Development of a video coding structure to record Active, Blended, and Collaborative pedagogical practice

David Evenhouse[a], Claudio Cesar Silva de Freitas[a], Nimit Patel[a], Rohit Kandakatla[a], Austin Zadoks[a], Taylor Prebel[a], Nick Stites[a], Edward Berger[a,b], Jennifer DeBoer[a,b], Jeffrey Rhoads[b], Charles Krousgrill[b]

[a]School of Engineering Education
[b]School of Mechanical Engineering
Purdue University
West Lafayette, Indiana, USA

freeform@purdue.edu

***Abstract**: Classroom observation protocols (COPs) are valuable tools for collecting information about and characterizing classroom instruction. The purpose of this work is to describe the development and final structure of a specialized COP for use in coding video recordings of an active, blended, and collaborative (ABC) learning environment known as Freeform. Coding these ABC elements in the context of the instructor's actions allows the research team to empirically describe what occurs in the Freeform classroom, generating a unique and rich dataset containing information about the duration and sequencing of these ABC instructional elements. This data informs ABC instruction, facilitates further implementation of Freeform, and can inform further innovation based on reliable and valid observation data. Additionally, we believe documentation about the development process itself will empower other researchers to explore the use of COPs in analyzing and characterizing their own pedagogical practices in undergraduate STEM education.*

Introduction

Recent calls in the Engineering Education community for greater use of research based, innovative pedagogical practices (Jamieson & Lohmann, 2012; Lattuca, Bergom, & Knight, 2014) have increased the demand for tools that can be used to measure the effects, and support the implementation, of such innovative instruction. At times this requires collecting data during class meetings, observing how the various elements of the innovation, the instruction, and the institutional setting interact. Classroom Observation Protocols (COPs) can be used to generate reliable and detailed descriptions of what happens during an observed class period. On their own, or combined with other data such as student academic performance metrics, COPs can act as valuable tools to inform our understanding of pedagogical innovations.

Freeform is an innovative learning environment developed at Purdue University – West Lafayette in the School of Mechanical Engineering. It intentionally combines best practices in active, blended, and collaborative (ABC) learning to improve student acquisition of conceptual knowledge and problem solving skills (Rhoads, Nauman, Holloway, & Krousgrill, 2014). Since Freeform's inception in the Mechanical Engineering undergraduate Dynamics course, the number of students who earn a D, fail (F) or withdraw (W) from their course (the so-called DFW rate) each semester has decreased significantly (DeBoer et al., 2016). To better understand what Freeform looks like in practice, and to empower Freeform developers to bring these benefits to other instructors, courses, and institutions, our team has been using classroom observation techniques to identify characteristics of these Freeform classes through highly structured video analysis.

In this paper, we present our approach to crafting a COP for use in the analysis of video recordings of Freeform class meetings. In so doing, we address the research question:

> What are the necessary components of a COP to capture ABC practices, and how do we adapt or develop such a COP for the unique needs of our research context?

We have developed and are currently using this new COP to evaluate instruction within the Freeform environment. Its structure, and method of creation, can inform the work of others as they examine and assess innovative teaching practices in undergraduate STEM education.

Background: Foundations of a classroom observation protocol

Instructional context: Freeform

Freeform combines ABC practices to enhance student learning in conceptual knowledge and problem solving skills. Any COP employed in Freeform research must therefore be able to capture each of these ABC elements in the form of valid and reliable data. The validity of a COP is demonstrated if the instrument being used successfully measures the variables that it is intended to measure (DeMonbrun, Finelli, & Shekhar, 2015). The first step in our process of selecting or developing a COP is to rigorously define those aspects of Freeform that we intend to measure. In this paper, we demonstrate the validity of this instrument through the intentional alignment of our COP with the nature of Freeform and its broader institutional context.

Active learning in Freeform is defined as learning that requires students to engage cognitively and meaningfully with the materials (Bonwell & Eison, 1991). It requires action (Chi & Wylie, 2014) on the part of the students, and cannot include traditional lecturing on the part of the instructor (Freeman et al., 2014). Blended learning is defined as the combination of in-class learning with web-based online learning elements, and does not necessarily require a fully flipped classroom (Oliver & Trigwell, 2010) [Note: we do not view Freeform as a flipped classroom]. Finally, collaborative learning occurs when two or more students work interactively as they progress toward specified learning outcomes (Berkeley, 2014), in this case within the setting of the classroom environment. Collaborative learning is also considered to be a form of active learning (Chi & Wylie, 2014).

