Spring 5-2015

A Multi-Institution Investigation Of Educational Practices And Strategies In Stem Courses

Scott James Merrill
University of Maine

A MULTI-INSTITUTION INVESTIGATION OF EDUCATIONAL PRACTICES AND STRATEGIES IN STEM COURSES

by

Scott James Merrill

A Thesis Submitted in Partial Fulfillment
of the Requirements for a Degree with Honors
(Biology)

The Honors College

University of Maine

May 2015

Advisory Committee:
Michelle K. Smith, Assistant Professor, School of Biology and Ecology, Advisor
François G. Amar, Professor of Chemistry, Dean of the Honors College
Paula Lemons, Associate Professor
Karen Pelletreau, Lecturer in Biology Education
MacKenzie R. Stetzer, Assistant Professor of Physics

Abstract

This study examines the teaching practices of faculty participating in the Automated Analysis of Constructed Response (AACR) project. The AACR project focuses on using short-answer assessment questions to elicit the mixed models students have about key concepts in STEM courses. The 19 faculty from six different institutions who participated in this project are all teaching biology courses, asking biology AACR questions, and participating in Faculty Learning Communities (FLCs). FLCs are a method of faculty professional development in which groups of faculty regularly meet to discuss issues of teaching and learning. Here I use a combination of classroom observation data and surveys where faculty self-report on teaching practices to answer three research questions: 1) What instructional practices are currently being used by the AACR FLC faculty? 2) What instructional practices do AACR FLC faculty think that they are using in their courses? and 3) How closely do AACR FLC faculty's perceptions of their teaching align with their measured teaching practices? Results from the classroom observations show that instructors participating in FLCs utilize a variety of teaching practices ranging from lecture to collaborative learning. Survey data show that faculty self-awareness of their own teaching practices varies depending on the types of questions asked. Taken together, these data establish a baseline from which to monitor changes in teaching practices and self-perceptions of teaching practices of the FLC faculty as a result of their participation in the AACR project.

Acknowledgements

I would like to thank a number of people without whom this project would not be possible. First and foremost, I would like to thank my advisor, Professor Michelle K. Smith, for giving me the opportunity to work on this project and guiding me the entire way. Next, I would like to thank Dr. Karen Pelletreau for her help in guiding my presentation and writing in a coherent direction. I would like to thank Professor Paula Lemons for her help in guiding me through some of the qualitative survey data that her lab collected at the University of Georgia. I would also like to thank Dean François Amar and Professor MacKenzie Stetzer, in addition to Karen and Paula, for agreeing to serve on my Honors thesis committee.

Outside of my committee, I would like to thank all of the observers at the six institutions that conducted the observations of the FLC faculty. I would like to thank Jeremy Smith for the work that he put into processing the COPUS observation data and dealing with my constant issues with file extensions. I would like to thank Dr. Jill Voreis at the University of Georgia for her assistance in processing the survey data and for sorting through the demographic data. I would also like to thank my fellow members of the Smith lab that I have not already mentioned (Jonathan Dumont, Leif Johnson, Justin Lewin, Christopher Nashi, Dr. Mindy Summers, Darlene Turcotte and Erin Vinson) for providing feedback for my numerous presentations. Finally, I would like that thank my friends and family for their constant support throughout this long and arduous process. It has been a long road and I would not have made it to the end without the support of the aforementioned individuals!

Table of Contents

List of Tables

List of Figures

Introduction

A Need for Reforming Teaching

In 2012, there was a call to action from the White House emphasizing the need to increase production of college graduates with STEM (science, technology, engineering and mathematics) degrees by nearly one million over the next decade (PCAST 2012). Numbers from the PCAST report indicate that over three-quarters of the requested new STEM graduates could be produced by focusing on retaining the 60% of students who start in STEM majors but fail to graduate with a STEM degree. One factor that has been shown to affect retention rates is student experience in introductory courses (Seymour 2002). The first step suggested by PCAST towards improving STEM retention rates is to increase the use of research-validated instructional practices such as asking students to solve problems in class or working with peers to discuss questions. Current research indicates that students learn more and are less likely to drop out of STEM courses implementing research-validated active learning approaches (Freeman *et al.,* 2014). In spite of numerous research endeavors highlighting the benefits of reforming teaching practices, there has been little effort to document actual teaching practices of university faculty (Wieman and Gilbert, 2014; Smith *et al.,* 2014). This lack of information makes it difficult to document the current state of educational practices of university faculty, and monitor how instruction changes over time.

What Types of Professional Development Opportunities are Available for Faculty

Concern across the nation about maintaining our international STEM rankings has led to a tremendous interest towards changing the basic approach to teaching STEM courses (Henderson *et al.*, 2011). As a result of this interest in transforming STEM education, there are numerous resources designed to facilitate professional development that are now readily available. The modes of professional development that are currently in use are diverse in nature, including: workshops, seminars, mentoring programs and action research (Emerson and Mostellar, 2000). The varieties of available professional development materials were developed by different groups of researchers approaching the issue by focusing on the individual instructor and/or faculty environment and structure (Henderson *et al.,* 2011). Efforts targeting change of individual faculty are mainly focused on the development of reflective faculty and dissemination of curricular materials with pedagogy. Efforts to change environment and structure include developing a common vision for teaching and implementing policies that promote the use of new teaching practices (Henderson *et al.,* 2011).

Currently, there are a variety of formats being used to disseminate common curricular materials and pedagogical knowledge. This reform approach is centered on researchers identifying faculty who would benefit from reform and then showing them new ways to organize the curriculum and/or teach the subject of interest (Henderson *et al.,* 2011). A key aspect of providing instructors with new curriculum is to impress upon them the importance and value of the new materials and pedagogical strategies. The best curricular materials and pedagogical techniques in the world will have little impact on STEM education if the value of the new materials is not substantial enough to convince instructors to implement them in their courses.

Although disseminating "best practice" curricular materials as a part of talks and workshops is one popular mode of professional development, it falls short of the goal of changing teaching practices. For example the Silverthorn research group (Silverthorn et al. 2006), designed activity modules to help instructors implement active learning activities into their introductory physiology courses. They recruited numerous instructors at a summer professional development conference to use these new curricular materials in their courses the following semester. However, within a few weeks after the semester, over a quarter of the instructors who volunteered to use the modules had backed out. This attrition of volunteers continued until over half of the instructors had withdrawn from the study prior to the start of the semester. Of those who remained in the study, none used the modules and the attrition continued to increase. Numerous obstacles were listed by the faculty for why they failed to participate in the study, ranging from lack of time, increased class sizes, increased teaching loads, lack of support from leadership, etc. The researchers in this study also noted that those who dropped out of the study prior to the start of the semester were the instructors who had prior experience implementing active learning strategies in their courses. The attrition of instructors with experience in using active learning activities gave the perception that instructors who had already developed active learning activities of their own felt less inclined to use the new modules instead. Also, the instructor population that remained in the study consisted of those who had little to no experience using active learning strategies in their courses. The behavior of the instructors in this case emphasizes the point that it is difficult to encourage instructors to use active learning strategies in their courses without providing guidance.

Other research groups have also determined that producing common curricular materials and disseminating them is not sufficient to produce reformed teachers. Analysis of various professional development efforts have shown that simply providing the “best practice" materials is not effective for enacting changes in practice (Clark *et al.,* 2004; Henderson *et al.,* 2011). Furthermore, there is a perception in the STEM research community that assessment data showing the benefits of improved curricular materials is a sufficient argument in and of itself for the adaptation of the materials (Clark *et al.,* 2004). For example, Clark et al. (2004) studied professional development efforts at the University of Massachusetts-Dartmouth (UMD) and the Rose-Hulman Institute of Technology (RHIT). They noted that while the UMD group and RHIT group had similar goals; the overall outcomes were distinctly different at each institution. At UMD, the results of the pilot study of new curricular materials showed positive improvements in student learning according to assessments. However, these positive results did not lead to successful implementation of the new materials as a "data war" ensued where the assessment results were used by different faculty to argue for and against the implementation of the new curricular materials. Notably, there was significant backlash at this institution from the faculty that were not involved in writing the materials.

In contrast to the UMD case, the RHIT case showed that the adaptation of new curricular materials is made much easier by the inclusion of a large, diverse group of faculty in the development of the materials. At this group, they held weekly meetings for all interested faculty at the institution to attend. From the attendees, a faculty group was formed that designed the new curricular materials. Members of the committee then met with faculty in each department to present the materials and answer questions. Any questions and concerns raised by faculty at the presentations were addressed point by point by the research team. The perception with the RHIT case was that faculty that were not directly responsible for developing the materials felt more involved with the overall process when they were allowed to provide feedback and observe the changes their feedback enacted on the curricular materials. Both the UMD and the RHIT cases demonstrate the importance of convincing faculty and administrators that the revised materials are better than the old (Clark *et al.,* 2004). Deficiencies in available methods to advocate that new materials are better have severely limited the effectiveness of developing and distributing "best practice" curricular materials for the purpose of enacting changes in teaching practice (Henderson *et al.,* 2011). The weakness in advocating for new materials combined with the lack of voice given to instructors when new materials are being designed forms a complex issue that needs to be rectified in order to facilitate effective professional development in the future.

In addition to developing "best practice" materials, there are numerous research efforts that are working to develop instructors that are more reflective of their teaching practices. Many of these efforts rely on summer professional development workshops to facilitate change. In general, workshops tend to focus on the importance of implementing student-focused teaching practices in the course as opposed to teacher-focused strategies. The goal of these efforts is to change faculty beliefs about teaching to produce better learning outcomes for students (Connolly and Millar, 2006). The success of these workshops is typically measured based on self-reported data from the faculty. An issue with collecting data in this way is that many times instructors report their teaching practices to be distinctly different from what they actually implement in the classroom (Connolly and Millar, 2006). One study focusing on the results of one summer workshops noted that 89% of instructors who participated reported transforming their teaching practices towards active learning and student-centered instruction (Ebert-May *et* *al.*, 2011). Contrary to what the faculty self-reported, this study found that three-quarters of the participants in the summer workshop continued to use teacher-centered, lecture based teaching practices in their courses.

Focusing on student evaluation of faculty and quality of instruction, Davidovitch and Soen (2007) also noted the ineffectiveness of summer workshops in producing meaningful changes in teaching practices. Overall, instructors participating in the workshops had lower scores pertaining to overall evaluation and clarity of instruction compared to those who did not participate. In this study, faculty participated in extensive workshops led by senior professionals in academia with a specific focus on improving instructional quality. Combined, the research of Ebert-May et al. (2011) and Davidovitch and Soen (2007) suggest summer workshops could use improvement in an effort facilitate meaningful professional development.

As described above, professional development opportunities currently available do not always encourage changes in faculty teaching practices. Instructors cite numerous obstacles that prevent them from adopting active learning, student-centered teaching practices in their courses (Silverthorn *et al.*, 2006; Clark *et al.,* 2004). Those who participate in professional development workshops perceive that they have reformed their teaching practices to a much greater extent than the reality (Ebert-May *et al.,* 2011; Davidovitch and Soen, 2007). Despite the diverse approaches being taken to enhance teacher-centered teaching practices, there are clearly significant obstacles preventing reformation of teaching practice. In order to facilitate sustainable change, there needs to be a fundamental change in approach that focuses on supporting faculty *through* these professional development experiences with time and incentives (Andrews and Lemons, 2015).

One method that is distinctly different from dissemination of curricular materials and summer workshops and focuses on supporting faculty is the formation of Faculty Learning Communities (FLCs). FLCs have been documented as being effective in developing reflective teachers and increasing faculty interest in teaching and learning (Henderson *et al.,* 2011; Cox, 2004). Simply defined, an FLC is a small group of cross-disciplinary faculty and staff collaborating in an extended program to enhance teaching and learning (Cox, 2004). One way FLCs recruit members is to target faculty at the beginning of their careers. When initially taking up their first teaching position, most scientists have not extensively considered the available knowledge concerning pedagogy (Lynd-Balta *et al.*, 2006). The deficiency in pedagogical knowledge is a direct result of the emphasis of their doctoral degrees being discipline specific, with minimal training provided for formal teaching. Therefore, FLCs are advertised to new faculty as a method for them to expose themselves to more complex aspects of pedagogy. A unit such as an FLC functions to provide faculty the opportunity to explore pedagogy with their peers, which may be a novel experience for some participants.

Several studies have analyzed the impacts of FLCs on implementing successful professional development. In 2004, an entire volume of *New Directions for Teaching and Learning* was dedicated to FLCs and the growing body of research focusing on them. Numerous aspects of FLCs were considered in this issue, such as how to: develop FLC facilitators (Sandell *et al.*, 2004), manage numerous FLCs (Barton and Richlin, 2004), and assess FLCs (Hubball *et al*., 2004). The latter of those three issues, evaluating and assessing FLCs, address a point of great concern. It has been postulated that FLCs have positive outcomes for students and both pretenure and tenured faculty based on the results of self-reported surveys of FLC participants (Cox, 2004). However, there is not a consensus for the best manner in which to monitor the progress and outcomes of the FLCs relative to their respective goals. To date, there is little to no available literature that use direct observation data to study the effectiveness of FLCs in enacting change in teaching practices of the participants. One aspect of this thesis work is to measure, through observation, the effects of FLC participation on faculty.

Reformed Teaching in Large Enrollment Courses

A central theme to many efforts to reform undergraduate STEM education is to diversify the student experience in the classroom. The National Science Foundation (NSF) has provided substantial funding for numerous grants under its Widening Implementation & Demonstration of Evidence Based Reforms (WIDER) program (http://www.nsf.gov/pubs/2013/nsf13552/nsf13552.htm). The focus of these WIDER grants is to increase the use of evidence-based teaching practices in higher education, as recommended in the PCAST report on STEM education. My thesis project is a part of one WIDER grant, building on previous research completed by the Automated Analysis of Constructed Response (AACR) research group. The current AACR project is designed to provide faculty the opportunity to integrate formative assessment into their courses. As opposed to posing multiple-choice questions, the AACR project aims to provide instructors with large enrollment courses the option to pose short answer questions to students, sort the student answers using a computer program, and obtain feedback with sufficient time to use the results to inform their teaching. Figure 1 shows an example AACR question.

The AACR project can be broken down into three strands: 1) question development & computer analysis, 2) student learning & misconceptions, and 3) Faculty Learning Communities (FLCs, Figure 2). The first strand focuses on the development of AACR answer questions for the students and on improving the computer analysis technology that reviews the student responses. The instructor can then inform his or her teaching by using the results from student responses as formative assessment. The second strand focuses on the student responses processed by computer analysis and working to formatively assess the mixed models students have in the process of learning biology. The third strand focuses on the formation and development of Faculty Learning Communities (FLCs) at the six institutions participating in the AACR grant (Michigan State University, University of Maine, University of Georgia, University of Colorado at Boulder, University of South Florida, and Stony Brook University). For my project, I worked to characterize the teaching practices used in courses taught by the FLC faculty. Specifically, I measured the classroom activity in FLC courses and gauge the awareness that FLC faculty have for their own teaching practices.

The specific way in which these AACR FLCs are being run is unique in that observations are being conducted in courses taught by FLC faculty as they participate in the professional development. These FLCs are designed around the use of common AACR assessment questions and the use of the observation data to monitor changes in faculty teaching. The collection of observation data from the AACR-FLC faculty will directly address noted difficulties in measuring the progress of participants in professional development enterprises due to a lack of available observation data (Ebert-May *et al.,* 2011). The primary function of FLCs is to improve student learning by providing faculty with a forum in which to discuss their interests, challenges, and success with teaching (Cox 2001). In the AACR project, FLCs are led by principal investigators (PI) at each institution and consist of faculty who agreed to implement AACR questions in their course. Each group contains faculty who volunteered to participate in this effort to transform their teaching practices. The members of the FLC meet three to four times throughout the semester to discuss their courses and their use of the AACR materials.

Documenting and Surveying Teaching Practices

In starting the AACR research project, it was imperative to understand who the FLC members are as instructors. There are multiple ways to ascertain an understanding for who faculty are as instructors. One way is to gather qualitative data through surveys and interviews. Another method is to gather quantitative data on teaching practices and perceptions of teaching practices of the FLC faculty.

