. Educ.* 8**, 3,** 203–213 (2009).
12. Preszler, R. W. Replacing lecture with peer-led workshops improves student learning. *CBE-Life Sci. Educ.* 8**, 3,** 182–192 (2009).
13. Prince, M. Does active learning work? A review of the research. *J. Eng. Educ.* 93**, 3,** 223–231 (2004).
14. Springer, L., Stanne, M. E. & Donovan, S. S. Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Rev. Educ. Res.* 69**, 1,** 21–51 (1999).
15. Barkley, E. F. *Student Engagement Techniques: A Handbook for College Faculty*. John Wiley & Sons. (2009).
16. Kane, T. J., Taylor, E. S., Tyler, J. H. & Wooten, A. L. Identifying effective classroom practices using student achievement data. *J. Hum. Resour.* 46**, 3,** 587–613 (2011).
17. Waxman, H. C., Weber, N. D., Franco-Fuenmayor, S. E. & Rollins, K. B. in *Teaching at Work* (eds. Li, Y. & Hammer, J.). Sense Publishers. 9–27 (2015).
18. Stuhlman, M. W. & Pianta, R. C. Profiles of educational quality in first grade. *Elem. Sch. J.* 109**, 4,** 323–342 (2009).
19. Waxman, H., Padrón, Y., Franco-Fuenmayor, S. & Huang, S. Observing classroom instruction for ELLs from student, teacher, and classroom perspectives. *TABE J.* 11**, 1,** 63–95 (2009).
20. Smith T. W., David S., Toward a prototype of expertise in teaching, *Journal of Teacher Education,* 55, **4,** 357 – 371 (2004).
21. Hill, H. & Grossman, P. Learning from teacher observations: Challenges and opportunities posed by new teacher evaluation systems. *Harv. Educ. Rev.* 83**, 2,** 371–384 (2013).
22. Raphael, L. M., Pressley, M. & Mohan, L. Engaging instruction in middle school classrooms: An observational study of nine teachers. *Elem. Sch. J.* 109**, 1,** 61–81 (2008).
23. Piburn, M. Reformed teaching observation protocol (RTOP) reference manual. Technical Report. (2000).
24. Smith, M. K., Jones, F. H. M., Gilbert, S. L. & Wieman, C. E. The classroom observation protocol for undergraduate STEM (COPUS): A new instrument to characterize university STEM classroom practices. *CBE-Life Sci. Educ.* 12**, 4,** 618–627 (2013).
25. Shekhar, P., Demonbrun, M., Borrego, M., Finelli, C., Prince, M., Henderson, C., & Waters, C., Development of an observation protocol to study undergraduate engineering student resistance to active learning. *Int. J. Eng. Educ.* 31**, 2,** 597–609 (2015).
26. Evenhouse, D. et al. Development of a video coding structure to record active, blended, and collaborative pedagogical practice. *2017 Research in Engineering Education Symposium*. (2017). Manuscript submitted for publication.
27. Hallgren, K. A. Computing inter-rater reliability for observational data: An overview and tutorial. *Tutor. Quant. Methods Psychol.* 8, **1,** 23–34 (2012).
28. McHugh, M. L. Interrater reliability: the kappa statistic. *Biochem. Medica.* 22**, 3,** 276–282 (2012).
29. Emam, K. E. Benchmarking kappa: Interrater agreement in software process assessments. *Empirical Software Engineering*. 4**, 2,** 113–133 (1999).
30. Jorgensen, F., Fridley, and J. L., Jorgensen, J. E. & Lamancusa, J. S. Benchmarking: A process basis for teaching design. *Frontiers in Education Conference.* 960–967 (1997).

31. Kasser, J., Hitchins, D., Frank, M. & Zhao, Y. Y. A framework for benchmarking competency assessment models. *Syst. Eng.* 16**, 1,** 29–44 (2013).
32. Cohen, J. A power primer. *Psychological Bulletin* 112, **1,** 155–159 (1992).
33. Sheskin, D. *Handbook of parametric and nonparametric statistical procedures*. Chapman & Hall/CRC. Boca Raton, FL, (2004).
34. Henderson, C., Beach, A. & Finkelstein, N. Facilitating change in undergraduate STEM instructional practices: An analytic review of the literature. *J. Res. Sci. Teach.* 48**, 8,** 952–984 (2011).