Methods for establishing validity and reliability of observation protocols

Mr. Robert Matthew DeMonbrun, University of Michigan

Matt DeMonbrun is a Ph.D. student at the Center for the Study of Higher and Postsecondary Education (CSHPE) in the School of Education at the University of Michigan. His research interests include college student development theory and teaching practices and how they relate to student learning outcomes in engineering education.

Dr. Cynthia J. Finelli, University of Michigan

Dr. Cynthia Finelli, Director of the Center for Research on Learning and Teaching in Engineering and research associate professor of engineering education at University of Michigan (U-M), earned B.S.E.E., M.S.E.E., and Ph.D. degrees from U-M in 1988, 1989, and 1993, respectively. Prior to joining U-M in 2003, she was the Richard L. Terrell Professor of Excellence in Teaching, founding director of the Center for Excellence in Teaching and Learning, and associate professor of electrical engineering at Kettering University. In her current role, she coordinates faculty and TA professional development in the College of Engineering, conducts rigorous engineering education research, and promotes the growth of engineering education both locally at UM and nationally. Dr. Finelli's current research interests include evaluating methods to improve teaching, studying faculty motivation to change classroom practices, and exploring ethical decision-making in engineering students. She also has established a national presence in engineering education; she is a fellow in the American Society of Engineering Education, is an Associate Editor of the IEEE Transactions on Education, and past chair of the Educational Research and Methods Division of ASEE.

Mr. Prateek Shekhar, University of Texas, Austin

Prateek Shekhar is a PhD student in the Department of Mechanical Education at the University of Texas at Austin. His research is focused in understanding students' and faculty's reaction to adoption of active learning based teaching methods in engineering classrooms. He holds a M.S. in Electrical Engineering from University of Southern California and B.S. in Electronics and Communication Engineering from India.

Methods for Establishing Validity and Reliability of Observation Protocols

Classroom observations can be a useful tool in conducting research on a myriad of interactions and events that occur in an educational setting. Research in much of the social sciences and industrial management involves the use of trained human observers as a primary source of data collection to answer research questions about interactions, faculty/student behaviors, instructor effectiveness (performance evaluations), and teaching methods utilized by instructors. Observations have been used in engineering education research to investigate faculty use of different types of instructional methods, interactions between faculty and students in various courses, and the relationship dynamics of student teams in design courses, among other research areas.

Our research team has utilized classroom observations to examine students' resistance to faculty's use of active learning methods in the engineering classroom. Specifically, we have used observations to further understand the various ways in which students exhibit resistance to active learning methods and the ways faculty respond to this challenge. Since trained observers who are not involved parties in the classroom (i.e., neither students nor instructors) are conducting our observations, we have had to continually reflect on the precise detection, perception, recognition, and judgment of certain events to ensure our observations are accurately capturing what is occurring in the classroom. This experience is much different than training observers to obtain inter-rater agreement, used often by researchers to ensure that observations are reliable across many different observers. Instead, we have examined ways in which we can confirm the events we are recording are a **valid depiction of classroom behaviors**.

In this paper, we discuss the essential steps in confirming valid observations of classroom behaviors. We start with an overview of the concepts of reliability and validity in conducting observations, using previous techniques from the social science literature. Next, we discuss our project on student resistance and steps taken to ensure reliability and validity in our observations. Although our instrument was developed from other published observation protocols, we used multiple approaches to ensure the accuracy of both our instrument and the observations we were conducting. These included videotaping classroom behaviors and conducting student focus groups to confirm the precision of our observation processes.

Validity and reliability in observation protocols

Observations, especially those in the classroom, are a research methodology within the continuum of ethnography, which finds its origins in 19th century Western anthropology[1]. Ethnography has become a commonly accepted means of qualitative research, especially when researchers want to capture research subjects in their settings[2]. Hammersley and Atkinson[1] note that ethnography is defined by several common features across fields in the social sciences: selecting and sampling cases for study, securing access to participants and/or settings, conducting observations and interviews, recording and storing data, and analyzing data for writing reports on findings. Although each of these features comes with its own distinct difficulties, the process of analyzing ethnographic data is often one of the most difficult steps for researchers to navigate during the research process. Much of this confusion comes from attempts to demonstrate an understanding of what was actually observed.