Several instruments have been designed to capture quantitative teaching practices data through classroom observations. When considering the available instruments to choose from, it is critical to consider the circumstance in which the instrument is going to implemented. Those that can be used reliably by a variety of observers and provide clear data pertaining to classroom activity measurements are preferable in large-scale situations such as the AACR project. The Classroom Observation Protocol for Undergraduate STEM (COPUS, Smith *et al.* 2013) is a protocol for gathering observation data and only requires a few hours of training to obtain a high inter-rater reliability (IRR). With COPUS, a trained observer can measure instructor and student activity. Compiled data provides instructors a complete picture for what they are doing and what their students are doing throughout a typical class period. COPUS has previously been used to describe the range of instructional practices present in a large number of STEM courses (Smith *et al.,* 2014). Using COPUS data, it is also possible to profile a given course as a subset of one of the four major instructional styles used in higher education (Lund *et al.,* 2015). These profiles are used as an extension of the COPUS data to provide a more simplified view of the teaching practices being employed by a given set of instructors.

There are also numerous instruments that have been used in other studies to ascertain the level of self-awareness instructors have for their own teaching. In the review of changes in STEM education literature, Henderson (2011) noted that one of the most cited works was that of Trigwell and Prosser (2004). These two researchers conducted a study to determine the relationship between approaches to teaching and teaching intentions. Their work resulted in the synthesis of the Approaches to Teaching Inventory (ATI; Trigwell and Prosser, 2004). The ATI is a 16-item survey scored on a 1-5 Likert scale, and it identifies the teaching strategy and intention used by the instructor for a given course. By administering the ATI to instructors prior to when they teach each semester, researchers can measure the extent to which the course of interest is taught with student-focused and teacher-focused strategies.

Another instrument used to assess instructor's awareness of their teaching practices is the recently developed Teaching Practices Inventory (TPI; Wieman and Gilbert 2014). A 72-item survey composed of objective questions; the TPI measures the extent to which a given instructor uses research-based teaching practices in their course. It was designed to be supplementary to the data being measured using the COPUS protocol by targeting elements of the course that are measurable both inside and outside of the classroom. Using the scoring rubric of the TPI, it is possible to quantitatively assess the extent to which an instructor employs research-based teaching practices in their course and compare different courses in this way (Wieman and Gilbert, 2014). Previously, the TPI has been used to document the wide variety of teaching practices being implemented at institutions such as the University of British Columbia and the University of Maine (Wieman and Gilbert, 2014; Smith *et al.,* 2014). It was also found to have strong correlations with classroom observation data collected using COPUS (Smith *et al.,* 2014). Simply put, the TPI has been shown to be effective in documenting the alignment between classroom activity and instructors perception of classroom activity in STEM courses.

Perceptions and Assessment of Teaching Practices

I have used quantitative data to address the following three major research questions: 1) what instructional practices are currently being used by the AACR FLC faculty, 2) what instructional practices do AACR FLC faculty think that they are using in their courses and 3) how closely do AACR FLC faculty's perceptions of their teaching align with their measured teaching practices? The results provide a solid foundation from which to track changes in instructional practices over the course of the project. Collection of these data allows us to measure the effectiveness of professional development and to better target the needs of FLC members. Significant emphasis is placed on relating the data collected with COPUS and from the ATI and TPI to see if FLC members' perceptions of how they teach their courses align with the classroom activity for both the students and themselves. I also utilized additional demographic data for each course taught by FLC members, such as teaching experience and class size, to determine if these factors impact their instructional practices. I conclude with a discussion of what these data indicate about the initial profiles of the FLCs as a whole and how continued participation in the FLCs could impact the profiles of FLC members moving forward.

Methods

Selection of Observers from Participating Institutions

The primary objective of my thesis was to collect data relating to instructor and student classroom activity in the STEM courses taught by AACR faculty using the Classroom Observation Protocol for Undergraduate STEM (COPUS) (Smith *et al*., 2013). COPUS is a protocol that allows observers to objectively classify classroom activities using 25 codes. Observers record the activities in which the instructor (13 codes possible) and students (12 codes possible) engage in two-minute increments for the duration of a class period. A sample of the COPUS protocol is shown in Figure 3. A list of the COPUS codes and their definitions is shown in Table 1 for instructors and for students. Individuals participating as observers in the AACR project range from undergraduate students to post-docs. To date, I have collected and compiled data for 27 courses taught by 19 different FLC members during 2014.

Training of COPUS Observers

All observers from the six participating institutions were simultaneously trained in the use of the COPUS at the onset of the Spring 2014 semester using an online training protocol. The training period consisted of an introduction to the COPUS protocol, an explanation of the explicit meaning of each code, and the general methods for data collection. Following the introduction, observers practiced coding videos of STEM courses. After watching two-minute intervals of the videos, observers discussed the selection of codes with both peers at their institution and with the broader online group participating in the training.

When collecting COPUS data from multiple institutions for analysis, it is important ensure that the protocol is being implemented in a uniform manner. Therefore, at the conclusion of the initial training session, observers were instructed to independently code Video #1 (a 8:22 STEM lecture video, http://harvardmagazine.com/2012/02/interactive-teaching). Their results were entered into a COPUS excel sheet and submitted to us at UMaine for analysis. The data were processed for a pairwise comparison, and we calculated the inter-rater reliability (IRR), which measures agreement between raters, using SPSS (IBM, Armonk, NY). Specifically, the IRR Cohen's kappa score was calculated for all possible observer pairs, and then those values were averaged to obtain the kappa score for the group as a whole (Cohen, 1960). Cohen's kappa scores range from 0-1.0, with larger values corresponding to greater agreement between observers and lesser values attributing agreement to chance. The target average kappa value for observers in the program is $\kappa > 0.80$, which indicates "almost perfect" agreement between raters (Landis and Koch, 1977). The average κ score for the first assignment was calculated to be 0.696. While this kappa value indicates substantial agreement between observers, it does not exceed the target value. Therefore, additional training was deemed necessary to obtain sufficient IRR.

For the next round of training, we generated a heat map of the codes selected by the observers for Video #1. A sample heat map of results from Video #1 can be seen in

Figure 4. Heat maps summarized the 25 COPUS codes used in each two-minute interval selected by all the observers. A white-black color gradient corresponds to the frequency with which a code was selected. The greater the number of observers that selected a given code in a given interval, the darker that square will appear on the heat map. For example, "L" (listening) was coded by all observers for the 8-10 minute interval and therefore is represented as a black box on the heat map (Figure 4, row 8, column L).

To improve the IRR scores, Video #1 was observed again as a group in two-minute segments. At the end of each two-minute segment, the group discussed the heat map results, and went over any discrepancies between their codes and the heat map. Any codes that were improperly coded were discussed in-depth. An additional video (Video #2, 10:00 STEM lecture video) was used for practice during the training. As before, the video was coded in two-minute segments followed by intra-institutions and inter-institution discussions while comparing individual responses to the answer key.

Upon completion of the second training session, Video #3 (a 14-minute video STEM lecture video) was assigned for them to code individually. As with Video #1, the COPUS results from this video were used to calculate the IRR. The average kappa score was calculated to be 0.871, indicating "almost perfect" agreement between raters (Landis and Koch, 1977). A kappa score greater than the target value of 0.80 indicated that the project was ready to proceed to the next step. A heat map of results was disseminated to the observers along with notes about any codes that were not unanimously selected by the group.

To ensure this IRR would transfer to a full-length lecture, the observers coded Video #4 (a 52-minute STEM lecture video) to simulate the experience of coding a full-length class period. An average κ score of 0.830 was calculated using SPSS. This kappa score indicates again "almost perfect" agreement between the observers for this assignment, and was consistent with what was observed with the shorter video assignment. Observers were then instructed to conduct live observations of classes taught by faculty in the FLCs at their institutions.

At the start of the Fall 2014 semester, a new round of training was conducted for new observers and the entire group. For the new observers, we discussed the COPUS codes and had them practice coding videos in two-minute segments like other first time training sessions. Following the first session, Video #2 was assigned for homework for both the new and returning observers. Results from this homework assignment were used to calculate an average kappa score of 0.835, again indicating "almost perfect" agreement between observers. Protocols from the homework were processed to generate a heat map for the next training period. A second training session was held where the heat map of Video #2 was reviewed by each two-minute time interval. Video #5 (another 10:00 STEM lecture video) was coded for additional practice and discussion. At the conclusion of the second training session, the observers were instructed to code Video #6 (a 20-minute STEM lecture video) as an individual assignment. We analyzed the results from Video #6 and calculated an average κ =0.80. Because this kappa score was right at the target level, approval was given by the PIs to conduct live observations. The one observer who submitted results from the assignment that were below the average was provided one-on-one training to clear up the misconceptions they had with the COPUS protocol.

Conducting Classroom Observations

For the Spring 2014 semester, observers were instructed to observe each faculty member at least twice to obtain baseline data for their instructional activities (Table 2). The purpose of conducting multiple observations was to ensure that the data are an average of the classroom activity of a course and not one-time phenomena.

During the Fall 2014 semester, the observation plan was revised. Each faculty member was to be observed three times (Table 2). Two of those observations, one early and one late in the semester, would contribute to that instructors baseline data. The third observation was to be conducted on the day an instructor was discussing an AACR question.

Processing Observations

With observation data being submitted at infrequent intervals throughout the semesters, it became increasingly important to keep a detailed observation record of what work had been completed. A record was maintained for each semester detailing the dates of observation, FLC faculty member observed, observer name, COPUS protocol file name, and whether a pair of observers or a lone observer conducted the observation.

Two pie charts were generated from observation(s) of a single class period: one detailing the percentage of codes coded for instructor activity and one for the percentage of codes coded student activity for that class period (Figure 5). For example, if a total of 50 codes were coded for an instructor during a class period, and 20 of those codes were for lecturing (Lec), then the instructor pie would be 40% lecture. In trying to compare observations, it is difficult to get a general sense of what both the students and instructors are doing when considering all 25 COPUS codes individually (Smith *et al.* 2014). To overcome this difficulty, the 25 COPUS codes were collapsed into four categories of instructor activity and four categories of student activity (Table 3).

When a pair of observers observed a class, the COPUS data required an extra step with regards to processing. The protocols generated by each observer were analyzed and the codes agreed upon by both observers (ones that they both coded or both did not code) generated the code count for that class period. For example, if observer 1 coded lecture for the first 24 intervals of 25 intervals and observer 2 coded lecture for only the last 24 intervals, that observation would reflect lecture being coded 23 times. If the observers had 25 codes agreed upon for instructor activity, then that instructor pie would reflect lecture 23/25, or 92% of instructor activity. Moving forward with data analysis, the refined count was used to represent the observations of any class by a pair of observers.

Once collapsed code counts from observations of single class period were calculated, data from multiple observations of the same instructor were averaged. For example, if an observation of one class period had coded presenting 25 times out of 30 total instructor codes and an observation of a second class period coded presenting 23 times out of 32 instructor codes, then the percent "Presenting" for that instructor would be 48/62, or 77%.

To provide a more global view of the COPUS data, we collaborated with researchers from the University of Nebraska to use a clustering analysis method for analyzing COPUS data (Lund *et al.*, 2015). This new method of analysis considers five instructor COPUS codes (Lec, RtW, FlUp, CQ, MG) and five COPUS student codes (CG, WG, OG, SQ, AnQ) as eight cluster codes (student CQ, WG, and OG condensed to a new group work GW code) to describe the instructional style for a course. The profile method differs from using the collapsed COPUS codes in that it looks at each code a percentage of possible time blocks during which it could been coded. For example, if an observer coded an instructor lecturing for 24 intervals and real-time writing for 10 intervals for a 25-interval class, then the instructor would be profiled as lecturing for 96% of class and real-time writing for 40% of the class. Reducing the scope of the original COPUS data, similar to what was done with the collapsed codes, allows for the classification of the course into four statistically different clusters pertaining to the most common instructional styles used at the college level.

Perceptions of Teaching

To complement the classroom observation data from COPUS, each participating FLC member completed a survey at the beginning of the semester. The survey asked some demographic information on the course being taught during the forthcoming semester. The demographic information was used to give a more descriptive look at what types of instructors participate in FLCs, their years of teaching and their professional development experience. The survey also included the Approaches to Teaching Inventory (ATI) (Trigwell and Prosser, 2004). The ATI was administered again to participants in the fall semester of 2014. This time, the survey (Fall 2014) also included the Teaching Practices Inventory (TPI) (Wieman and Gilbert, 2014). Both surveys are designed to identify faculty member's perceptions of their teaching.

Approaches to Teaching Inventory

The ATI is designed to show the relationship between a teacher's approach to teaching and student's approaches to learning (Trigwell and Prosser, 2004). The instrument measures a teacher's focus in a particular classroom context, i.e., teacher-focused or student focused, as well as a teacher's intentions for a particular classroom, i.e., information-transfer or conceptual change. In the development of the inventory, it was found that a teacher-focused strategy was always paired with an information-transfer intention. Conversely, a student-focused strategy was consistently paired with a conceptual-change intention. These strategy-intention pairings form the two scales of the inventory: information-transfer/teacher-focused (ITTF) and conceptual-change/student-focused (CCSF). The inventory consists of sixteen items, with eight items for the ITTF scale and eight items for the CCSF scale (Table 4). Each item is graded on a 1-5 Likert scale, with the sum of the eight items responses being the score for that given item. Two ATI scores: the ITTF score and the CCSF score (min = 8, max = 40 for either section) are reported for each faculty member. This numerical score provides a reference point from which changes can be monitored over time, provided that the context is the same.

Teaching Practices Inventory

The TPI is designed to measure the extent to which research-based teaching practices are being used in a given STEM course (Wieman and Gilbert, 2014). It consists of 72 items that are broken down into eight categories. Using the scoring rubric, it is possible to assign a numerical score to quantify the use of research-based instructional practices. As with the ATI, the TPI numerical score provides a point of reference, which can be tracked over time to monitor changes in the practices of a given faculty member. Breaking down the scores by the category, it is also possible to target practices that are lacking in a given course. For comparison with the COPUS data, the category III of the inventory, "In-class features and activities," is of particular interest due to the clear parallels between the nature of the inventory questions and data collected via COPUS.

For example, this category of the TPI asks instructors questions such as how often they conduct reflective activities, lecture in class, and have students give presentations. Each faculty member who completes the TPI will have a total score (max = 72) that can be broken down into multiple category scores (ex. my thesis emphases category III, max score = 15).

Results

Documenting Instructional Practices of FLC Members

Using COPUS, I documented the instructional practices of 19 FLC members during the Spring 2014 and Fall 2014 semesters. Observation data of the collapsed COPUS codes from both semesters are shown in Figures 6 & 7 for instructor and student activity, respectively (for codes, see Table 3). It is important to note that these data, although presented as a percentage, do not represent the percentage of time in the classroom engaged in an activity, but rather the percentage of the codes assigned. With the instructor activity (Figure 6) there is a wide continuum of percent of "Instructor Presenting," ranging from 6%-100%. Furthermore, as the percentage of "Instructor Presenting" decreased, the percentage of "Instructor Guiding increased." For every course observed, a small amount (<10%) of codes pertained to administration and other activities.

Data from all instructors is identified using pseudonyms (Figure 6, y-axis) and includes a note of which semester during which the course was taught in 2014. Some instructors were observed over multiple semesters, for example "Allison" and "Kate" (Figure 6). When comparing different semesters taught by the same instructor, most had similar profiles (e.g., "Kyle, Laura,") so they clustered near each other on the graph. However, some instructors had noticeably different profiles from one semester to the next. The change in profile from one semester to the next can be explained by instructors teaching different courses (e.g., "Doug") or by instructors altering the structure of the course, such as flipping the classroom (e.g., "Allison" who taught a traditional class in Spring 2014 and a flipped class in Fall 2014).

