Classroom Research Working Paper Series

Volume 5

Projects conducted in 2015-2017 by graduate and postdoctoral scholars and faculty/staff collaborators from four institutions

Cornell University
David Bassen
Mary Centrella
Aaron Chiou
Ashley Earle
Katryn Evinson
Frances Fairbairn
Annetta Fotopoulos
Kathleen Hefferon
Jing-Mao Ho
Derek Holyoak

Reece Kearney
Yehudah Pardo
Kristen Pueschel
Shivem Shah
Catherine Smith
Katherine Thorsteinson
Allison Truhlar
Laura Tupper
Lucy Xiaolu Wang
M. Todd Walter
Kimberly Williams

Columbia University
Elizabeth Case

University of Rochester
Emily Hangen

Yale University
Larry Bowman
Kelly Culhane
Kaury Kucera
Samantha Lin
Ahrume Julie Park

Center for Teaching Innovation

Editors: Kimberly Williams, Eugene Law, and Runjini Raman

Table of Contents

Enhancing Student Engagement and Understanding of Failure in a Biomechanics Classroom

David Bassen

Meinig School of Biomedical Engineering, Cornell University
dmb457@cornell.edu

Abstract

This study presents a planning methodology for lecture heavy class periods that incorporates active learning techniques along with an example of purposeful, observational reflection. Through the analysis and understanding of two class periods, it is demonstrated that the success of student driven class time hinges on successful implementation of the lecture material. Further, minute-reflection papers are shown to be a low-cost way of getting student feedback to help both students and the instructor.

Introduction

There are few pedagogical studies specifically focusing on teaching biomechanics that involve active learning. Further, at least one study (Savinainen, Nieminen, Mäkynen, & Viiri, 2013) states that there is little help in the teaching literature for instructors of physics, a much broader field that is fundamental to biomechanics. Many of these papers contain very technical and involved advice on teaching pedagogy and discuss methods that the average instructor probably would not implement due to the time and effort required to change established course designs. One exception is Huston's paper (1999), a retrospective on the author's decades of teaching biomechanics. While it does not include the transformative techniques promoted by the other studies, it includes many practical insights that could increase educator awareness of key problems without the need to create entirely new course designs and teaching materials. Knudson (2016) also provides an example of a pedagogical study in biomechanics by gathering misconceptions for students in courses on mechanics and motion that could be addressed in any class format. A study assessing modest changes to a standard lecture would be helpful and approachable to many instructors while being novel in its simplicity. In my study, I actively alternate between lecture and problem solving in class, directed questions, and think-pair share exercises. In doing so, the standard lecture format is broken up by deliberately timed switching to active learning methods. Most importantly, I wish to demonstrate that instructors do not need a radically different format to make class time significantly more effective. They can use existing or classical lecture materials that already contain opportunities for well-timed interactions.

Literature Review

Difficulties in Teaching Biomechanics

Huston (1999) articulates many of the issues that make biomechanics difficult to teach. From my perspective, the two most critical aspects are: (1) students not having sufficient prerequisite background when starting a biomechanics course, and (2) students not understanding the expectations on them when they "discover how broad biomechanics is" leading to them "often [becoming] discouraged." Class time should be set aside as needed to cover any remedial review: either through an intense and thorough review of mechanics or basic biological concepts that students might be missing entirely. Huston also suggests that the main application areas of biomechanics are performance, injury, and rehabilitation. Therefore, learning objectives should coincide with these areas. Yet a key issue with teaching biomechanics is the integral connection to basic mechanics. Knudson (2016) observes: "Teaching Newtonian mechanics in physics and an introductory biomechanics class can be difficult given that many students fear the course and often avoid it as long as possible." Framing problems in a biomechanics classroom is difficult, especially in classes that are required to cover multiple biological systems. For instance, in the case of understanding the physics of walking, one could decide to lecture on the physics of joints, tendons, and muscles, or to focus on the kinematics and general forces of moving a leg. Since an entire class could be devoted to this subject and interrelating the different functional scales, problems must be framed clearly to avoid distractions and must clearly identify an objective, such as understanding injury, that can be realized within the limited framework.