In her study on curricular innovation, Hjalmarson (2008) defines a variety of different curricular systems based upon what innovations are being emphasized. Freeform focuses on the pedagogical strategies and learning resources being employed to present information to students, framing it as a pedagogically-based curricular system (Hjalmarson, 2008). Though the resulting environment may be considered student driven, the core of Freeform's impact on education lies in the alignment of instructors' in-class pedagogical choices with both research-based instructional innovations, and the Freeform learning resources made available to students online. Observations of the instructors' actions in their classrooms are therefore essential to understanding Freeform and informing its future implementations.

Institutional context: Resources, data, and limitations

In our work to characterize Freeform, the team decided to video record class meetings, aiming for more accurate, efficient, and less intrusive analysis than can be provided by in-class, real-time observation approaches. Several unique limitations and opportunities arose from the video capture and analysis methods employed. Because Freeform is pedagogically-based, our videos intentionally track instructor actions rather than the student responses to those actions. Coding actions and interactions of the students, for example, would be impossible given the nature of our video samples. However, tracking the actions of the instructor and recording their pedagogical choices provides us with the clearest picture of how Freeform has been applied to the classroom we are observing. This focus also mitigates ethical and professional concerns brought up by the University's Institutional Review Board related to recording students.

We chose to use the StudioCode software package to assist us with our video analysis, as it affords us a variety of unique opportunities. In StudioCode, coding structures such as COPs may be intentionally crafted with internal inclusions and exclusions, building reliability (consistency in coding across users and datasets) into the instrument and allowing for coding multiple instances, or events, simultaneously. Video speed can also be increased or decreased to facilitate time efficiency or coding precision as needed. This all allows researchers to easily extract qualitative and quantitative data from videos for subsequent analysis.

Establishing the need for a specialized COP

Previously-documented COPs have not been designed to capture all the features of interest to us within the Freeform environment. For example, Smith et al. developed an observation protocol to characterize university STEM classroom practices (Smith, Jones, Gilbert, & Wieman, 2013). This Classroom Observation Protocol for Undergraduate Stem (COPUS), which was heavily influenced by the pre-existing Reformed Teaching Observations Protocol (RTOP), was developed to provide a more targeted and context-appropriate tool for classroom observation in collegiate STEM courses. However, COPUS does not explicitly address both conceptual and problem solving knowledge, does not distinguish among ABC learning practices, incorporates observations of the students in addition to the instructor, and employs real-time coding at two minute intervals rather than continuous, offsite coding of video recordings (Smith et al., 2013). Likewise, Finelli et al. developed a coding scheme that observed the degree of student engagement in active learning, a point of interest to the Freeform team (Finelli et al., 2014). However, the innovations of the Freeform environment focus on the pedagogical choices and subsequent actions of the instructor. Thus, it would be more beneficial to us for the purpose of our project to observe the actions of the instructor and record the amount of student engagement the instructor intends to facilitate, a slight but important variation from the work of Finelli et al. Finally, the COP developed by Finelli et al. does not capture blended learning, nor several types of instructor actions, which prevents us from simply taking and adapting this work to our research.

The instrument we developed improves upon existing COPs for application within the Freeform framework by: (i) using a video-based offline procedure (enabling continuous and flexible coding), (ii) differentiating of ABC pedagogies, and (iii) discerning both conceptual and problem solving foci of instructor actions. To do this we integrated Finelli's description of active learning with robust coding definitions from the COPUS instrument, supplemented these with blended and collaborative codes, and finally created a new COP calibrated for observations within an ABC environment.

Developing a Freeform video coding protocol

Overall structure of the COP

We conceptualized the new COP to include instructor Events, Characterizations of those events, and the Degree of Engagement of students in learning activities. Events form the backbone of our coding protocol by denoting mutually exclusive, in-class applications of distinct pedagogical practices or behaviors. This is much like what may be seen in other prominent COPs that demonstrate a focus on instructor actions. For example, the COPUS instrument incorporates twelve distinct codes to record what the instructor is doing during a given classroom observation (Smith et al., 2013). Events are coded continuously in StudioCode (rather than at specified intervals), an affordance of using offline video analysis rather than real-time, in-class observation.