In their seminal book, "Writing Ethnographic Fieldnotes," Emerson, Fretz, and Shaw[2] discuss techniques for writing effective fieldnotes in a variety of observational settings. They state that fieldnotes can often be written from multiple perspectives. Using a *first-person point of view*, researchers are able to describe specifically what they observe or experience during the data collection process. This is particularly useful when the researcher is a member of the group s/he is observing because this provides readers with an "insider's" point of view. In contrast, the *third-person point of view* allows researchers to capture exactly what individuals say or do during the observation, offering a narrative of what has occurred.

But, how does a researcher write about others' thought processes or behaviors while doing some task or event? Emerson, Fretz, and Shaw[2] describe this type of observation as the *focused third-person point of view*. In this situation, the researcher is observing through a third-person lens by documenting what is happening in the setting, but the researcher also includes details that only the individual being observed might notice. Emerson, Fretz, and Shaw[2] note that, "though the researcher might make inferences about thoughts and feelings, he would base them on observable facial expressions, gestures, and talk, and describe these from the [individual's] perspective" (p. 98).

Ensuring good reliability and validity in such an observation is particularly difficult for any researcher. After all, how does one know that what s/he has seen and interpreted is actually what the observed individual is experiencing? Good **reliability** in an observation protocol certifies that observations will be consistent across time or observers[3]. Good **validity** in an observation protocol ensures that the observation instrument actually measures what it is intended to measure.[4] Although this might be easy for some protocols that only encompass content observations (e.g., the student asks a question, the instructor uses an example in a real-world context), those that involve behavioral observations (e.g., students appear disengaged, student is off-task on his or her computer) are much more difficult because they assume that the observer knows precisely what the subject is doing.

Previous research has determined that the reliability of observational studies can be decomposed into two measurement theories: classical test theory and generalizability theory[5-6]. Classical test theory assumes that all measurements are generally equivalent in their content, mean, variance and covariance. Generalizability, on the other hand, recognizes that there are a multitude of potential errors that can cause different results during observations. These errors could include observer bias, differing subject demographics, and situational effects (e.g., technology is broken in the classroom, weather was particularly problematic for students, Friday afternoon classes vs. Monday morning classes)[6].

There are several ways to control for such errors in generalizability. One way is through *duplicate generalizability studies*[7]. For example, researchers could have two observers view the same video of a class, write down their observations, and discuss answers as a group. Another way to control for these errors is through *session generalizability studies*[7]. In this case, researchers could have observers view the same course and alternate their observations by minutes or instances of the event. Finally, *developmental generalizability studies*[7] can be conducted, in which case a test-retest approach is applied. For instance, one observer might view classes at the beginning of the semester, while another (or possibly even the same) observer views classes at the end of the semester to illustrate differences across the term. Each of these cases of reliability testing is only meaningful if observers then come to a conclusion regarding differing observations, and a protocol is established for effectively labeling such cases in the future, or while training other observers.

The reliability of an observed behavior is also closely linked to the validity of the observation. As described by Gardner[8], "it is important to note that reliability of a measure sets limits on its validity." (p. 188). In other words, without a **valid** protocol for observation, the **reliability** of the measurement tool is not useful. Validity for observations does not require the same "gold standard" analysis that is used in quantitative research[9]. Rather, qualitative researchers try to protect against the various ways in which their research analyses could be misleading. Maxwell[4] calls this "validity threat." Therefore, validity in qualitative research consists of a researcher's conceptualizations of all possible types of "validity threats" and strategies by which they attempt to deal with these threats, assuming they are plausible.

Maxwell[4] makes recommendations for ways in which researchers can protect against these validity threats, and several of those are particularly useful for observational data. The first is planning for intensive, long-term involvement with the research study[10]. Little interpretation can be made from one or two cases, but several observations made over and over again with similar populations can lead to trends and potential theories. Second, researchers should plan to collect "rich" data to get a full picture of what is happening in the observations[11]. Observers may very well be effective at capturing all that is happening during an observation, but observations backed up with audio or video recordings allow researchers to go back and reexamine what the observer documented. Third, triangulation using multiple data sources and participants can improve the validity of one's dataset, and ensure that results do not apply to only one observation or sample population[12]. Finally, validation by the individuals being observed, otherwise known as "member checking," ensures that the conclusions of the observer match what is actually happening during the observation[13-14]. This last form of validation is the focus of our paper. One example of how to conduct "member checks" is discussed further below.