With student activity collapsed codes (for codes, see Table 3), a similar, yet more narrow, continuum was present with respect to percentage of "Student Receiving" (Figure 7). Arranged from greatest to least percent "Student Receiving," courses observed in 2014 ranged from 35% to 100% "Receiving" by the students. As with instructor collapsed codes, students collapsed codes allowed for a glimpse of what students were doing in each course. A decrease in "Student Receiving" was complemented by an increase in the percent of "Students Talking to Class" and "Students Working" in class. It is worth noting that "Students Talking to Class" code was prevalent across the observations (Figure 7). However, the "Students Working" collapsed code was not as widespread or as consistent along the continuum compared to the "Students Talking to Class" code. Students in courses with lower amounts of "Instructor Presenting" spent more time working in class (Figure 7). All of these data indicate great diversity of classroom activity in courses taught by FLC members.

Profiling Instructors by Cluster Analysis Method

Using COPUS data, it is possible to profile a given course as a subset of one of four major instructional styles. As described recently by Lund and colleagues (2015), these profiles are: Lecture, Socratic, Peer Instruction and Collaborative Learning. A profile is generated by completing a cluster analysis looking at multiple COPUS observations as a percentage of time blocks as opposed to as a percentage of codes assigned. The profiles are reflected in Figure 8, showing the placement of each course (n=26) in the cluster analysis (Lund *et al.,* 2015). Several instructors' data from the FLCs placed them in between two instructional styles, so they were classified as "hybrid" groups to describe instructors that exemplified aspects of two major profiles of teaching practices. For example, if an instructor had observations classified as Lecturing (Lec) and Socratic (Soc), then they would fall into the "Lec/Soc hybrid" cluster. The distribution reflected in Figure 8 shows a range of profiles with the "Lecture" (n=5) and "Peer Instruction" (n=7) clusters being the most common. For the hybrids, the most common by far was the "Peer Instruction/Collaborative Learning" hybrid (n=7). There is one course from the 2014 data set that is not reflected in the profile histogram because it is classified as a "N/A" profile using the cluster analysis. Six additional observations profiled as "N/A", but additional observations from a different date generated a profile that was used for classification. It is important to note that the COPUS profile model is a statistical model that, as with all models, has some degree of inherent error. The intrinsic error is what is preventing all observations from being neatly profiled by the cluster analysis. The "N/A" problem could be resolved through further revisions of their model.

Ascertaining FLC Member Self Awareness of Own Teaching Practices

To measure the FLC member's awareness of their own teaching practices, each faculty member was asked to complete a survey at the beginning of each semester during which they were being observed. Organizing the ATI scores from lowest ITTF to score to highest ITTF score, (Figure 9), the faculty with the lowest ITTF scores had the highest

CCSF scores, as would be expected. However, the faculty with the highest ITTF scores had roughly numerically equal CCSF scores. These results suggest that FLC members exist along a continuum ranging from courses that are very student centered to those that are both instructor and student centered.

As the ATI is designed to capture the strategy and intention used by the instructor, I correlated the ATI ITTF score to the eight COPUS collapsed codes (Table 3) to see how well these two instruments aligned. Correlations were completed between these two data sets to see if there was any relationship between the two metrics. Overall, ITTF score was a poor predictor of classroom activity measured by COPUS ($p>0.05$, Figure 10). The only notable exception is the significant negative correlation ($p<0.05$) between the COPUS collapsed code “Students Talking to Class” and the ITTF score (Figure 11), indicating that faculty with lower ITTF scores are more likely to have their students talk in class. The remaining collapsed codes show weak non-significant trends in the direction that would be expected. For example, it would be expected that the collapsed “Instructor Guiding” code to be negatively correlated with ITTF score. In other words, faculty who have higher ITTF scores are less likely to use “Instructor Guiding” behaviors such as asking clicker questions, and moving and guiding throughout the classroom.

A similar regressions analysis was run to correlate the ATI CCSF score to collapsed code COPUS data (Table 5). It is expected that the more student focused the course (higher CCSF score), the less prevalent “Instructor Presenting” and “Student Receiving” codes. All five collapsed codes showed the hypothesized positive or negative correlation with CCSF score. However, all of the correlations were weak and not significant ($p > 0.05$). Therefore, the CCSF category score of the ATI is a poor predictor of classroom activity measured by COPUS.

The TPI was added to the instructor survey in the Fall 2014 in order to find out more about which instructional practices the instructors were using in their classes. The total survey scores and scores from category III ("In-class Features and Activities") are shown in Figures 12a and 12b, respectively. Total TPI scores (Figure 12a), arranged from lowest to highest total score, showed a wide range of scores (30-49) with respect to the maximum total score possible (72). Category III scores (Figure 12b) showed a similarly wide range (4-12) with respect to the maximum score for the category (15).

To investigate the relationship between TPI scores and COPUS collapsed codes, I correlated the total TPI score against the four COPUS collapsed codes. Designed to measure the extent of use of research-based teaching practices in a course, the total TPI scores were expected to have positive correlations with the following collapsed codes: 1) "Instructor Guiding," 2) "Students Talking to Class," and 3) "Students Working." These positive correlations are expected because of the implied positive relationship between using research-validated teaching practices (higher TPI scores) and more active classrooms (more "Instructors Guiding," "Students Talking to Class," and "Students Working"). Conversely, I hypothesized the total TPI score would have a negative correlation with the "Instructor Presenting" and "Students Receiving" collapsed COPUS codes. These negative correlations are expected because of the implied negative relationship between lesser use of research validated teaching practices (lower TPI scores) and less active classrooms (more "Instructor Presenting" and "Students Receiving"). Percent of "Instructor Presenting" had a negative correlation with total TPI score, albeit a weak ($R^2 = 0.22$) and statistically non-significant ($p > 0.05$) one (Figure 13). Results from correlations of the five collapsed COPUS codes of interest are reflected in Table 6. None of the five correlations of total TPI score to collapsed codes were statistically significant. Overall, this result indicates that the total TPI score is not a strong statistically significant predictor of classroom activity measured using COPUS. This result makes sense given that the TPI asks instructors about several features outside the class period such as the number exams, frequency of homework assignments etc. and the COPUS is focused only on measuring in class behaviors.

Because the TPI category III focuses on in-class behaviors, I next examined the relationship between the COPUS collapsed codes and TPI category III scores. I hypothesized that the collapsed COPUS codes would correlate with the category III score in the same way that they did with the total TPI score. Looking at the instructor collapsed codes, “Instructor Presenting” (Figure 14) had a strong negative correlation ($R^2 = 0.59$) with TPI category III score that is statistically significant ($p < 0.05$). “Student Receiving” (Figure 16) also had a statistically significant ($p < 0.05$) negative correlation ($R^2 = 0.63$) with TPI category III score. The “Instructor Guiding” score (Figure 15) and “Student Working” score (Figure 18) both had strong positive correlations ($R^2 = 0.62$ and 0.65, respectively) that are statistically significant ($p < 0.05$). The only collapsed code that did not have a statistically significant correlation with TPI category III score is “Students Talking to Class” (Figure 17), which had a weak negative correlation ($R^2 = 0.11$). In general, this result indicates that the TPI category III is a strong predictor of classroom activity measured using COPUS, with the exception of the prevalence of “Students Talking to Class.”

Assessing Impacts of Class Size on Teaching Practices

Several studies have shown that faculty often cite external factors that influence whether or not they can teach in an interactive manner, including class size (Silverthorn *et al.,* 2006; Clark *et al.,* 2004). In an effort to determine whether class size impacts instructional practices in courses taught by FLC faculty, class sizes were plotted against COPUS data. There was a weak ($R^2 = 0.00070$) negative correlation between "Instructor Presenting" and class size that was not statistically significant ($p > 0.05$; Figure 19). Furthermore, all five COPUS codes (Table 7) were not significantly related to the class size of the course being observed. This result indicates that class size does not greatly impact the diversity of activity being observed using COPUS.

In terms of relating the use of research-based teaching practices and class size, the TPI category III scores were also correlated to class size figures. Figure 20 shows the weak negative correlation ($R^2 = 0.0017$) that existed between these two variables. This regression was also found to be non-significant ($p > 0.05$), further supporting the claim that class size does not restrict the variety of teaching practices implemented in courses taught by FLC members.

Investigating Impacts of Teaching Experience on Teaching Practices

Using demographic information collected in the surveys, I related COPUS observation data to the years of teaching experience of each instructor. The percent of "Instructor Presenting" was weakly correlated with teaching experience ($R^2 = 0.013$) but was not statistically significant (Figure 21). All five collapsed codes of interest (Table 8) had statistically non-significant ($p > 0.05$) correlations with teaching experience. These data indicate that teaching experience does not alter teaching practices in a manner measurable using COPUS.

Discussion

Characterizing teaching practices across institutions for professional development

The goal of my thesis was to characterize the teaching practices of faculty participating in the FLCs of the AACR project (Figure 2). Specifically, I wanted to measure the teaching practices used in the classrooms and document the perceptions instructors have for their own teaching practices. Additionally, I wanted to explore how factors like class size and teaching experience affect both classroom activity and perceptions of teaching. Using COPUS, I was able to show that there were a variety of teaching practices (Figure 6) and instructional styles (Figure 8) being used by FLC faculty in their courses. In surveying for faculty perceptions, I showed that the ATI was a poor predictor of classroom activity within FLC courses (Table 5 & Figure 10). The TPI was a much stronger predictor of classroom activity (Table 6), especially when relating the "In-Class Features and Activities" category score to COPUS data (Figures 14-18).

What teaching practices do AACR FLC faculty utilize?

In the AACR project, FLCs are being used as the agents of change for professional development of interested faculty. As previously noted, FLCs are but one of many options for professional development that are available to faculty. FLCs were selected as the mode of professional development in the AACR project because they have been previously documented as being effective in developing reflective instructors

(Henderson *et al.,* 2011). Two goals of this project are to 1) determine if FLCs continue to be beneficial when there are multiple FLCs collaborating across institutions and 2) to provide stronger evidence that FLCs are effective in effecting change. Difficulties have been noted in other studies related to lack of means to measure the success of professional development, such as an independent assessment of classroom practice (Ebert-May *et al.,* 2011). In order to be able to measure the progress of FLC members over time, I used a combination of surveys and classroom observations to establish a baseline characterization of the teaching practices of AACR FLC faculty.

Many publications to date view teaching practices used by instructors in higher education as a contrast between either traditional lecturing or active learning, which has been noted to be counterproductive in facilitating professional development (Smith *et al.*, 2014). The AACR faculty appear similar to the University Course Observation Program faculty studied by Smith *et al.* (2014) based on the wide continuum of "Instructor Presenting" ranging from 6%-100% (Figure 6). The diversity in "Instructor Presenting" is complemented by variety in "Instructor Guiding" students. Similarly, you can clearly see that there are a variety of instructional styles used by AACR faculty using the COPUS cluster analysis method (Figure 8). These profiles aim to place instructors along a continuum of instructional styles ranging from teacher focused (lecture) to student focused (collaborative learning). All variety of styles were represented by the faculty in the FLCs, with a large number of instructors profiled as utilizing "Peer Instruction" and "Peer Instruction/Collaborative Learning Hybrid." In this sense, it is quite clear that we cannot neatly categorize AACR faculty into the two canonical groups often used to describe teaching practices.

One concern at the outset of the AACR project was that because the faculty members who volunteer to participate in the project are already incorporating active learning into their classes, the conclusions from the project could not be broadly applied. The continuum of instructor activity (Figure 6) and COPUS profiles (Figure 8) clearly indicate a variety of instructional styles are used by faculty in the project, and can directly alleviate this concern.

To further investigate factors that might generate the continuum of instructor activity, course demographic data was compared to the classroom activity data. I found that class size did not restrict the diversity of classroom activities in courses taught by FLC faculty. Class size has been cited both anecdotally and in literature as a significant barrier to encouraging classroom activity (Murray and Macdonald, 1997; Silverthorn *et al.,* 2006; Clark *et al.,* 2004). Instructors note that as class size increases, they feel forced to lecture due to the difficulty and impracticality of implementing active learning strategies in these larger classes. The findings from the FLC data directly contradict the notion that large class size prevents teaching in an interactive manner. In the FLCs, we see a statistically non-significant correlation between class size and all COPUS collapsed codes, such as "Instructor Presenting" (Figure 19 & Table 7). Furthermore, these data indicate that some FLC faculty with large enrollment courses teach in very active ways while other FLC faculty with small enrollment courses teach with a more teacher-focused style.

In addition to tracking class size, I also obtained the years of teaching experience for each instructor in the demographic sections of the surveys. In the literature, there is a perception that professional development needs to be targeted to new teachers (Barlow and Antoniou, 2007). Correlations between measured classroom activity and teaching experience (Figure 21) showed there was no statistically significant correlation between the COPUS classroom observation and class size demographic data sets from the FLCs. The lack of correlation indicates that AACR faculty do not follow a clear developmental timeline transitioning from teacher-centered to student-centered in their instructional practices. Within the FLCs, there were faculty members with very little experience who taught in very active ways along with experienced faculty who taught in highly teacher-focused manners. If the AACR-FLC faculty are a representation of types of university faculty, the lack of correlation between teaching practices and teaching experience indicates that when designing professional development, faculty at all career levels should be included.

One way to ascertain if the FLC faculty are representative is to compare COPUS observations to see if similar patterns in teaching practices are observed. Data collected at the University of Maine in the University Classroom Observation Program (UCOP), resulted in a similarly wide continuum of instructor activity seen with both the student and instructor COPUS data (Smith *et al.* 2014). Observing instructors from a variety of disciplines, as opposed to the primarily biology instructors of the AACR project, the authors of this study saw similar patterns in terms of a continuum of "Instructor Presenting." However, the UCOP data showed a lack of statistical significance between course size and classroom activity that was seen in with the FLC data. The UCOP program found a significant, but not large positive correlation between class size and "Instructor Presenting" collapsed code. This correlation is an interesting result indicating that patterns related to classroom activity seen within one discipline at multiple institutions are similar, yet not identical, to those obtained across an entire public research-intensive institution like the University of Maine. When considering the scope of the AACR project, the differences with the UCOP data have positive implications for the work of the FLCs in terms discounting obstacles that have been encountered in other studies.

How aware are faculty for their own teaching practices?

To define the pedagogy FLC members apply to instruction, I used several surveys to document their perceptions of how they teach their courses. Using the Approaches to Teaching Inventory (ATI), I investigated the extent to which instructors think their courses are teacher or student centered. The relationship between ATI scores and COPUS data suggests that the instructors' intention for the course does not align with classroom activity on a typical day. Research implies that how faculty perceive classroom activity in their courses is distinctly differently from what actually occurs (Ebert-May *et al.*, 2011; Fung and Chow, 2002). Correlating the COPUS observation data to the ATI scores, this perception holds true with respect to ATI survey data. Of the ten correlations of ATI scores with COPUS data, only the ITTF score has a statistically significant correlation with a single COPUS collapsed code ("Students Talking to Class," Figure 8). From the FLC population, I saw a variety of strategies being used, ranging from heavily student focused to a mixed focus on student and teacher (Figure 9). The variety of intentions shown in the ATI data parallels the variety seen in teaching practices measured using COPUS, but the lack of alignment calls into question the usefulness of the ATI to survey instructor self perceptions of teaching.

In an effort to complement some of the teaching philosophy data collected with the ATI, the decision was made to implement the Teaching Practices Inventory in addition to the ATI. This newly published instrument is designed to complement COPUS data and also addresses some other pedagogical aspects of the course through a series of objective questions (Wieman and Gilbert, 2014). Instructors with higher TPI scores use a variety of research-based teaching practices in their courses. Comparing the total TPI score to COPUS, there were no statistically significant correlations. However, when the scope of TPI data being correlated was refined to include only the "In-Class Features and Activities" category III score, the TPI had statistically significant correlations with four of the five collapsed codes measured via COPUS (Figures 14-17). The only collapsed code that did not have a statistically significant correlation with TPI category III score was "Students Talking to Class," which had significant correlations with the ATI ITTF score. Not only did the TPI category III score have statistically significant correlations with COPUS collapsed codes, but it also confirmed implications made using COPUS data about the effect of class size on instructor activity.