To capture the presence or absence of ABC learning practices, we include a set of Characterization codes, activated in tandem with the Event codes. Characterization codes indicate that the instructor is explicitly incorporating active, blended, or collaborative practices in their teaching. For example, an Event code denoting that the instructor in conducting an assessment could include Characterization codes indicating the assessment is both active (students are solving a problem) and collaborative (in the case of a group quiz) in nature.

Using Characterization codes alongside Event codes informs our understanding of an instructor's in-class intentions in an efficient and manageable way.

Broader literature indicates that the degree to which students are actively engaged is significant to their eventual learning outcomes (Chi & Wylie, 2014; Finelli et al., 2014). Several Degree of Engagement codes capture the extent of student involvement in the learning activities, expressed as 'All', 'Some', or 'None'. These codes capture instructor intentions, not actual student actions, because our data collection and analysis is entirely focused on pedagogical decisions of the instructor. For example, an in-class quiz would be tagged with a Degree of Engagement of 'All', because all students in the class would be involved in completing the quiz.

Lastly, two catch-all terms were included in the coding scheme. This was done later in the development process to facilitate data retrieval, synthesizing existing information rather than generating new data. The 'Assessment' term combines all data from Event codes that involve assessment of students, while 'Direct Instruction' combines all codes related to direct, passive (i.e. lecture) instruction.

StudioCode also enables forced inclusions or exclusions, linking activated codes to other, related codes in an automated way. For example, if an in-class quiz were to be coded as a 'Graded Assessment' Event, the Characterization code 'Active' and the Degree of Engagement code 'All' would automatically be applied to the event by the StudioCode environment itself. These conceptual linkages between Events, Characterizations, and Degrees of Engagement facilitate consistent and efficient coding of instructional events.

Codes and definitions

Multiple rounds of testing and feedback were required to establish the final coding scheme and to determine the finer points of the coding definitions. This process included meetings with Freeform professors and researchers, application of the coding scheme, and reliability analysis using Cohen's Kappa statistic (Spitzer, Cohen, Fleiss, & Endicott, 1967). This progression can be seen in Figure 1.

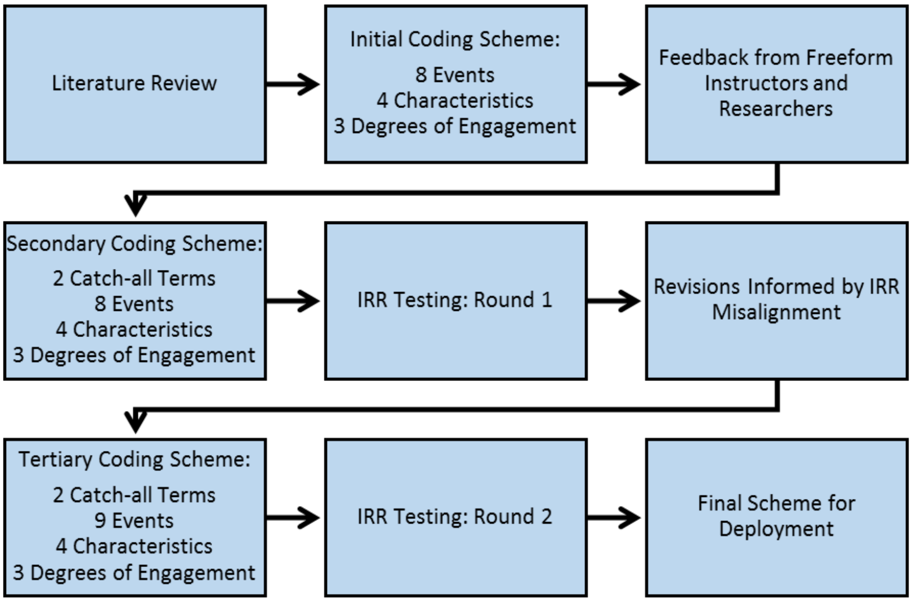

Figure 1: A visual summary of our coding scheme development process.