Student resistance to active learning in the engineering classroom

Instructors have reported many barriers that discourage their use of active learning practices in the classroom. These barriers include: (a) concerns about student resistance, (b) questions about the efficacy of the teaching method, (c) concerns about preparation time and (d) concerns about ability to cover the syllabus[15-19]. Among these barriers, one area most in need of additional research is that concerning student resistance. Although student resistance can be a significant discouragement to faculty attempting new teaching practices, it is a natural response to new teaching methods not typically used in the classroom. Weimer[20] states that student resistance to active learning methods is often a result of the additional effort needed on behalf of the student, causing anxiety about their ability to succeed within this new classroom environment. Weimer[20] also noted that student resistance can often take a number of forms. Open resistance is characterized by emotional complaints, arguments about the usefulness of the task, and verbal objection to performing the tasks. Passive, non-verbal resistance occurs when individuals exhibit an overall lack of enthusiasm and often refuse to participate in the activity. Finally, partial compliance occurs when individuals perform the bare minimum of responsibilities to complete the task, or rush through the task in order to finish as quickly as possible.

Classroom observational instruments tend to vary significantly in terms of how they are utilized. For example, the Teaching Dimensions Observation Protocol[21] (TDOP) asks observers to conduct ratings in 2-minute time intervals, while the Reformed Teaching Observation Protocol[22] (RDOP) asks observers to make ratings over the entire class period. Given our desire to understand student reactions to active learning instructional practices, we decided to develop an instrument that could be used to rate individual instances of active learning. In other words, each instance of active learning used during the class period constitutes a separate observation, and observations can be compared to one another during a class period or throughout the semester.

Our classroom observation protocol[23] consists of two protocols used at the beginning of the semester and throughout the academic term. The *first-day protocol* documents the first day of class and any mention of active learning practices to be used throughout the term. Afterwards, the *daily classroom observation protocol* is completed for each instance of active learning that occurs during each of the class periods. This protocol documents several aspects of active learning: 1) basic course details, including start and stop times for the activity, 2) information about each active learning instance, including the level of difficulty and novelty of the material being discussed, 3) the type of active learning, 4) the degree of faculty participation in the activity, 5) how the instructor introduces the instance of active learning, and 6) student response during the activity.

In order to ensure the reliability of our observations and validate our protocol, we conducted three steps while designing and refining our instrument. First, we crafted our instruments from several previously validated observation instruments developed for or adapted to STEM courses, including: the Reformed Teaching Observation Protocol (RTOP)[22], UTeach Observation Protocol (UTOP)[24], Teacher Behaviors Inventory (TBI)[25], Teaching Dimensions Observation Protocol (TDOP)[21] and VanTH Observation System (VOS)[26]. Since most of these protocols focus on the extent to which students are actively engaged in their learning, we adapted the instrument to measure Weimer's[20] instances of student resistance.

Second, we developed a training protocol to ensure inter-rater reliability across multiple observers. Observers were asked to view a short videotape of a class session with an instructor who incorporated several instances of active learning. Observers then completed an observation form for each instance of active learning they observed during the video. Afterwards, observers met to discuss findings on the overall number of active learning instances, as well as details regarding each observation. In doing so, observers were able to compare and contrast results from their observations in order to ensure accuracy in our findings across all observers.

Finally, we conducted student focus groups at two sites involved in the research study, in order to ensure that our observations were accurate interpretations of students' experiences during active learning instances in the classroom. This portion of our validation process is the focus of the remainder of this paper.

Methodology

After developing our initial observation protocol, we conducted focus groups with students at two institutions to validate the accuracy of our observations. Students were asked to participate in an hour-long focus group, addressing students' resistance to active learning practices in the engineering classroom. One of our sites, Institution A, included six students recruited from two observed classes in our pilot study. The second site, Institution B, recruited 16 students from multiple engineering classes (observations were not conducted at this institution).

These focus groups consisted of two parts. For the first part, we asked students about the ways their instructors used active learning practices in the classroom. We stepped through the separate sections of our observation protocol with students. For example, we provided examples of the types of active learning practices included in our protocol, and asked students, "Here are some ways that we've seen instructors use active learning practices in the past. What other things that your instructors do are not on the list?" For other parts of our protocol, we did not provide students with examples. Rather, we asked students to participate in freethinking exercises around specific aspects of their classroom experiences. For example, we asked students, "How does your instructor introduce or talk about active learning activities?" rather that provide them with a list of ways that instructors typically introduce activities. In doing so, we ensured that our protocol confirmed students' reflections of their classroom experiences, rather than our own impressions tainting their responses.