Another key aspect in engineering is estimating

quantities that are not known exactly. Such estimations are usually made using information from available evidence and literature but may also need to be done rapidly in the field with incomplete information. Shakerin (2006) provides activities to help assess and practice estimation problems for engineering students. In this study, students demonstrated a poor ability to make estimations. Students struggled due to two elements of estimation: (1) by default these are situations where they do not have all information required, and (2) there is not necessarily a right answer to get the question "correct". When students were tasked with estimating quantities for which known values are available they are usually not accurate. Estimations are often required to form the boundary conditions of biomechanical problems. Therefore, during my module, students need to estimate quantities while they learn to setup and solve complex mechanics.

Specific Teaching Methodologies

Scott, Volz, and Athanasiou (2006) present a course project that heavily uses a teaching as learning strategy. Students must learn a mechanical model and explain it to the class using a classic lecture style: with a whiteboard or chalkboard and colored markers or chalk. The obligation to "prepare others to use information immediately to solve problems" is cited as a main driver of student learning over simply giving a PowerPoint presentation, which may cause students to they think they are simply being evaluated (Scott et al., 2006). Overall, students found this experience worthwhile, even when surveyed one year later. Students state that the main benefit was learning how to teach a lecture and how much work is required to do so, with one student explicitly stating that it forced him to learn material taught in the course. The effectiveness of student teaching can be capitalized in any classroom without needing to flip the classroom completely. Students can act as teachers when class time is provided for think-pair-share (TPS) activities. In this case, students are aware during the "think" stage that they will need to present their approach to a partner and not only work on the problem but think about how they can articulate their own successes and confusions. An appropriate time to use TPS could be at the start of solving a problem, where students can complement each other by inherently using different methods to reach the same end or by getting stuck at different parts of a problem.

Another challenge in teaching biomechanics is that often our final models contain many complex parts that would be overwhelming to learn all at once. An equally competing problem is that smaller aspects of

the model system may seem meaningless without this context. Clyne and Billiar (2016) state: "In problem based learning, the problem is introduced first, and it then becomes the motivation for learning the subject material." Their study focuses on a student-centered approach to problem based learning where the professor is a facilitator to a larger discussion. Their study states that faculty who are accustomed to lecturing can slowly add active learning sections to their lectures. Presenting the fully realized problem at the start of a class or course module is a method itself and does not necessitate an active learning strategy. However, assessing students' initial understanding of the problem can help the instructor use the remaining lecture time more efficiently, in addition to the use of TPS.

Savinainen and others (2013) investigate different methods of representing physical problems. Their study also identifies the value of pairing students, which allows them to feel comfortable making mistakes in front of their peers. Overall, students should be encouraged to draw depictions of physical systems that are most appropriate for representing that system, such as using free body diagrams that show forces or kinetic diagrams that show accelerations.

Methods

Constraints

The time available for this module was two 75-minute class periods. As a new BME class in the new undergraduate program, the class consisted of just five students. The course had already addressed a wide range of course materials, including fluid shear stress and extracellular signaling, and the students were already accustomed to visiting instructors. Students were all juniors with engineering backgrounds that had taken physics but not engineering dynamics.