Events: Our final structure includes nine Event codes. These events are largely informed by the COPUS and RTOP instruments, with slight adjustments having been made to better align them with the purpose and interests of the Freeform project. For example, Talking/Lecturing (T/L) and Real Time Writing (RTW) codes directly reflect the COPUS framework, which distinguishes between the two as in Freeman et al. (2014). Likewise, our Demonstration code is derived from the COPUS D/V (Demonstration/Video) coding, and our Other/Admin code combines both the W (Waiting) and Adm (Administration) codes from COPUS (Smith et al., 2013). Opportunities for student and instructor feedback, assessment, and interactive learning are then captured with the Questions and Assessment codes. The full list of nine Events is below:

- *Conceptual T/L*: 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
- *Problem Solving T/L*: 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 a problem statement, etc.
- *Conceptual RTW:* The instructor is explaining some concept (e.g., free body diagrams, equation derivations, etc.) by writing on the board.
- *Problem Solving RTW:* The instructor is solving some example problem on the board, demonstrating the application of equations, or enumerating a problem solving process.
- *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.
- *Graded Assessment:* This categorization includes instances where the instructor asks a question or series of questions of the students that all of the students are expected to answer for a grade; for example, quizzes, 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 of the students are expected to answer but their responses are not graded; for example, feedback forms, problems or examples given to solve in class.
- *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.
- *Other/Admin:* 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.

Characterization: Characterization codes classify the events above to verify the presence, and record the duration, of ABC learning practices. Here, we see the influence of the essential nature of Freeform as an ABC learning environment (Rhoads et al., 2014) as well as literature on each of the ABC practices. Events may be classified using four characteristic codes:

- *Active (A):* Includes all interactive, constructive, and active events in the classroom (Chi & Wylie, 2014; Freeman et al., 2014).
- *Blended (B):* Includes any use or mention of online resources (Bowen, Chingos, Lack, & Nygren, 2012; Means, Toyama, Murphy, Bakia, & Jones, 2009)
- *Collaborative (C):* Includes any peer-to-peer interaction within the classroom (Dillenbourg, 1999; Jeong & Chi, 2007).
- *Passive (P):* A null instance representing the lack of active learning, does not expressly promote any activity or interaction on the part of the students.

'Passive' and 'Active' are, by our definitions, mutually exclusive, a nuance programmed directly into our coding scheme. Likewise, collaborative events are assumed to be active, thus classifying an event as 'Collaborative' forces it to be 'Active' as well. Our final scheme is laid out in StudioCode such that most event codes are automatically characterized to facilitate greater coding reliability. For example, selecting a T/L or RTW event code will automatically characterize the event as 'Passive', as demonstrated in Figure 2.

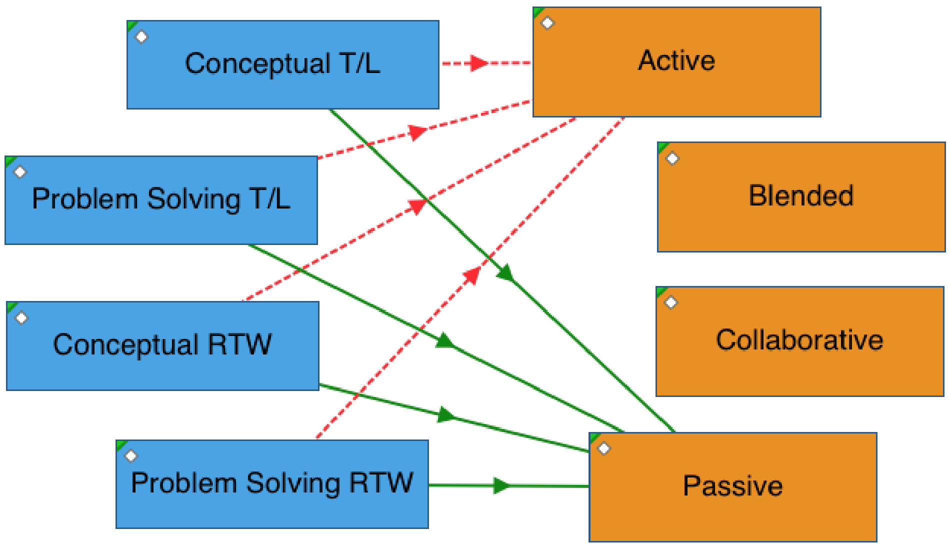

Figure 2: The relationships between T/L and RTW Events and their Characterization codes. Solid lines indicate a forced activation, while dotted lines indicate a deactivation.