The second part of our focus group included questions about ways students might demonstrate resistance during active learning experiences. We presented open access videos and images of students participating in active learning activities. Students viewed photos and videos grouped into four scenarios and were asked questions regarding observed student behaviors during each scenario. During the first and second scenarios, we presented an active learning instance where students were asked to work together in a group to design a solution to an industry specific engineering problem or discuss a programming technique with the whole class. The primary focus of these scenarios was to validate student engagement during activities involving either multiple groups or the entire class. In the third scenario, we presented a video with an active learning exercise asking students to work in pairs (or triads) and sharing their findings with the group. Finally, we presented a scenario focusing on individual work where the students were being asked to work on a problem by themselves. All four of these scenarios matched the types of active learning that were represented in our observation protocol. After each video and picture, we asked about which students were and were not actively engaged in the activity. Then, we asked students in our focus group to clarify how they knew these students were or were not engaged.

Findings

The first part of our focus groups confirmed the initial design of our protocol. When students reviewed sections of our protocol, they gave examples of when they had seen certain active learning teaching practices used in the classroom, how instructors introduced these activities, and the types of student reactions they had exhibited or observed from other peers. Hearing more open-ended questions regarding instructor's attempts to engage students in active learning, or peers' reactions to working in teams allowed us to group responses not listed in our protocol under one of the current categories. If students felt that these categories did not accurately capture something they had observed, we documented these disparities to discuss with our team of researchers for possible additions and amendments to the protocol.

The second part of our focus groups helped us to shape future training processes for observers. Our findings indicated that focus group participants utilized student facial expressions and body language to differentiate between engaged and disengaged students. For facial expressions, most of the participants reported that the students who were smiling or laughing in the picture were engaged in some off-task discussion, especially if it did not involve direct eye contact with the instructor or all group members. They indicated that an exception to this situation might be if the instructor offered positive feedback to the group, in which individuals would naturally be inclined to respond with positive facial expressions.

Participants also indicated that, within groups, smiling or laughing was often a sign of the difficulty of the material or novelty of the exercise. They stated it was often a natural reaction to joke or mingle with other group members after an activity was over, or before the activity had started because students were comfortable with the material and did not feel the need to practice it further. Furthermore, participants indicated students were often not comfortable with group activities if they were unfamiliar with the material. They mentioned an apprehension when they discussed new topics with others because they did not want to be observed as confused or lacking a thorough understanding of the material. This feedback led directly to the development of a new section of our observation protocol, which is intended to measure the uniqueness and novelty of the activity.

For body language, opinions of student engagement were often mixed among participants. A majority of the participants identified students sitting in a casual manner (e.g., slouching, lowered shoulders, erratic eye contact with the instructor) as disengaged. When we further questioned to obtain a consensus on this identifying factor, some students suggested that sitting in a casual manner was often a personal preference and not necessarily a sign of disengagement. However, a few participants pointed out that if the student was looking at a notebook or a calculator while sitting casually, then they would identify that student as being engaged. Focus group participants reported that the students leaning towards the other members and making eye contact appeared to be engaged in a relevant discussion. An attentive posture, where students were writing in their notebooks while looking at the instructor, was reported as an indicator of student participation in the activity. Students encouraged the research team to look into the context of what was being asked of students (e.g., working in groups with new material), rather than judging student engagement purely by facial expressions or body language.

Finally, focus group participants offered mixed reactions to the use of laptops or tablets during class. On one hand, participants stated that laptops and tablets were often used as a part of class discussion or assessment. For example, students would often access recently published notes from the instructor and use these to help them work through problem-solving exercises or discussions with groups. Additionally, instructors often used online platforms as a substitute for clicker technology, and would ask students to respond to multiple-choice questions using their laptops or tablets. On the other hand, participants indicated laptops and tablets offered a quick distraction to activities in class, and students would often use laptops or tablets to pretend like they were accessing notes, when they were engaged in off-task materials.