Correlating class size to TPI category III score (Figure 20), I was also able to show that there was no statistically significant relationship between the number of students in the class and the extent to which instructors use research-based teaching methods in their course. The lack of correlation between both the measured (COPUS) and perceived data (TPI) to class size is an intriguing find with respect to designing professional development. In prior research, instructors remarked that class size, among other things, was an obstacle to implementing student-centered techniques in their courses (Silverthorn *et al.*, 2006). Few studies have used actual correlations of teaching practices to class size as an argument in favor of student-focused practices. Therefore, the class size data could be used to support the use of student-focused materials in future professional development.

Overall, we saw that the TPI category III score was the strongest predicator of classroom activity measured using COPUS. It is the only instrument that we used that produced statistically significant correlations with COPUS data. In general, the total TPI score and ATI scores lacked statistically significant correlations with COPUS data. One explanation for the limited application of the ATI may be that the instrument is designed to measure constructs like knowledge, conceptions, and thinking rater than classroom behavior. With the total TPI score, the lack of correlation may be due to the scope of the questions being asked. Many aspects outside the physical instructing of students are included in the total TPI score, such as asking what kind of supporting materials they provide for students or what kind of feedback they give students in their course. Furthermore, by reducing the scope of the TPI data to the category III "In-class Activities" score, we were able to directly compare faculty responses to questions such as "what percentage of time they spend lecturing" to the observation results from the COPUS protocol.

Using this combination of classroom observation and survey data provided a more complete picture of the FLC faculty. Without the COPUS data, there would be no way to confirm the activities instructors and students are engaging in during class time. Without the TPI, there would be no data that can sufficiently gauge the instructor's perception of how they structure their courses (total score) and the specific teaching practices they employ (Category III score). Using these two instruments together, you can effectively show the clear relationship between classroom activity and instructor perceptions of classroom activity. Additionally, both of these data sets indicated that classroom activity was not restricted by class size, in direct contradiction to previously noted research.

Impacts of AACR FLC data on Future Professional Development

My work with the FLCs has been focused on the collection and analysis of quantitative data using the COPUS, ATI, and TPI instruments. As I stated in the introduction, there are more ways to understand who the FLC members are as instructors aside from quantitative data. Researchers in the University of Georgia FLC led by Professor Paula Lemons are working to analyze interview transcripts from interviews conducted with members of each FLC during each semester of teaching. The interview data can be used to expand the picture that we have of each FLC faculty as we use faculty's own words and thoughts to further refine our understanding of their position upon entering the project. In conjunction with the COPUS cluster analysis method, we can use the interview data, survey data, and profiles to tell a complete story of the evolution of each instructor's teaching practices over the course of the project.

Moving forward with the AACR project, FLC faculty are now starting to design common instructional activities to be used to teach the material assessed by AACR questions. My project has placed a heavy emphasis on collecting observation data and surveying the faculty it the inception of their participating in the FLCs. Literature clearly states that the synthesis and dissemination of curricular materials to other faculty members is insufficient in propagating change in instructional practices (e.g. Clark *et al.* 2004). Therefore, the AACR project needs the observation data to monitor the changes that take place during the development of the new curricular materials. Reviews of STEM education practices have emphasized the need for reform efforts to last over long periods of time, involve evaluation and feedback of use of the curricular materials, and be explicitly focused on changing faculty's perceptions about teaching (Henderson *et al.*, 2011; Gibbs and Coffey, 2004; Ho *et al.*, 2001). By measuring teaching practices and perceptions from the beginning, we can provide meaningful feedback over the course of this long-term project to help understand practices and perceptions about teaching.

Establishing a baseline for how FLC members teach their courses is critical to the long-term goals of the AACR project. In order for FLCs to function as they are intended to, we need to ascertain what types of ideas are being shared in the FLCs relating to teaching practices. A number of the faculty participating in the FLCs were clearly implementing varying degrees of student-focused teaching practices in their courses based on their COPUS, ATI, and TPI data. However, how effectively they are implementing said techniques into their courses is something that merits further investigation. The COPUS observations can also help discern effectiveness of teaching practices. For example, the COPUS data highlighted a distinct misalignment between the range of instructor activity (Figure 6) and student activity (Figure 7). Students were spending a noticeably larger percentage of time receiving information compared to the percentage of time instructors spent presenting material. The disparity of result indicates that there were instances where instructors are doing activities classified as "Guiding" yet, the students were spending that time receiving information. One common instance that illustrates this point is when instructors were following-up (FlUp) to an activity that students were completing, but doing so by lecturing through the follow-up. The extended follow-up by the instructor led the instructor profile to appear as if they were employing a student-focused strategy, but the students were experiencing a teacher-focused strategy. Instances of extended follow-up and others causing instructors to appear active while students remain passive are contributing factors for why fewer instructors are profiled as "Collaborative Learning" using the COPUS cluster analysis. Tracking for changes in this activity misalignment and in profile distribution can serve as an indicator of the effect of FLC participation on faculty. By establishing this thorough baseline using easily repeated measures, it is now possible to continue these efforts for the duration of the AACR project and monitor for changes in perceptions and practices in teaching by the FLC faculty.

Limitations of the Study

This work provides a comprehensive analysis of teaching practices utilized by instructors participating in FLCs as a part of the AACR project. As a result, I have to consider the possibility that the population of instructors being sampled may not be entirely representative of the entire population of STEM instructors in higher education. This, in part, could account for different conclusions being drawn relating to class size from my data as opposed to other studies conducted in an analogous manner (Smith *et al.,* 2014). It is also worth noting that this study was occasionally hindered by the segmented relationship between the FLCs at different institutions. Collecting data from all participating faculty requires them to completely fill out all required fields in surveys and be observed in the manner requested by the trained COPUS observers at each institution. Variability in the sample sizes of the different data sets is a direct result of difficulties encountered in garnering full participation from all FLCs.

Conclusions

I have documented the teaching practices used by the FLC faculty at the start of the project. Coupled with measuring their classroom activity, I was able to ascertain FLC faculty's level of awareness of their own teaching practices by surveying their teaching practices with the TPI and teaching philosophies with the ATI. Through various modes of analysis, I was able to determine that FLC faculty's perceptions of their own teaching did not always align with their measured classroom activity. I have established the baseline by which changes in teaching practices will be assessed over the course of the project. These data should have profound effects on the outlook of each FLC and the professional development resulting from their collaboration. In planning professional development, I have also validated the use of surveys such as the TPI to probe the variety of teaching practices being employed by instructors participating in professional development. In terms of helping to change their perceptions with respect to teaching practices, I have presented data that strongly suggest that the number of students in the course and instructor teaching experience do not restrict the diversity of teaching practices that can be employed in any course. Future professional development efforts should be able to utilize this data to tell a make a compelling case in transitioning teaching practices to more student-focused methods by changing instructor's perceptions about teaching in general.

References

Andrews TC, Lemons PP (2015). It's Personal: Biology Instructors Prioritize Personal Evidence over Empirical Evidence in Teaching Decisions. CBE Life Sci Educ *14*, 1-18.

Barlow, J., & Antoniou, M. (2007). Room for improvement: The experiences of new lecturers in higher education. Innovations in Education and Teaching International, 44(1), 67–77.

Barton MA, Richlin L (2004). Managing multiple faculty learning communities. New directions for teaching and learning. *97*, 81-85.

Clark, C. M., Froyd, J., Merton, P., & Richardson, J. (2004). The evolution of curricular change models within the foundation coalition. Journal of Engineering Education, 93(1), 37–47.

Cohen J (1960). A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement 20 (1), 37-46.

Connolly MR, Millar SB. (2006). Using workshops to improve instruction in STEM courses. Metropolitan Universities 17: 53–65.

Cox, M. D. (2001). Faculty learning communities: Change agents for transforming institutions into learning organizations. *To Improve the Academy, 19*, 69-93.

Cox, M. D. (2004). Introduction to faculty learning communities. In M. D. Cox & L. Richlin (Eds.), Building faculty learning communities: New directions for teaching and learning, no. 97 (Vol. 2004, pp. 5–23). San Francisco: Jossey-Bass.

Davidovitch, N., & Soen, D. (2006). Using students' assessments to improve instructors' quality of teaching. Journal of Further and Higher Education, 30(4), 351–376.

Ebert-May D, Derting TL, Hodder J, Momsen JL, Long TM, Jardeleza SE (2011). What we say is not what we do: effective evaluation of faculty professional development programs. BioScience *61*, 550–558.

Emerson, J. D., & Mosteller, F. (2000). Development programs for college faculty: Preparing for the twenty-first century. In R. M. Branch & M. A. Fitzgerald (Eds.), Educational media and technology yearbook 2000 (Vol. 25, pp. 26–42). Englewood, CO: Libraries Unlimited, Inc.

Freeman S, Eddy S, McDonough M, Smith MK, Okoroafor N, Jordt H, Wenderoth MP (2014). Active learning increases student performance in science, engineering, and mathematics. Proc Natl Acad Sci USA *111*, 8410–8415.

Fung L, Chow LPY (2002). Congruence of student teachers' pedagogical images and actual classroom practices. Educ Res *44*, 313–321.

Gibbs, G., & Coffey, M. (2004). The impact of training of university teachers on their teaching skills, their approach to teaching and the approach to learning of their students. Active Learning in Higher Education the Journal of the Institute for Learning and Teaching, 5(1), 87–100.

Henderson C, Beach A, Finkelstein N (2011). Facilitating change in undergraduate STEM instructional practices: an analytic review of the literature. J Res Sci Teach *48*, 952–984.

Ho, A., Watkins, D., & Kelly, M. (2001). The conceptual change approach to improving teaching and learning: An evaluation of a Hong Kong staff development programme. Higher Education, 42, 143– 169.

Hubball H, Clarke A, Beach AL. Assessing faculty learning communities. New Dir Teach Learn 2004; 97: 87–100

Landis JR, Koch GG (1977). The Measurement of Observer Agreement for Categorical Data. Biometrics, *33*(1), 159-174.

Lund TJ, Pilarz M, Velasco JB, Chakraverty D, Rosploch K, Undersander M, Stains M. (2015) The Best of Both Worlds: Building on the COPUS and RTOP Observation Protocols to Easily and Reliably Measure Various Levels of Reformed Instructional Practice. CBE Life Sci Educ *14*, 1-12.

Lynd-Balta, E., Erklenz-Watts, M., Freeman, C., & Westbay, T. D. (2006). Professional development using an interdisciplinary learning circle: Linking pedagogical theory to practice. Journal of College Science Teaching, 35(4), 18–24.

Murray K, Macdonald R (1997). The disjunction between lecturers' conceptions of teaching and their claimed educational practice. High Educ *33*, 331–349.

Sandell KL, Wigley K, Kovalchick A (2004). Developing facilitators for faculty learning communities. New directions for teaching and learning, *97,* 51-62.

Seymour, E. (2002). Tracking the processes of change in US undergraduate education in science, mathematics, engineering, and technology. *Science Education, 86*(1), 79-105.

Silverthorn, D. U., Thorn, P. M., & Svinicki, M. D. (2006). It's difficult to change the way we teach: Lessons from the Integrative Themes in Physiology curriculum module project. [Article]. Advances in Physiology Education, 30(4), 204–214.

Smith MK, Jones FHM, Gilbert SL, Wieman CE (2013). The Classroom Observation Protocol for Undergraduate STEM (COPUS): a new instrument to characterize university STEM classroom practices. CBE Life Sci Educ *12*, 618–627.

Smith MK, Vinson EL, Smith JA, Lewin JD, Stetzer MR (2014). A Campus-Wide Study of STEM Courses: New Perspectives on Teaching Practices and Perceptions. CBE Life Sci Educ *13*, 624-635.

Trigwell, K., & Prosser, M. (2004). Development and use of the approaches to teaching inventory. *Educational Psychology Review, 16*(4), 409-424. doi: 10.1007/s10648-004-0007-9

White House (2012). Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics. In Report from PCAST, President's Council of Advisors on Science and Technology.

Wieman CE, Gilbert S (2014). The Teaching Practices Inventory; a new tool for characterizing college and university teaching in mathematics and science. CBE Life Sci Educ *13*, 552–569.

Tables

Table 1 – COPUS Codes. Description of each COPUS code, its abbreviation, and the definition used when coding.

Instructor COPUS Code	Abbreviation	Definition
Lecturing	Lec	Lecturing (presenting content, deriving mathematical results, presenting a problem solution, etc.)
Writing	RtW	Real-time writing on board, doc. projector, etc. (often checked off along with Lec)
Follow-up	Fup	Follow-up/feedback on clicker question or activity to entire class
Pose Q	PQ	Posing non-clicker question to students (non-rhetorical)
Clicker Q	CQ	Asking a clicker question (mark the entire time the instructor is using a clicker question, not just when first asked)
Answer Q	AnQ	Listening to and answering student questions with entire class listening
Moving/Guiding	MG	Moving through class guiding ongoing student work during active learning task
One-on-One	1o1	One-on-one extended discussion with one or a few individuals, not paying attention to the rest of the class (can be along with MG or AnQ)
Demonstration	D/V	Showing or conducting a demo, experiment, simulation, video, or animation
Administration	Adm	Administration (assign homework, return tests, etc.)
Waiting	W	Waiting when there is an opportunity for an instructor to be interacting with or observing/listening to student or group activities and the instructor is not doing so
Other	O	Other – explain in comments

Student COPUS Code	Abbreviation	Definition
Listening	L	Listening to instructor/taking notes, etc.
Individual thinking	Ind	Individual thinking/problem solving. Only mark when an instructor explicitly asks students to think about a clicker question or another question/problem on their own.
Clicker Q discussion	CG	Discuss clicker question in groups of 2 or more students
Worksheet group work	WG	Working in groups on worksheet activity
Other group work	OG	Other assigned group activity, such as responding to instructor question
Answer Q	Anq	Student answering a question posed by the instructor with rest of class listening
Student Q	SQ	Student asks question
Whole class discussion	WC	Engaged in whole class discussion by offering explanations, opinion, judgment, etc. to whole class, often facilitated by instructor
Predicting	Prd	Making a prediction about the outcome of demo or experiment
Student presenting	SP	Presentation by student(s)
Test/quiz	TQ	Test or quiz
Waiting	W	Waiting (instructor late, working on fixing AV problems, instructor otherwise occupied, etc.)
Other	O	Other – explain in comments

Table 2 - 2014 Observation Count. Number of courses that were observed a specific number of times during the given semester.

Observation Count	Observed One Time	Observed Two Times	Observed Three Times	Observed Four Times	Observed Five Times	Observed Six Times
Spring 2014 Number of Courses	3	10	1	0	0	0
Fall 2014 Number of Courses	2	5	3	2	0	1

Table 3 - Collapsed Codes Description. The organization of the 25 COPUS codes into the eight collapse codes: four for instructor and four for students COPUS codes. Table shows the category for each collapsed code, the collapsed abbreviation, the COPUS code abbreviation, and the corresponding description.

	Collapsed Code	Code	Description
Instructor is Doing	Presenting (P)	Lec	Lecturing or presenting information
		RtW	Real-time writing
		D/V	Showing or conducting a demo, experiment, simulation, etc.
	Guiding (G)	FUp	Follow-up/feedback on clicker question or activity to class
		PQ	Posing non-clicker question to students (non-rhetorical)
		CQ	Asking clicker question (entire time, not just when first asked)
		AnQ	Listening to and answering student questions to entire class
		MG	Moving through class guiding ongoing student work
		1o1	One-on-one extended discussion with individual students
	Administration (A)	Adm	Administration (assign homework, return test, etc.)
	Other (OI)	W	Waiting (instructor late, working on fixing AV problems, etc.)
		O	Other
Students are Doing	Receiving (R)	L	Listening to instructor
	Students Talking to Class (STC)	AnQ	Student answering question posed by instructor
		SQ	Student asks question
		WC	Students engaged in whole class discussion
		SP	Students presenting to entire class
	Students Working (SW)	Ind	Individual thinking/problem solving (explicitly asked by instructor to do so)
		CG	Discuss clicker question in groups of 2 or more students
		WG	Working in groups on worksheet activity
		OG	Other assigned group activity
		Prd	Making a prediction about a demo or experiment
		TQ	Test or quiz
	Others (OS)	W	Waiting (instructor late, working on fixing AV problems, etc.)
		O	Other

Table 4 – Approaches to Teaching Inventory (ATI) Category Items. Wording of each item of the ATI, separated into two categories: Information-Transfer/Teacher-Focused (ITTF) and Conceptual-Change/Student-Focused (CCSF). Each item is rated on a 1-5 Likert scale by the instructor. The sum of the eight items for each category is the total score for that category.