Degree of Engagement: Degree of Engagement codes draw heavily on the work of Finelli et al. (2014), and were adjusted to reflect our focus on the actions of the instructor. Like Finelli et al. (2014), we opted to use a three-tiered coding scheme. However, rather than recording student engagement, the 'None', 'Some', and 'All' codes capture instructors' intentions regarding student engagement through a given instructional behavior. Engagement codes again take advantage of built-in dependencies in StudioCode. For example, Passive events are automatically coded as having a Degree of Engagement of 'None'.

Whenever codes are selected, instances of the video being coded are saved to a timeline as is shown in Figure 4. These timelines may then be queried to generate numerical data such as the length of instances, the number of instances, or the amount of overlap between the various codes. Each of these instances also represents an individual video clip, which may be exported, watched, and even transcribed to generate qualitative data for analysis.

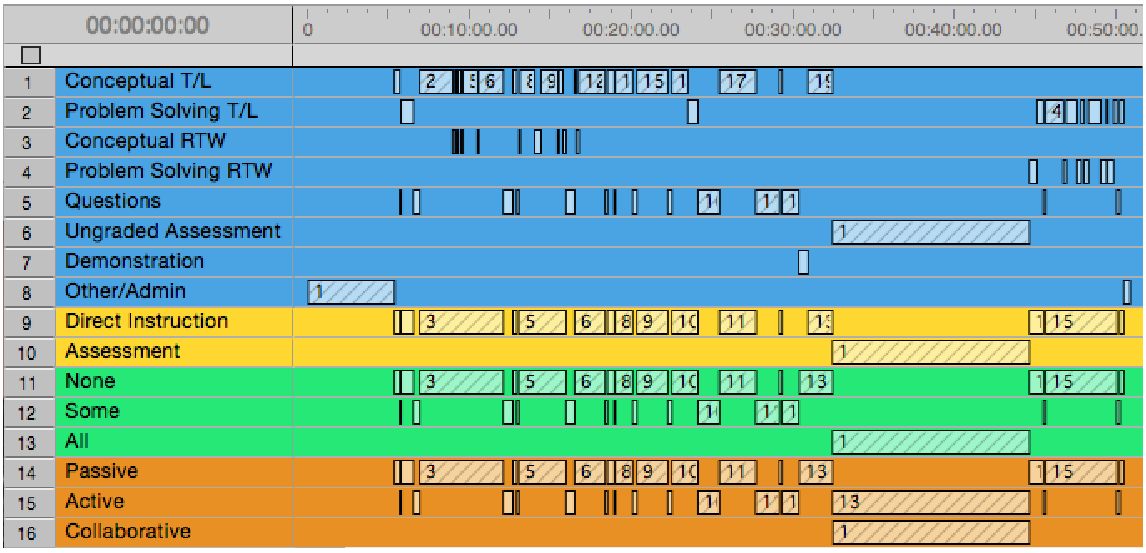

Figure 3. A representative StudioCode timeline. Code names are located in the leftmost column, with numbered instances arranged along the timeline to the right.

Validity, reliability, and further application

This COP has been intentionally designed to measure the pedagogical variables of interest to us within the Freeform environment, making it valid for use in this research context (DeMonbrun et al., 2015). Like in the work of DeMonbrun et al., we foster concurrent validity by drawing from the designs of existing, well-vetted observation protocols, and foster face validity by firmly rooting our code definitions in the theory behind our topics and phenomena of interest. Yet validity concerns do not begin and end with the development of our COP. To ensure that this validity is preserved moving forward, we will have to take care during our sampling process to ensure that the class recordings we take are truly representative of what occurs in the Freeform classroom. Well defined coding procedures can only prove useful if we are collecting the proper data for analysis (DeMonbrun et al., 2015).