As a resolution for observing the engagement of students who were on laptops, participants recommended that observers pay attention to the length of time the student was engaged with their laptop and try to discern what was being displayed on the screen. They acknowledged that students who spent more than a minute or so looking at a laptop without engaging in eye contact with the class or their fellow classmates could be considered off-task. Participants also recommended that, if possible, observers sit at the back of the room often during observations to see how students were utilizing laptops or tablets during class.

Limitations

There are several limitations to the specific validation processes we conducted in our study. First, these students' characterizations of engaged and unengaged students may not necessarily be representative of the same behaviors other students exhibit when engaged or unengaged in the material. Unfortunately, our IRB protocol did not allow for students to view videos of themselves or peers in the same classroom to confirm such behaviors, which we would recommend for future analyses. Additionally, just because students are able to critique their own or other peers' behaviors in the classroom does not guarantee that they will be honest in their representations of the behaviors that are occurring. They could be harsher towards certain peers, while more lenient towards another group with whom they identify closely, or they could share information they feel researchers want to hear. To counter these limitations, we recommend that researchers work to establish rapport with their students beforehand to build a sense of trust in sharing honest answers about student engagement. For more information about establishing rapport with research participants, see Maxwell[4].

Discussion

Our observation protocol was designed to capture student reactions and instructor engagement during active learning instances. Focus groups with students (the group of primary interest in our observations) allowed us to validate whether or not our protocol and observer training materials effectively captured (1) the types of learning activities students experienced in the classroom, (2) the way instructors set up and engaged in these exercises, and (3) the way students reacted to active learning in the classroom. In doing so, we were able to triangulate[12] our protocol to what we had already observed in the classroom. Just as interview transcripts offer an opportunity to confirm results during the interview process[13-14], individual meetings and focus groups can also be utilized to confirm findings during observations.

Our focus groups yielded several additions to our protocol that had not been previously considered. For example, we did not recognize the impact that the uniqueness or novelty of the material might have on resistance to these active learning practices until students indicated that this was a feature overlooked during our protocol development. Additionally, we had not considered the importance of the observer's location on the effectiveness of the observation. Observers would need to be located at the front of the class to gauge facial expression or body language, but a location at the back of the class would allow them to observe the materials in which students were actively engaged. Furthermore, focus group participants indicated that, if someone was watching their behavior from the front of the classroom (i.e., they could see they were being observed), they might be more inclined to exhibit more compliant behaviors for the fear of what might be reported back to the instructor. All of these additions improved the validity of our protocol because we were able to ensure that our observation effectively captured what was actually happening in the classroom.

Our findings are limited in that we were not able to use videos from actual classroom lectures in which students might have been involved. The use of these videos would have likely led to confirmation of our shortcomings in our previous observations. For example, we could have selected a video that a trained observer had already rated and asked the group questions regarding what the observer had already found. However, there are potential shortcomings to this approach. If students were actively pictured in the video they were viewing or knew others who were in the video, they might not be willing to admit when they or fellow peers were engaged in off-task behavior. Furthermore, if participants had, themselves, utilized some strategy to remain off-task while appearing on-task (e.g., staring at a tablet pretending to be engaged in the instructor's notes), they might be less willing to admit to such behavior because they do not want to give away their "secrets."

Conclusions and future work

Although the study above gives only one example of how researchers can further validate classroom observations, this research portrays the impact that these types of "validity checks"[4] can have on a research study. Our findings in the student focus groups led to substantial changes in both our observation protocol as well as training documents for future observers in our research study. Conducting these focus groups also allowed us to confirm prior measures in our protocol, and effectively address the validity of our data collection measure.

As our observation protocol is only one piece of our study on student resistance, we plan to further validate this protocol through the use of a quantitative measure of student resistance currently in development. Our hopes in using this measurement tool in conjunction with our observation protocol is to provide engineering education researchers with multiple mixed methods tools to examine student resistance to active learning practices in the engineering classroom.

This material is based upon work supported by the National Science Foundation under DUE Awards #1347417, 1347482, 1347580, and 1347718. Any opinions, findings, and conclusions or recommendations expressed are those of the authors and do not necessarily reflect the views of the NSF.