Information-Transfer/Teacher-Focused Category Items
I design my teaching in this subject with the assumption that most of the students have very little useful knowledge of the topics to be covered.
I feel it is important that this subject should be completely described in terms of specific objectives relating to what students have to know for formal assessment items.
I feel it is important to present a lot of facts to students so that they know what they have to learn for this subject.
In this subject I concentrate on covering the information that might be available from a good textbook.
I structure this subject to help students pass the formal assessment items.
I think an important reason for running teaching sessions in this subject is to give students a good set of notes.
In this subject, I only provide the students with the information they will need to pass the formal assessments.
I feel that I should know the answers to any questions that students may put to me during this subject.

Conceptual-Change/Student-Focused Items
In my interactions with students in this subject I try to develop a conversation with them about the topics we are studying.
I feel that the assessment in this subject should be an opportunity for students to reveal their changed conceptual understanding of the subject.
I set aside some teaching time so that the students can discuss, among themselves, the difficulties they encounter studying this subject.
I encourage students to restructure their existing knowledge in terms of the new way of thinking about the subject they will develop.
In teaching sessions for this subject, I use difficult or undefined examples to provoke debate.
I make available opportunities for students in this subject to discuss their changing understanding of the subject.
I feel that it is better for students in this subject to generate their own notes rather than always copy mine.
I feel a lot of teaching time in this subject should be used to question students' ideas.

Table 5 – Conceptual-Change/Student-Focused Regression Data Summary. R^2 values and significance of correlations between ATI CCSF section scores and five COPUS collapsed codes of interest. No significant correlations were observed ($p>0.05$).

Collapsed Code	Linear Regression	R^2	P-value
Presenting (P)	-	0.051	0.29
Guiding (G)	+	0.047	0.31
Receiving (R)	-	0.10	0.14
Students Talking to Class (STC)	+	0.082	0.18
Students Working (SW)	+	0.016	0.55

Table 6 – Teaching Practices Inventory (TPI) Total Score Regression Data Summary. R^2 values and significance of correlations between TPI total score and five COPUS codes of interest. No significant correlations were observed ($p>0.05$).

Collapsed Code	Linear Regression	R^2	P-value
Presenting (P)	-	0.22	0.15
Guiding (G)	+	0.23	0.13
Receiving (R)	-	0.31	0.073
Students Talking to Class (STC)	+	0.067	0.44
Students Working (SW)	+	0.30	0.083

Table 7 - Class Size Regression Data Summary. R^2 values and significance of correlations between class size and five COPUS codes of interest. No significant correlations were observed ($p>0.05$).

Collapsed Code	Linear Regression	R^2	P-value
Presenting (P)	-	0.00075	0.90
Guiding (G)	+	0.00073	0.90
Receiving (R)	-	0.0034	0.79
Students Talking to Class (STC)	-	0.0087	0.67
Students Working (SW)	+	0.029	0.44

Table 8 - Teaching Experience Regression Data Summary. R^2 values and significance of correlations between teaching experience and five COPUS codes of interest. No significant correlations were observed ($p>0.05$).

Collapsed Code	Linear Regression	R^2	P-value
Presenting (P)	+	0.013	0.58
Guiding (G)	-	0.0016	0.85
Receiving (R)	-	0.032	0.38
Students Talking to Class (STC)	+	0.095	0.13
Students Working (SW)	-	0.0019	0.83

Figures

The following DNA sequence occurs near the middle of the coding region of a gene.

DNA 5' A A T G A A T G G* G A G C C T G A A G G A 3'

There is a G to A base change at the position marked with an asterisk. Consequently, a codon normally encoding an amino acid becomes a stop codon.

1. How will this alteration influence DNA replication?
2. How will this alteration influence transcription?
3. How will this alteration influence translation?

Figure 1 - Sample AACR Question. An example AACR question used by AACR faculty learning community (FLC) members in their large enrollment STEM classrooms.

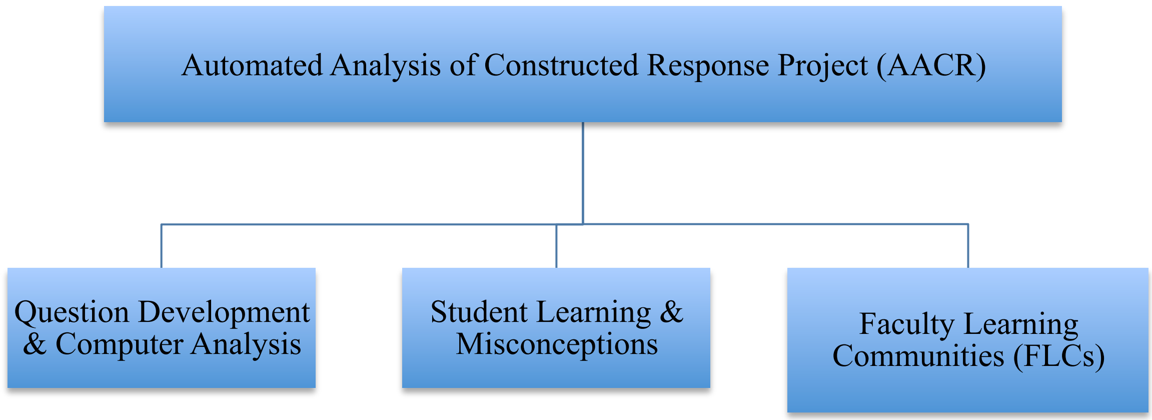


Figure 2 - Strands of the AACR Project. The AACR project can be broken down into three research strands. My thesis research focused on the work being done with Faculty Learning Communities.

	1. Students doing													2. Instructor doing													Comments
min	L	Ind	CG	WG	OG	AnQ	SQ	WC	Prd	SP	T/Q	W	O	Lec	RtW	Fup	PQ	CQ	AnQ*	MG	1o1	D/V	Adm	W*	O*	
0 - 2																										
2																										
4																										
6																										
8-10																										
	L	Ind	CG	WG	OG	AnQ	SQ	WC	Prd	SP	T/Q	W	O	Lec	RtW	Fup	PQ	CQ	AnQ	MG	1o1	D/V	Adm	W	O	
10-12																										
12																										
14																										
16																										
18 - 20																										

Figure 3 - Sample COPUS Protocol Sheet. An excerpt of the coding sheet used by COPUS observers in the classroom. The standard protocol includes more time intervals to allow for complete coding of class periods of variable lengths.

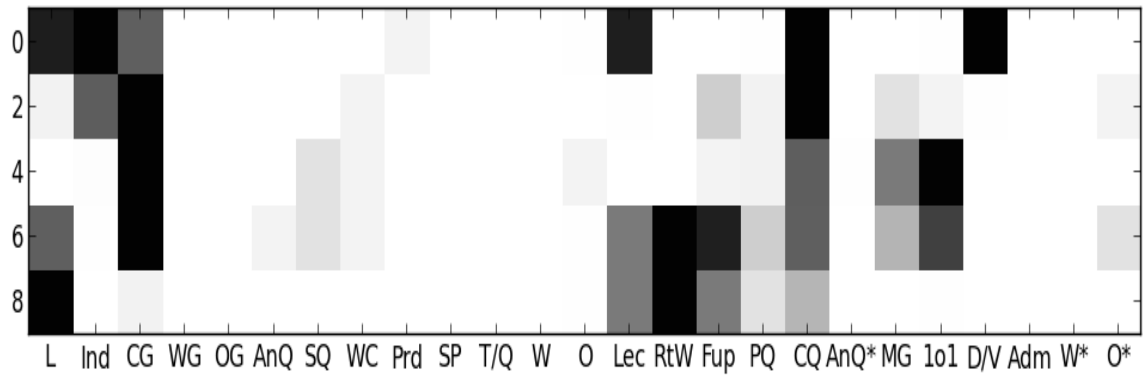


Figure 4 - Sample COPUS Heat Map. The heat map was generated by taking the sum of 14 observations of the same video during COPUS training. The darker a square, the greater the number of observers that used the same COPUS code during that time interval and the greater the agreement on that code. White squares indicate a code that was not coded by any observer during that time interval. The lighter squares indicated discrepancies in codes between observers and were used in training to clarify codes and observed behaviors.

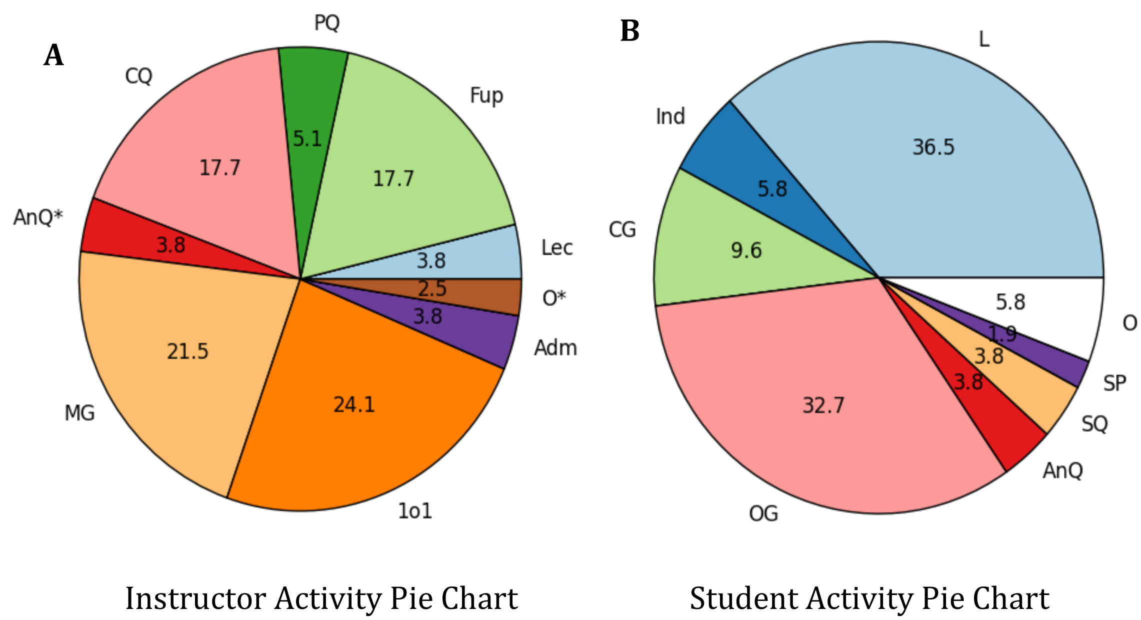


Figure 5 - Sample Pie Chart from COPUS Observations. Represents the percentage of COPUS codes from one classroom observation for instructor (a) and students (b). Calculated by taking the number of times a code was coded during the class period divided by the total number of codes coded. For the abbreviations of the COPUS codes, see Table 3.

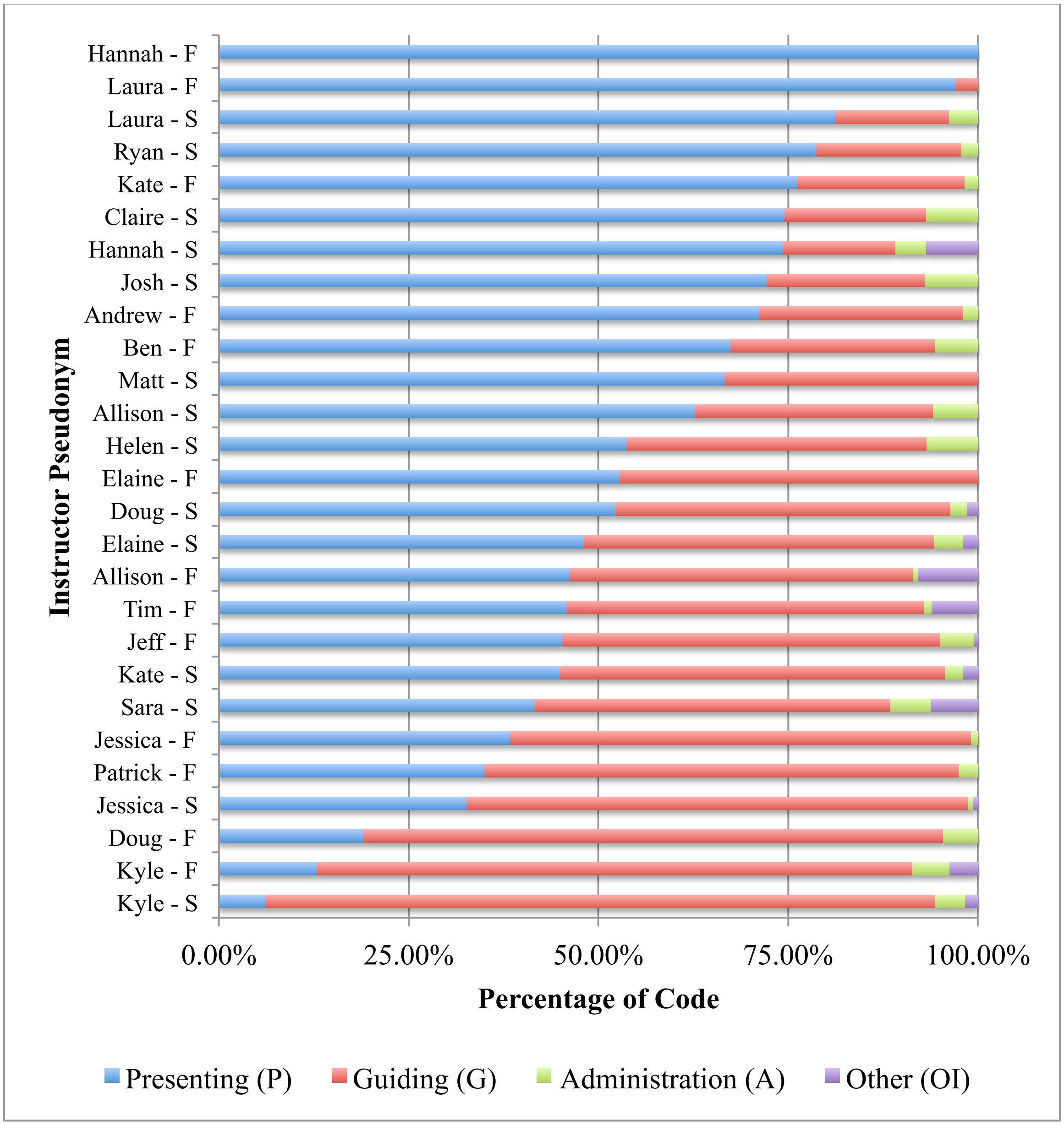


Figure 6 - COPUS Instructor Activity. Percentage of instructor collapsed COPUS codes from all observations in 2014. Each row represents a separate course (n=27) that is denoted by the instructor pseudonym and the semester (F = Fall, S = Spring). Multiple observations of the same course are averaged to produce the single bar (see *Methods*). For relationship between COPUS codes and collapsed codes, see Table 3.