In determining reliability, we employed Cohen's Kappa (Hallgren, 2012; Spitzer et al., 1967) as a means of calculating Inter-Rater Reliability (IRR) as seen elsewhere in STEM literature (El Emam, 1999; Kilgore, Atman, Yasuhara, Barker, & Morozov, 2007). During development, we had two coders analyze four videos for the purpose of IRR testing. Both coders were involved in the development of this coding scheme. After coding two videos, an initial Kappa was calculated in order to determine where there may be ambiguities in the coding structure or definitions. For example, an initial Kappa of 0.695 was calculated comparing the coding of Active and Passive events. This number is considered sufficiently reliable for the purpose of research (El Emam, 1999; Kilgore et al., 2007) but certainly suggests room for improvement. After a review of the coding structure and several theory-based revisions to the code definitions, another test was conducted which demonstrated a Kappa value of 0.798. Across all the components of the coding structure (Events, Characterization, and Degree of Engagement), and after adding additional coders, Kappa continues to be at or above 0.6, indicating moderate to substantial agreement among the coders who have now participated (El Emam, 1999).

Discussion

The development process and definitions presented here have resulted in a custom COP that targets key variables of interest within Freeform's ABC learning environment. Its components have been designed to align these target variables, the abilities of the software package available to us, and the insight gained from the literature at large, enabling us to produce a unique COP and resultant dataset detailing what exactly occurs within this innovative classroom environment.

This COP has already been put to use in research, and has demonstrated promising preliminary results. A publication by our team at the 124th ASEE Annual Conference and Exposition (Evenhouse et al., 2017) presents some of these preliminary findings, as well as their implications for Freeform instructors. For example, we are presently working to benchmark and characterize instruction within the Freeform environment across instructors and semesters. Doing so will inform future implementations of Freeform at our institution and others, as well as contribute to the broader body of literature on ABC instructional practices. Locally, this empirically generated dataset can facilitate empirically based professional development, fostering reflection and encouraging Freeform instructors to continuously improve on their own pedagogical practices (Creemers & Reezigt, 1996).

This work also expands on existing classroom observation literature by combining the coding of instructor behaviors with the coding of variables designed to evaluate the use of research-based, innovative instructional practices. In so doing, it demonstrates the potential of video coding to inform and facilitate the implementation of research-based instructional practices, a topic of interest within the broader engineering education community (Dancy & Henderson, 2008; Henderson, Beach, & Finkelstein, 2011; Litzinger & Lattuca, 2014).

Conclusion

This paper details the development of a customized COP to help us understand classroom practices in an ABC environment. Built upon prior COP instruments, our framework and its implementation extend the literature on classroom observation and evaluation, and provide an outline for other researchers to systematically build COPs for their specific needs. Although we do not recommend this tool for use in non-ABC environments, the methods described here can be used to guide the development of tools designed for different learning environments, where other teaching practices could be examined and observed. Though this coding scheme was developed to support the needs of the Freeform research group, we expect that the methods described in this paper can be useful in future applications in a variety of educational contexts.

References

Barkley, E. F. (2014). *Collaborative learning techniques: A handbook for college faculty* (2nd ed.). San Francisco, California: Jossey-Bass.

Bonwell, C. C. (1991). *Active learning: Creating excitement in the classroom*. Washington, D.C.: School of Education and Human Development, George Washington University.

Bowen, W. G., Chingos, M. M., Lack, K. A., & Nygren, T. I. (2012). *Interactive learning online at public universities: Evidence from randomized trials*. Ithaka S+R.

Chi, M. T. H., & Wylie, R. (2014). The ICAP Framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*(4), 219–243.

Creemers, B. P. M., & Reezigt, G. J. (1996). School level conditions affecting the effectiveness of instruction. *School Effectiveness and School Improvement*, *7*(3), 197–228.

Dancy, M. H., & Henderson, C. (2008). Barriers and promises in STEM reform. In *National Academies of Science Promising Practices Workshop*.

DeBoer, J., Stites, N., Berger, E. J., Rhoads, J. F., Krousgrill, C. M., Nelson, D. B., Zywicki, C., Evenhouse, D. (2016). Work in Progress: Rigorously assessing the anecdotal evidence of increased student persistence in an active, blended, and collaborative Mechanical Engineering environment. In *123rd ASEE Annual Conference and Exposition*.