References

1. Hammersley, M., & Atkinson, P. (2007). *Ethnography: Principles in practice*. Routledge.

2. Emerson, R. M., Fretz, R. I., & Shaw, L. L. (2011). *Writing ethnographic fieldnotes*. University of Chicago Press.
3. Mitchell, S. K. (1979). Interobserver agreement, reliability, and generalizability of data collected in observational studies. *Psychological Bulletin*, *86*(2), 376.
4. Maxwell, J. A. (2012). *Qualitative research design: An interactive approach* (Vol. 41). Sage.
5. Marcoulides, G. A. (1989). The application of generalizability analysis to observational studies. *Quality and Quantity*, *23*(2), 115-127.
6. Cronbach, L. J. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. John Wiley & Sons.
7. Mitchell, S. K. (1979). Interobserver agreement, reliability, and generalizability of data collected in observational studies. *Psychological Bulletin*, *86*(2), 376.
8. Gardner, F. (2000). Methodological issues in the direct observation of parent–child interaction: Do observational findings reflect the natural behavior of participants?. *Clinical child and family psychology review*, *3*(3), 185-198.
9. Putnam, H. (1992). *Realism with a human face*. Harvard University Press.
10. Becker, H. S., & Geer, B. (1957). Participant observation and interviewing: A comparison. *Human Organization*, *16*(3), 28-32.
11. Becker, H. S. (1970). *Sociological work*. Transaction Publishers.
12. Fielding, N. G., & Fielding, J. L. (1986). Linking data: the articulation of qualitative and quantitative methods in social research. *Beverly Hills (CA): Sage*, 41-53.
13. Bryman, A. (2003). *Quantity and quality in social research*. Routledge.
14. Lincoln, Y. S., & Guba, E. G. (1985). Naturalist inquiry. *Beverly Hills, CA: Sage*.
15. Dancy, M. H., Henderson, C., Rebello, N. S., Engelhardt, P. V., & Singh, C. (2012, February). Experiences of new faculty implementing research-based instructional strategies. In *AIP Conference Proceedings-American Institute of Physics* (Vol. 1413, No. 1, p. 163).
16. Finelli, C.J., Richardson, K.M., & Daly, S.R. (2013). *Factors that influence faculty motivation to adopt effective teaching practices in engineering*. Paper presented at the ASEE Annual Conference & Exposition, Atlanta, GA.
17. Froyd, J. E., Borrego, M., Cutler, S., Henderson, C., & Prince, M. J. (2013). Estimates of use of research-based instructional strategies in core electrical or computer engineering courses. *IEEE Transactions on Education*, *56*(4), 393-399.
18. Henderson, C., & Dancy, M. H. (2007). Barriers to the use of research-based instructional strategies: The influence of both individual and situational characteristics. *Physical Review Special Topics-Physics Education Research*,*3*(2), 020102.
19. Prince, M., Borrego, M., Henderson, C., Cutler, S., & Froyd, J. (2013). Use of research-based instructional strategies in core chemical engineering courses. *Chemical Engineering Education*, *47*(1), 27-37.
20. Weimer, M. (2002). *Learner-centered teaching: Five key changes to practice*. John Wiley & Sons.
21. Hora, M. T., & Ferrare, J. J. (2013). Instructional systems of practice: A multidimensional analysis of math and science undergraduate course planning and classroom teaching. *Journal of the Learning Sciences*, *22*(2), 212-257.
22. Piburn, M., Sawada, D., Turley, J., Falconer, K., Benford, R., Bloom, I., & Judson, E. (2000). Reformed teaching observation protocol (RTOP) reference manual. *Tempe, Arizona: Arizona Collaborative for Excellence in the Preparation of Teachers*.
23. Shekhar, P., DeMonbrun, M., Borrego, M., Finelli, C., Prince, M., Henderson, C., & Waters, C. (2015). Development of an Observation Protocol to Study Undergraduate Engineering Student Resistance to Active Learning. *International Journal of Engineering Education, (31)*2, 597-609.
24. Walkington, C., Arora, P., Ihorn, S., Gordon, J., Walker, M., Abraham, L., & Marder, M. (2012). *Development of the UTeach observation protocol: A classroom observation instrument to evaluate mathematics and science teachers from the UTeach preparation program.* Retrieved March 1, 2015 from http://cwalkington.com/UTOP_Paper_2011.pdf.
25. Murray, H. G. (1983). Low-inference classroom teaching behaviors and student ratings of college teaching effectiveness. *Journal of Educational Psychology*, *75*(1), 138.
26. Harris, A. H., & Cox, M. F. (2003). Developing an observation system to capture instructional differences in engineering classrooms. *Journal of Engineering Education*, *92*(4), 329-336.