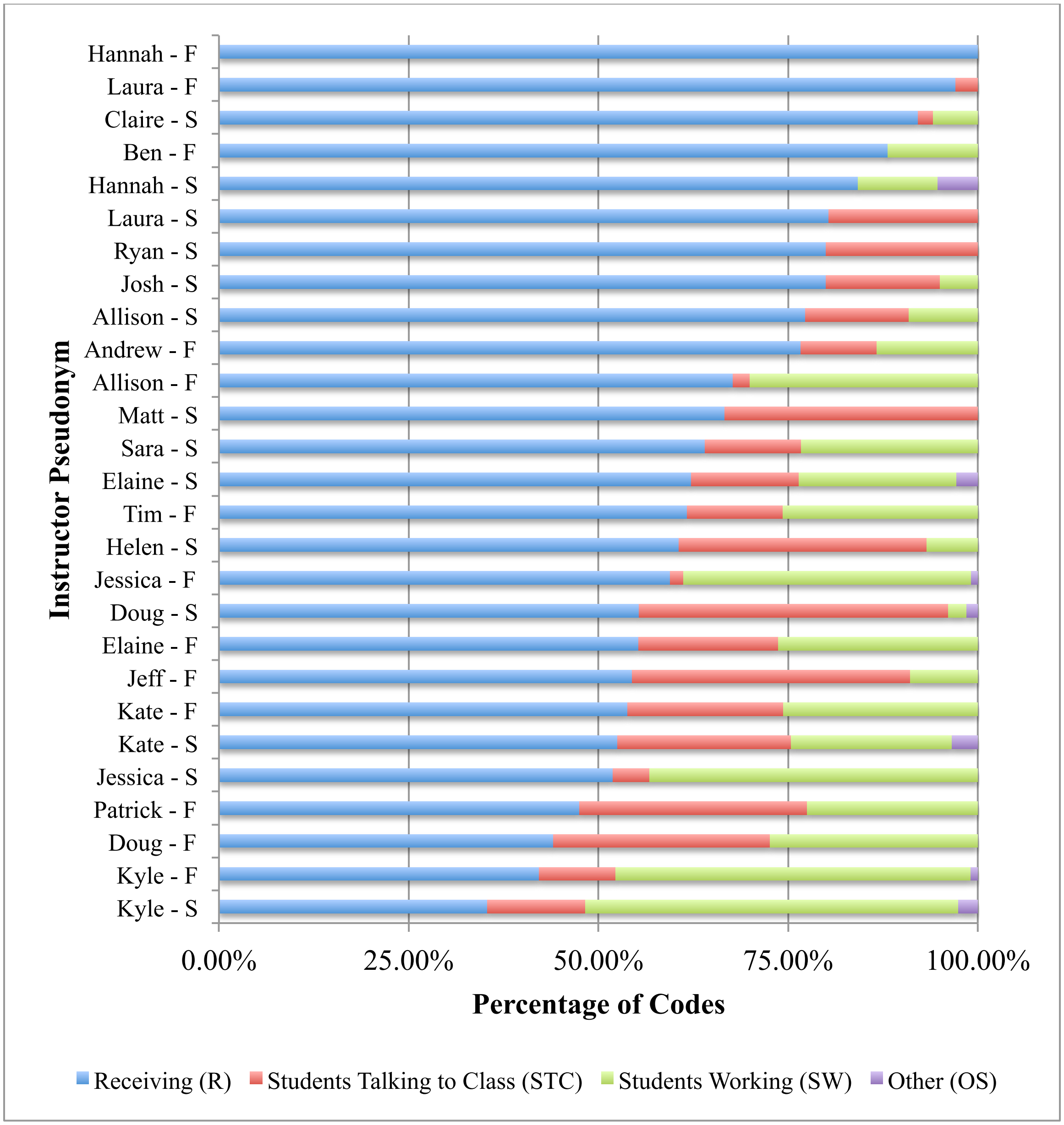


Figure 7 - COPUS Student Activity. Percentage of student collapsed COPUS codes from all observations in 2014. Each row represents a separate course (n=27) that is denoted by the instructor pseudonym and the semester (F = Fall, S = Spring). Multiple observations of the same course are averaged to produce the single bar (see *Methods*). For relationship between COPUS codes and collapsed codes, see Table 3.

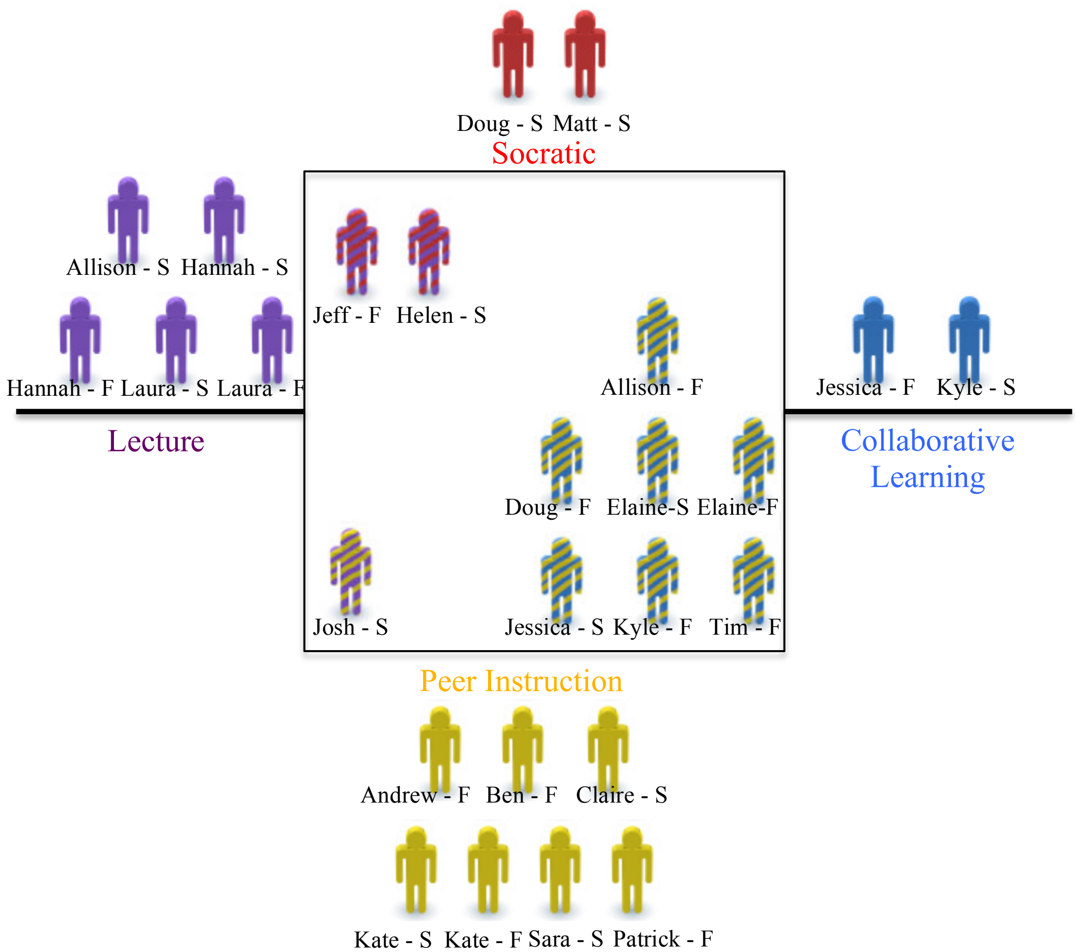


Figure 8 – COPUS Profile Distribution. Distribution of COPUS profiles, from COPUS cluster analysis, along the instructional style continuum. The solid colored stick figures located outside of the box are those who are profiled into one of the four labeled instructional styles. The striped stick figures that are in the box fall between two instructional styles and are referred to as “hybrid” profiles. For example, the purple/yellow striped stick figure (bottom left corner) refers the to “Lecture/Peer Instruction Hybrid” profile, indicating that this instructor had observations for both instructional styles.

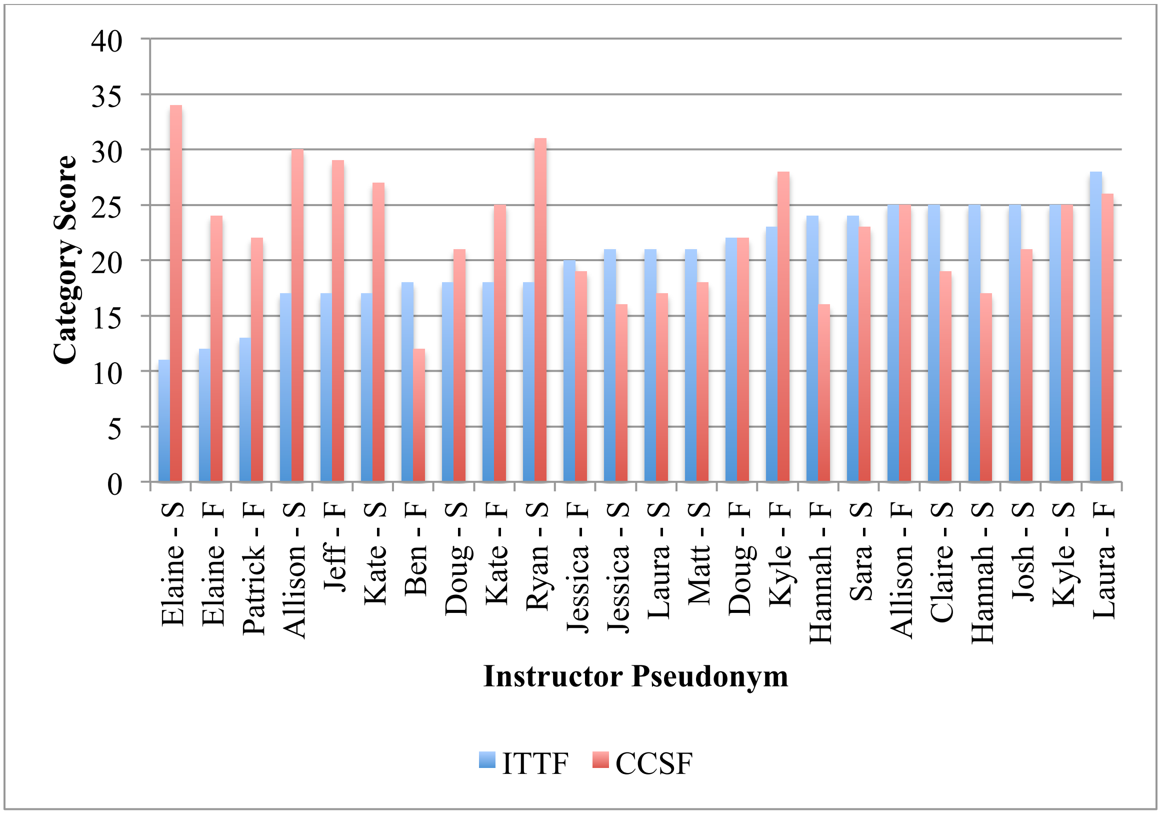
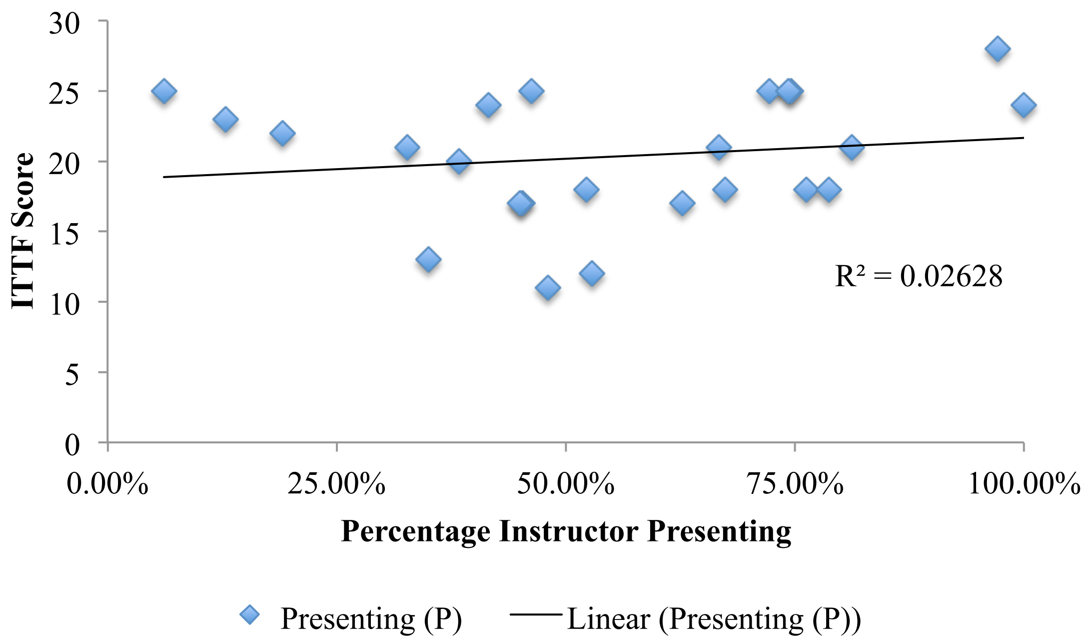


Figure 9 - Approaches to Teaching Inventory (ATI) Scores. ATI category scores [Information-Transfer/Teacher-Focused (ITTF) and Conceptual-Change/Student-Focused (CCSF)] for each instructor (n=24) by semester (F = Fall 2014, S = Spring 2014). Data are arranged by course ITTF score (increasing from left to right). The different population sizes for the ATI data compared to the COPUS data (n=27, Figures 4&5) results from not all instructors completing the survey as requested.

A



B

Collapsed Code	Linear Regression	R^2	P-value
Presenting (P)	+	0.026	0.45
Guiding (G)	-	0.043	0.33
Receiving (R)	+	0.13	0.090
Students Talking to Class (STC)	-	0.22	0.020
Students Working (SW)	-	0.0014	0.87

Figure 10 – Relationships between the Approaches to Teaching Inventory Information-Transfer/Teacher-Focused (ITTF) Category Scores vs. Percent of Codes for Instructor Presenting as Determined by COPUS Observations. A) A non-significant correlation was observed between percent of instructor presenting and instructor ITTF category score ($p > 0.05$) by course ($n = 24$). The maximum score in this category is a 40. B) Table summarizing all correlations between ITTF score and COPUS collapsed codes for instructor (light blue) and student activities (dark blue).

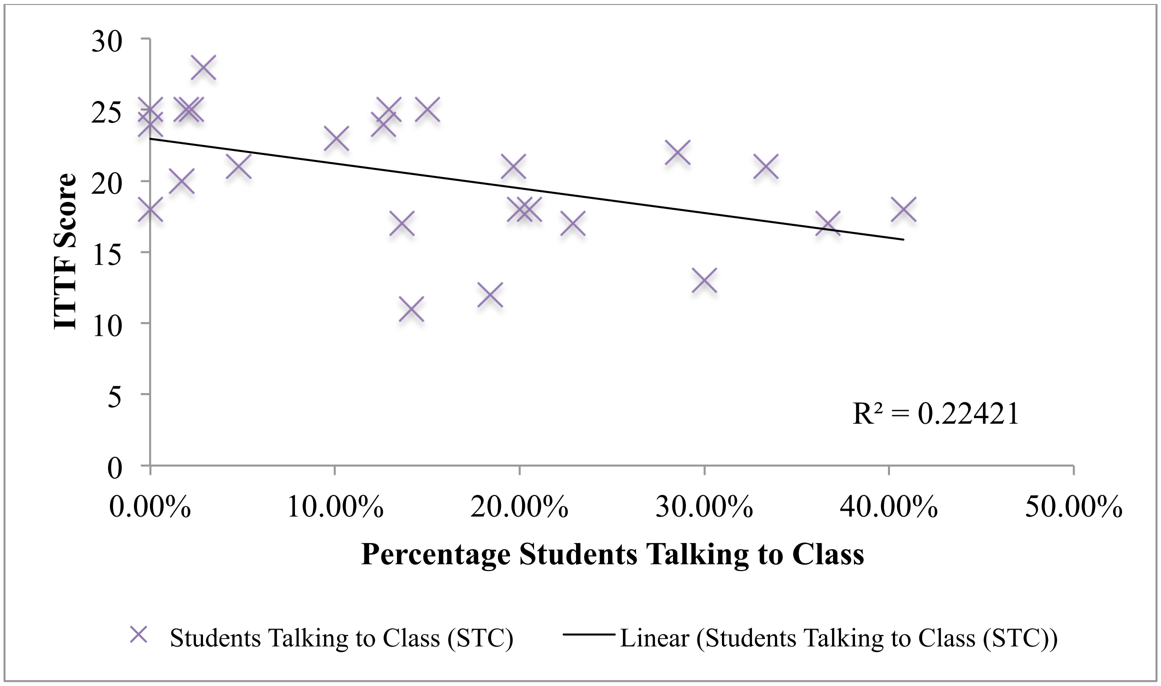


Figure 11 - Relationships between the Approaches to Teaching Inventory Information-Transfer/Teacher-Focused (ITTF) Category Scores vs. Percent of Codes for Students Talking to Class as Determined by COPUS Observation. A significant correlation was observed between percentage of student talking to class and instructor ITTF (ATI, Information-Transfer/Teacher-focused) category score ($p < 0.05$) by course ($n = 24$). The maximum score in this category is a 40.