DeMonbrun, R. M., Finelli, C. J., & Shekhar, P. (2015). Methods for establishing validity and reliability of observation protocols. In *122nd ASEE Annual Conference and Exposition*.

Dillenbourg, P. (1999). What do you mean by “collaborative learning”? *Collaborative-Learning: Cognitive and Computational Approaches.* (pp. 1–19). Elsevier.

El Emam, K. (1999). Benchmarking Kappa: Interrater agreement in software process assessments. *Empirical Software Engineering*, *4*(2), 113–133.

Evenhouse, D., Kandakatla, R., Stites, N., Patel, N., Zadoks, A., Prebel, T., Freitas, C., Krousgrill, C., Rhoads, J., DeBoer, J. (2017). What does an in-class meeting entail? A characterization and assessment of instructor actions in an active, blended, and collaborative classroom. In *124th ASEE Annual Conference and Exposition*.

Finelli, C. J., DeMonbron, M., Shekhar, P., Borrego, M., Henderson, C., Prince, M., & Waters, C. K. (2014). A classroom observation instrument to assess student response to active learning. In *2014 IEEE Frontiers in Education Conference (FIE) Proceedings*.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, *111*(23), 8410–8415.

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, *8*(1), 23–34.

Henderson, C., Beach, A., & Finkelstein, N. (2011). Facilitating change in undergraduate STEM instructional practices: An analytic review of the literature. *Journal of Research in Science Teaching*, *48*(8), 952–984.
Hjalmarson, M. (2008). Mathematics curriculum systems: Models for analysis of curricular innovation and developement. *Peabody Journal of Education*, *83*(4), 592–610.
Jamieson, L. H., & Lohmann, J. R. (2012). Innovation with impact. *American Society for Engineering Education*. Washington, DC.
Jeong, H., & Chi, M. T. H. (2007). Knowledge convergence and collaborative learning. *Instructional Science*, *35*(4), 287–315.
Kilgore, D., Atman, C. J., Yasuhara, K., Barker, T. J., & Morozov, A. (2007). Considering context: A study of first-year engineering students. *Journal of Engineering Education*, *96*(4), 321–334.
Lattuca, L. R., Bergom, I., & Knight, D. B. (2014). Professional development, departmental contexts, and use of instructional strategies. *Journal of Engineering Education*, *103*(4), 549–572.
Litzinger, T. A., & Lattuca, L. R. (2014). Translating research to widespread practice in engineering education. In A. Johri & B. M. Olds (Eds.), *Cambridge Handbook of Engineering Education Research*.
Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2009). *Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies*. Washington, D.C: US Department of Education.
Oliver, M., & Trigwell, K. (2010). Can "blended learning" be redeemed? *E-Learning and Digital Media*, *2*(1), 17–26.
Rhoads, J. F., Nauman, E., Holloway, B., & Krousgrill, C. M. (2014). The Purdue Mechanics Freeform Classroom: A new approach to engineering mechanics education. In *121st ASEE Annual Conference and Exposition*.
Smith, M. K., Jones, F. H. M., Gilbert, S. L., & Wieman, C. E. (2013). The classroom observation protocol for undergraduate stem (COPUS): A new instrument to characterize university STEM classroom practices. *CBE Life Sciences Education*, *12*(4), 618–627.
Spitzer, R. L., Cohen, J., Fleiss, J. L., & Endicott, J. (1967). Quantification of agreement in psychiatric diagnosis. *Archives of General Psychiatry*, *17*(1), 83–87.
Tileston, D. W. (2011). *Ten best teaching practices: how brain research and learning styles define teaching competencies* (3rd ed.). Thousand Oaks, CA: Corwin Press.

Acknowledgements

We would like to acknowledge our full Freeform research team for their help in this instrument's development. Likewise, we would like to thank Dr. Tara Kaczoroski, author of the blog "StudioCode for Educators", whose advice was invaluable to our coding and reliability testing. This material is based on work supported by the National Science Foundation (NSF), Division of Undergraduate Education under Grant No. 1525671. Any opinions, findings, conclusions, or recommendations expressed here are those of the authors and do not necessarily reflect the views of the NSF.

Copyright statement