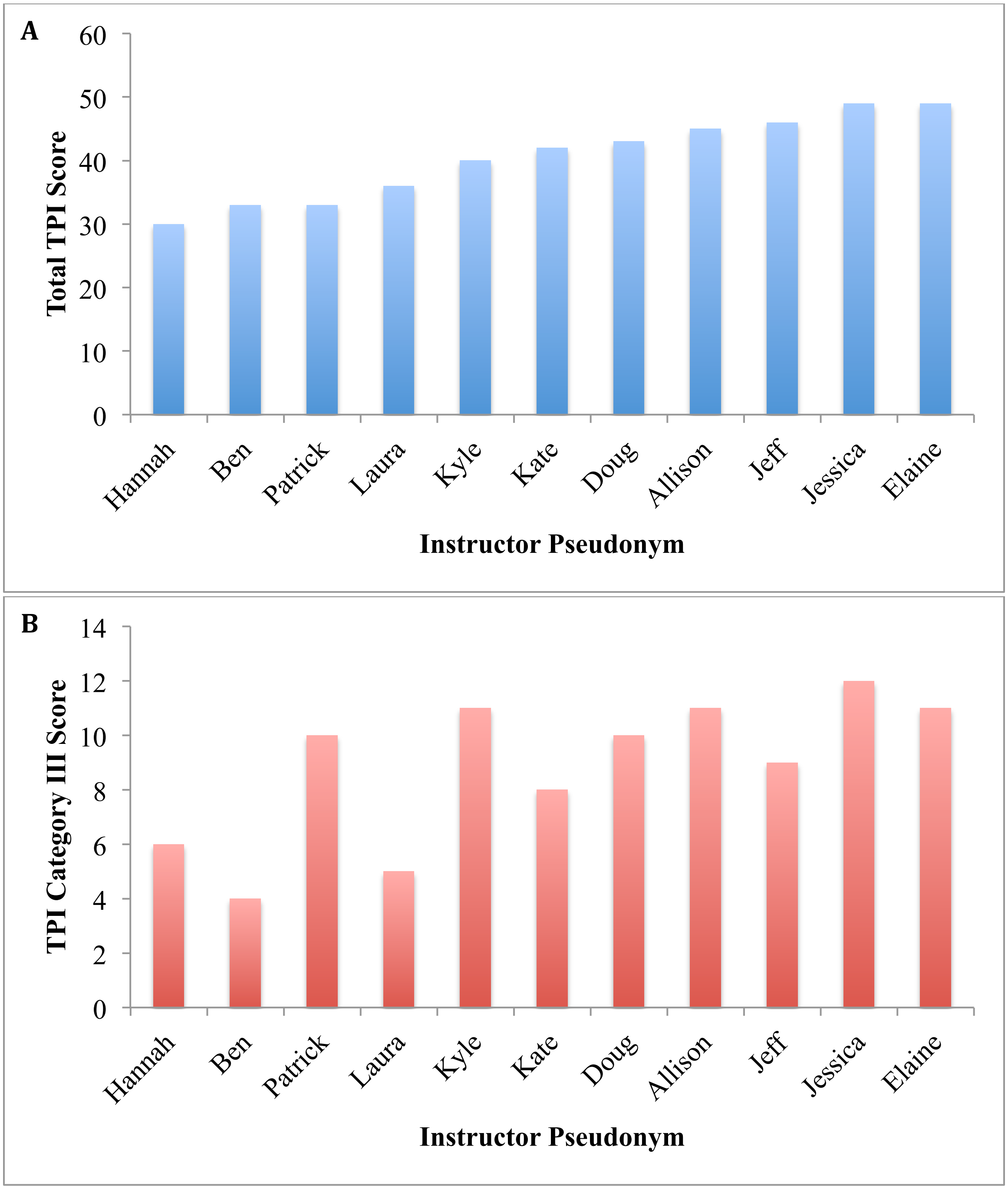


Figure 12 – Teaching Practices Inventory (TPI) Scores for All Categories and "In-Class Features and Activities" Category Alone. (A) TPI total scores arranged from least to greatest score by instructor pseudonym. The maximum score on the TPI is a 72. (B) TPI Category III scores ("In-Class Features and Activities") are arranged in the same order as in (A). The maximum score on this category of the TPI is 15.

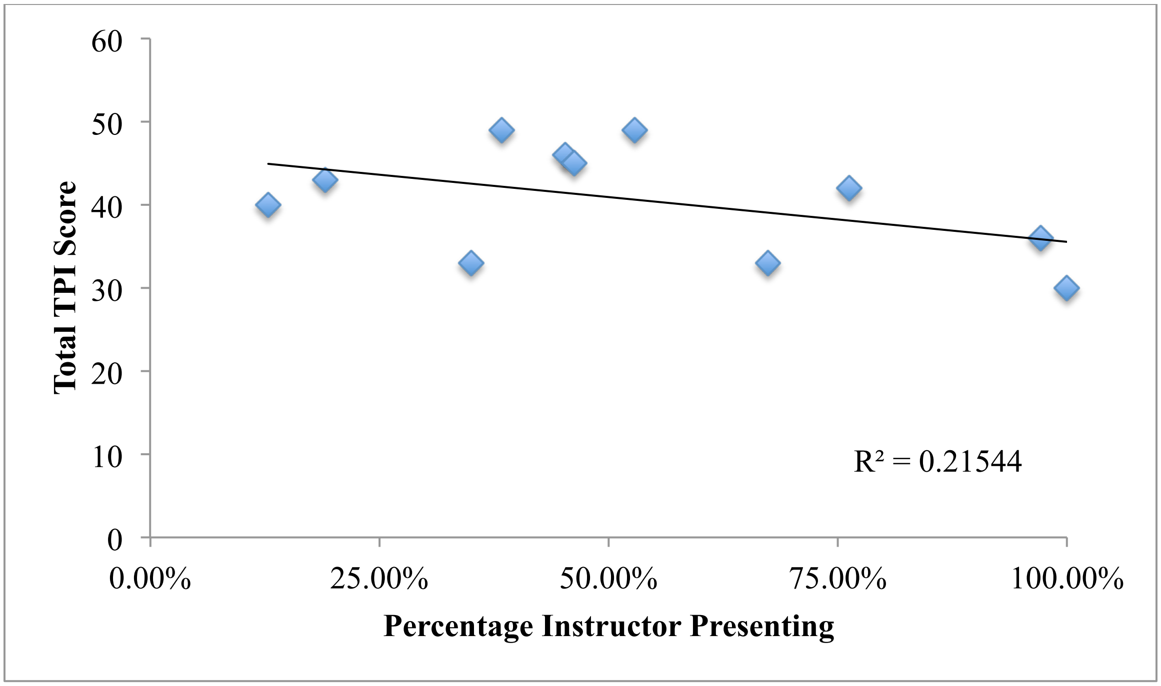


Figure 13 – Relationship between the Total Teaching Practices Inventory Score vs. Percent of Codes for Instructor Presenting as Determined by COPUS Observation. No significant correlation was observed between the instructor total TPI score (n = 11) and COPUS percent instructor presenting ($p > 0.05$) by course. The maximum score on the TPI is a 72.

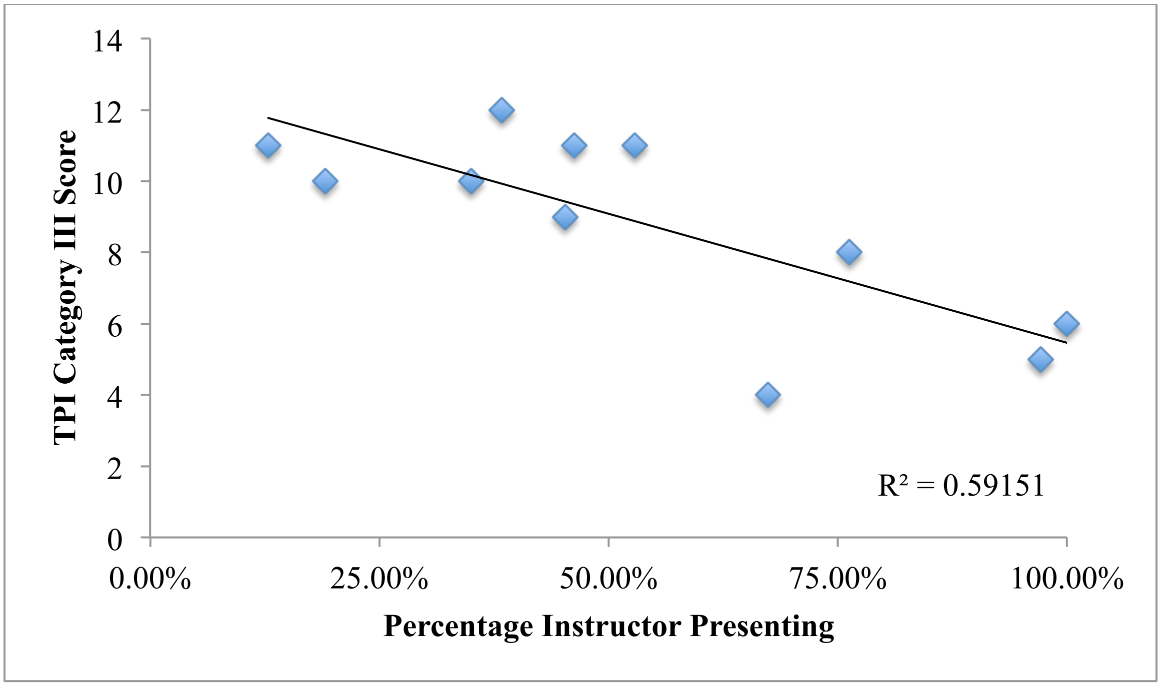


Figure 14 – Relationship between the Teaching Practices Inventory In-Class Features and Activities (Cat III) Score vs. Percent of Codes for Instructor Presenting as Determined by COPUS Observation. A significant correlation was observed between COPUS instructor percent presenting and instructor TPI Category III ("In-Class Features and Activities") score ($p < 0.05$) by course. The maximum score on this category of the TPI is 15.

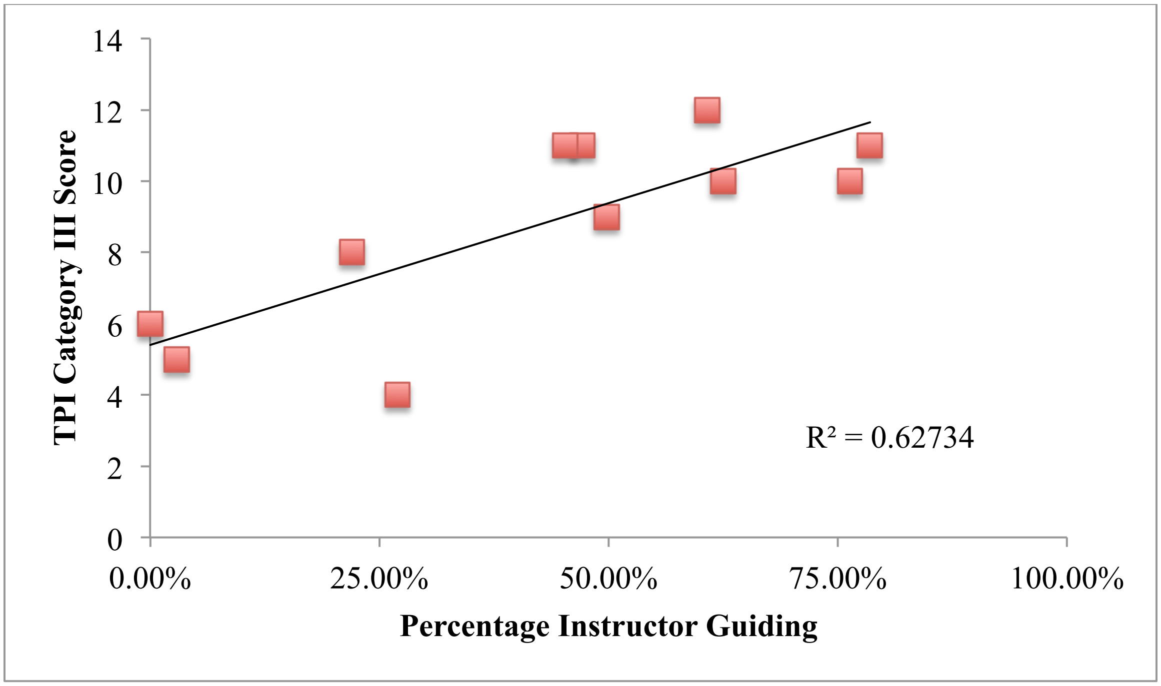


Figure 15 – Relationship between the Teaching Practices Inventory In-Class Features and Activities (Cat III) Score vs. Percent of Codes for Instructor Guiding as Determined by COPUS Observation. A significant correlation was observed between COPUS instructor percent guiding and instructor TPI Category III ("In-Class Features and Activities") score ($p < 0.05$) by course. The maximum score on this category of the TPI is 15.

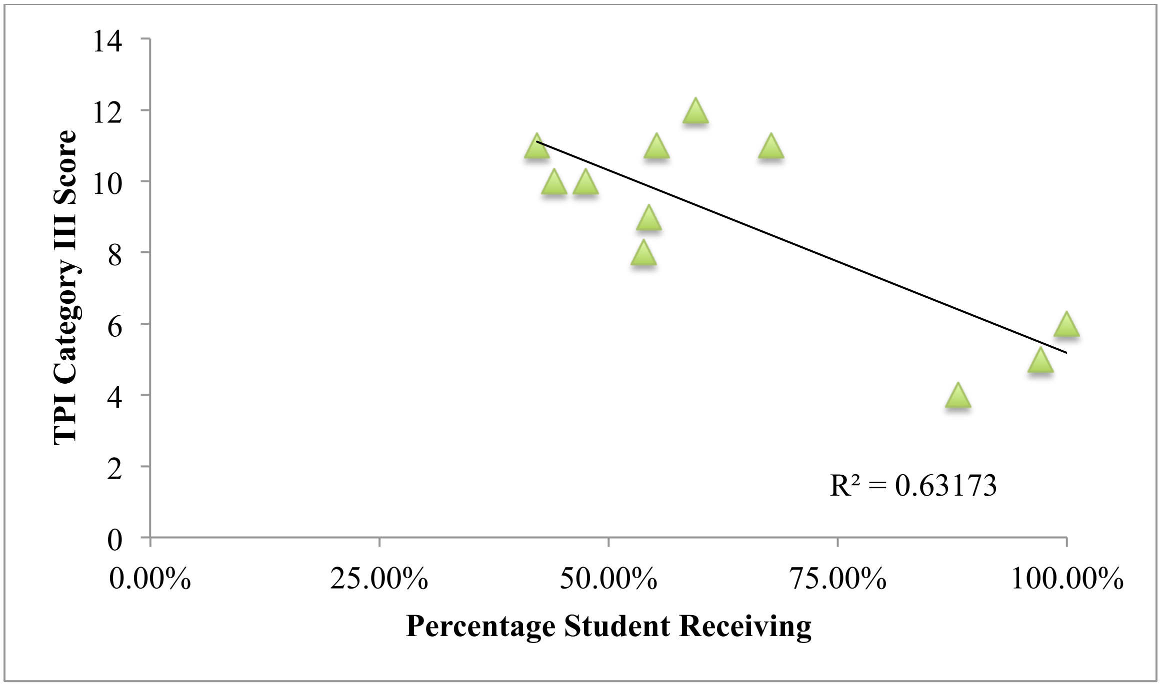


Figure 16 – Relationship between the Teaching Practices Inventory In-Class Features and Activities (Cat III) Score vs. Percent of Codes for Students Receiving as Determined by COPUS Observation. A significant correlation was observed between COPUS student percent receiving and instructor TPI Category III ("In-Class Features and Activities") score ($p < 0.05$) by course. The maximum score on this category of the TPI is 15.

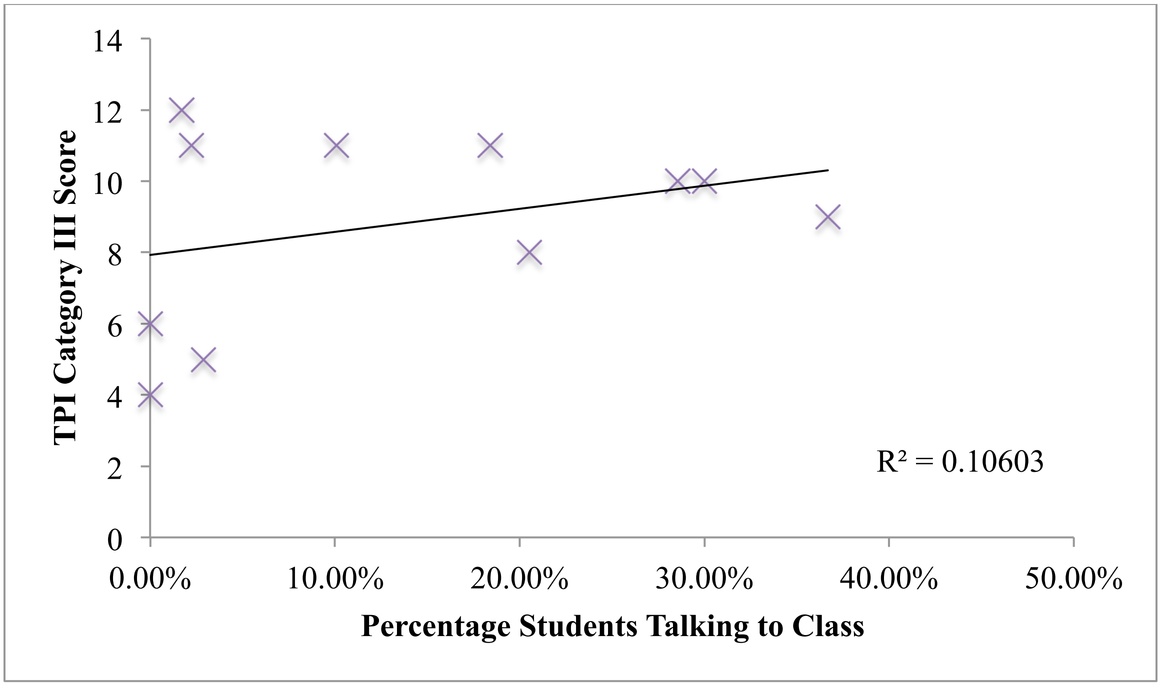


Figure 17 – Relationship between the Teaching Practices Inventory In-Class Features and Activities (Cat III) Score vs. Percent od Codes for Students Talking to Class as Determined by COPUS Observation. A non-significant correlation was observed between COPUS student percent receiving and instructor TPI Category III ("In-Class Features and Activities") score ($p > 0.05$) by course. The maximum score on this category of the TPI is 15.

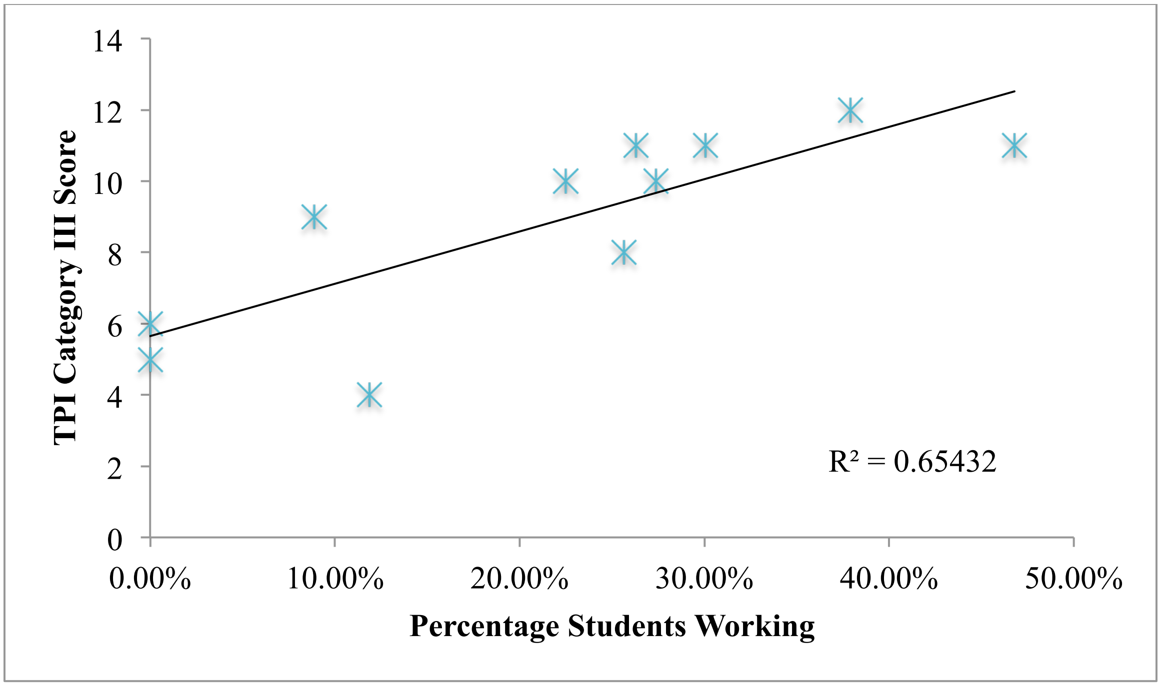


Figure 18 – Relationship between the Teaching Practices Inventory In-Class Features and Activities (Cat III) Score vs. Percent of Codes for Students Working as Determined by COPUS Observation. A significant correlation was observed between COPUS student percent working and instructor TPI Category III ("In-Class Features and Activities") score ($p < 0.05$) by course. The maximum score on this category of the TPI is 15.

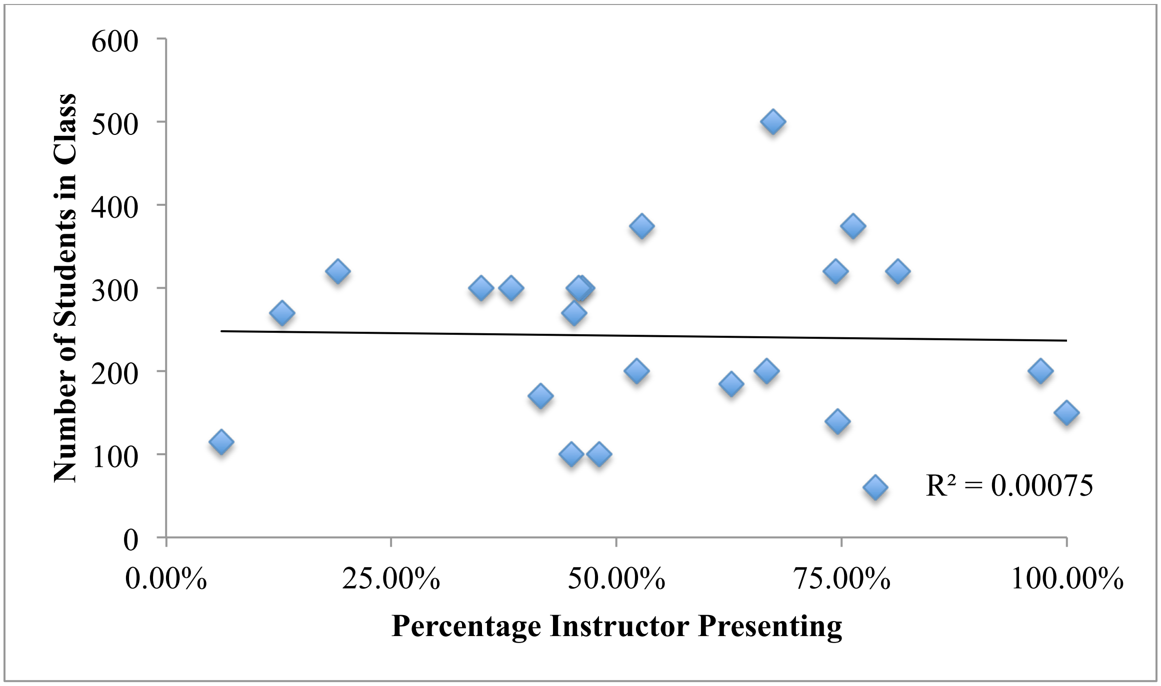


Figure 19 – Relationship between the Class Size vs. Percent of Codes for Instructor Presenting as Determined by COPUS Observations. There was no significant correlation between the number of students in the class and COPUS instructor percent presenting by course ($p > 0.05$).

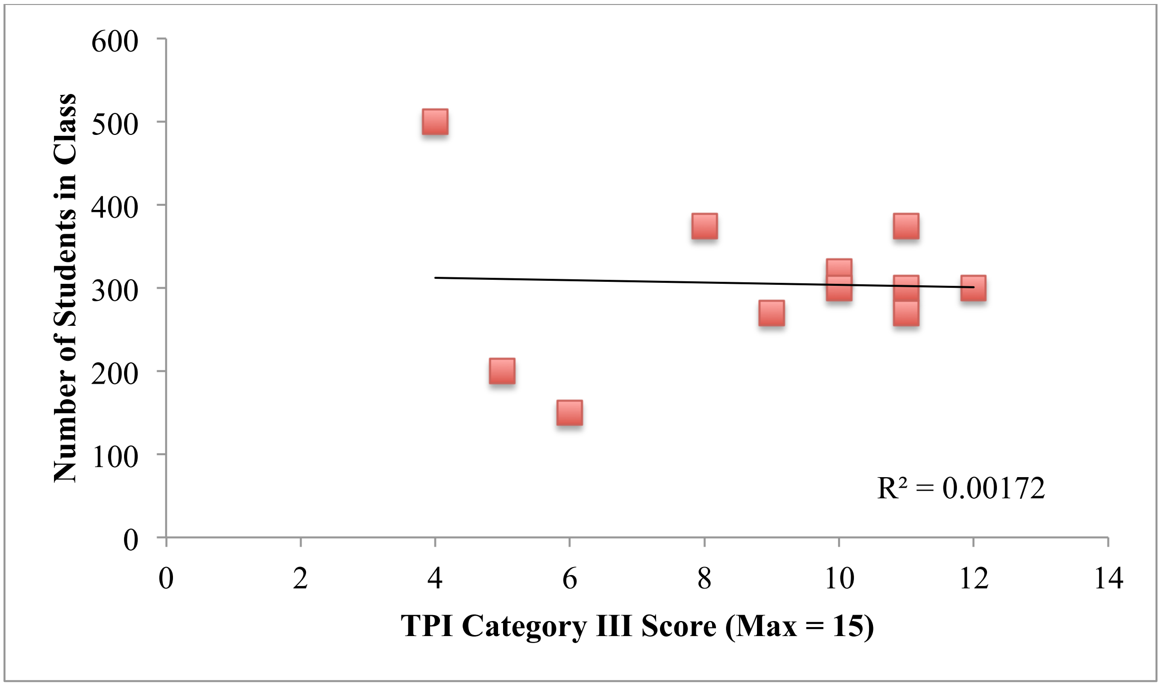


Figure 20 – Relationship between the Class Size vs. Teaching Practices Inventory In-Class Features and Activities (Cat III) Score. There was no significant correlation between the number of students in the class and TPI Category III ("In-Class Features and Activities") score ($p > 0.05$) by course. The maximum score on this category of the TPI is 15.

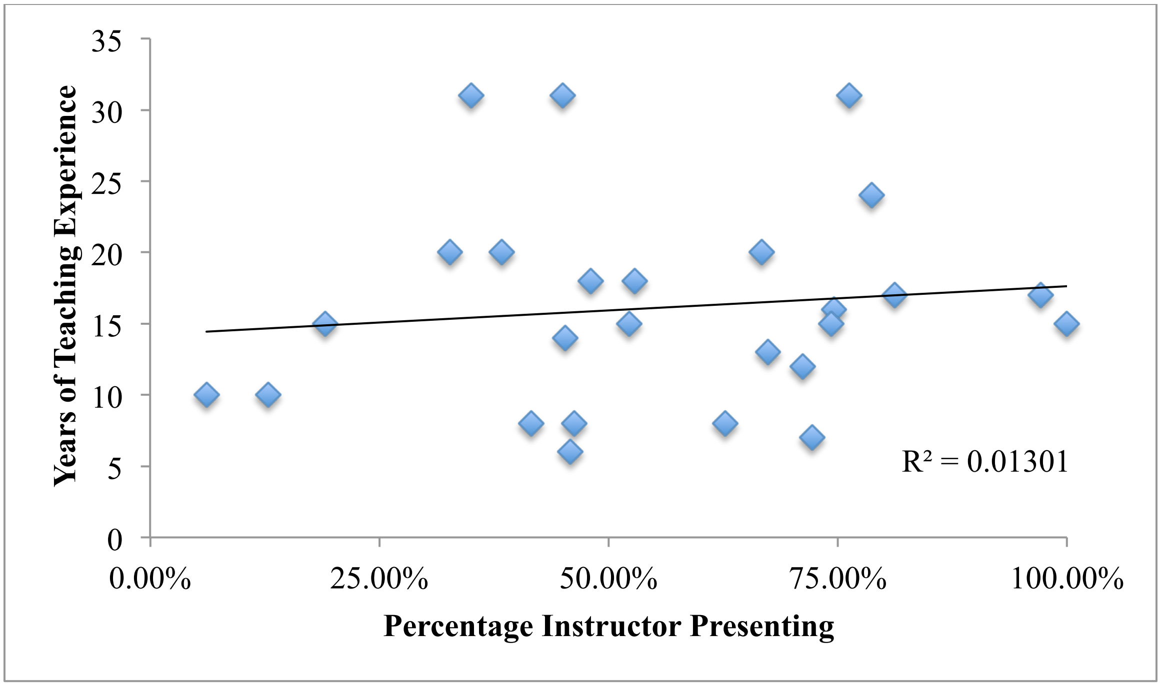


Figure 21 – Relationship between the Instructor Teaching Experience vs. Percent of Codes for Instructor Presenting as Determined by COPUS Observations. No significant correlation was observed between the number of years of teaching experience at the start of the FLC and COPUS percent instructor presenting by course ($p > 0.05$).

Author's Biography

Scott Merrill was born in Scarborough, Maine on August 8th, 1993. He was raised in Scarborough and graduated from Scarborough High School in 2011. Majoring in biology, Scott has minors in chemistry and premedical studies. He was a member of the University of Maine Men's Track and Field team for four years, serving as team captain during his senior year. In his time at UMaine, he has served as a chemistry lab teaching assistant, Maine Learning Assistant for biology and chemistry, a Peer Led Team Learning recitation leader for chemistry, and a student-athlete tutor in numerous subjects. During the summers following his sophomore and junior years, Scott served as a Student Undergraduate Research Assistant (SURA) for the Maine PSP under Professor Michelle Smith. His summer as a SURA culminated in his first publication (Smith MK, Merrill S. Why do some people inherit a predisposition to cancer? A small group activity on cancer genetics. CourseSource. 2015, http://coursesource.org/courses/why-do-some-people-inherit-a-predisposition-to-cancer-a-small-group-activity-on-cancer). His second summer as a SURA was focused on completing work for this thesis. In his time at the University of Maine, Scott was named to the Deans List numerous times, an annual member of the America East Commissioners Honor Roll, and the representative for the Men's Track and Field Team for Team Maine 2014 as a result of have the highest GPA on the team.

Upon graduation, Scott will enroll at Tufts University School of Medicine in Fall 2015. He is part of the Maine Track Program and will be pursuing a Doctor of Medicine (MD) degree.