Class Time Strategy

Normally, I would prepare a lecture covering the desired materials and have directed questions ready to ask the students. This is similar to most traditional engineering lectures where the instructor is speaking for most or all of class. My goal in designing these class periods was to incorporate the benefits of traditional lecturing, concise communication, and teaching by example, while mitigating the negatives of overwhelming attention spans and lack of confidence building exercises. This was done by incorporating brainstorming, think-pair-share, and one-minute reflection papers, in addition to pre-planned questions. To build an appropriate dynamics background three dynamics problems were taken from a
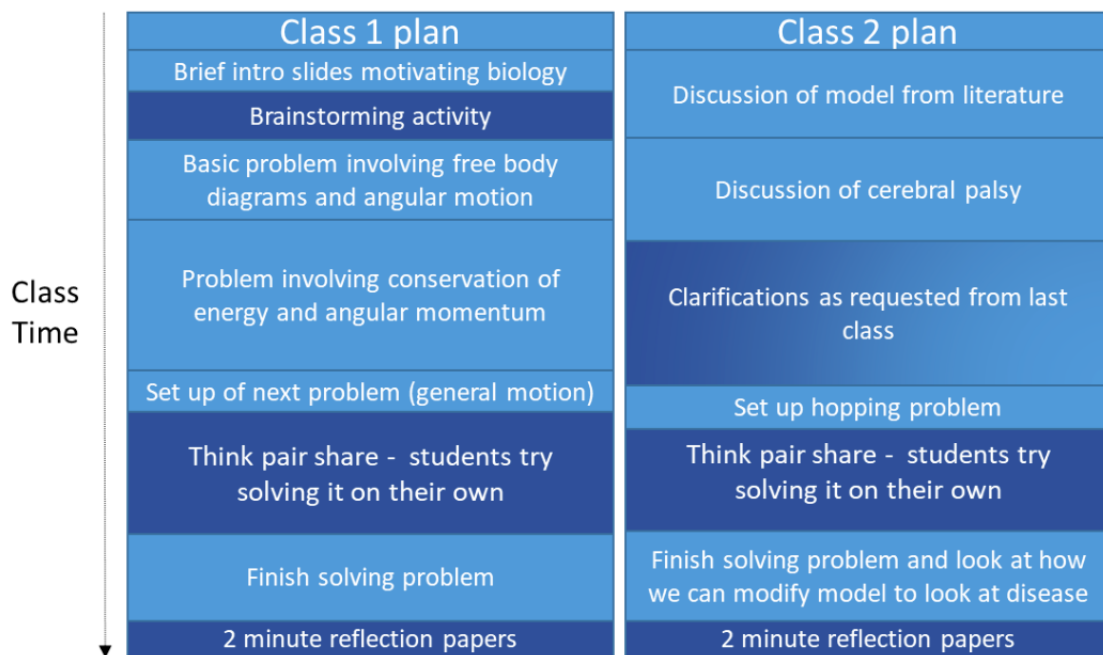
Figure 1: Schematics for class time planning. Shown are itineraries for each class period (75 minutes), with the size of each section qualitatively representing the relative length of each section. Dark blue sections indicate a deliberate usage of active learning.

standard engineering text (Hibbeler, 2009) that incrementally incorporate dynamics concepts useful for understanding limb motion.

Lesson Plan for Class 1

The problems incorporated into the background lecture contain basic dynamics review and concepts required to understand a complex model of the leg (Figure 1). The first problem is a simple rotating mass around a center to refresh students on angular coordinates and force balances. The second problem involves the energy from dropping linear mass being converted to the rotation of a wheel, covering conservation of energy and moments of inertia (problem 18-45) (Hibbeler, 2009). The third problem, to be first attempted by students, features a bar that is undergoing both translation and rotation (general motion), and introduces the summing of moments (problem 17-100) (Hibbeler, 2009).

Before focusing on dynamics principles with the above problems, a brief power point presentation was given to describe the basic stages of gait (compression, thrust, flight), followed by a brainstorming session about the composition and functions of legs to gauge previous student knowledge. The following prompts were provided: 1. What do you already know about legs and how they function?, and 2. What are legs made of?

Lesson Plan for Class 2

The beginning of the second class covered background for clinical motivations through study of a disease model (cerebral palsy), and analysis of a leg model from the literature to introduce students to the outputs of mechanical gate models. The remainder of the class was devoted to more dynamics based material. Clarifications on background questions from the previous class were addressed and a culmination leg model problem, involving a simplified leg with a single, femur-like member swinging to create a vertical translation.

Observation and Self-Reflection

To the best of my ability, I took notes during and after class to retain as many details about class time as possible, with a focus on student comments and perceived demeanor.

Minute Reflection Papers

Roughly two minutes at the end of class were set aside for reflection papers. The following prompts were given to students: 1) What is your biggest take away from the class?, 2) What are you still least clear about? ("muddiest point"), and 3) What would you like to do more of next time? Additionally, the following guidance was given: "Chose some (or all) of the following questions to respond to," and orally "responses can be strings of words, sentences, or a short paragraph." These instructions were given to the students to indicate that they may willingly determine

Figure 2: Minute reflection paper responses from each student. Specific pieces of advice or requests by students are bolded. Most other feedback was either descriptive or simply positive acknowledgment.

#1	- GOOD REVIEW OF DYNAMICS CONCEPTS - LOOKING FORWARD TO MORE APPLIED PROBLEMS (MECHANICS OF WALKING) NEXT CLASS - APPRECIATE STEP-BY-STEP BOARDWORK BUT **SOME MORE ARITHMETIC STEPS COULD BE SKIPPED TO SAVE TIME**
#2	- ENJOYED REVIEW OF BASIC PHYSICS - STILL CONFUSED ABOUT **[EQUATION] FROM PROBLEM 3 [WITH ACCELERATIONS], GO OVER NEXT TIME** - GOOD JOB OVERALL
#3	- TAKEAWAY, BASIC SUMS FROM DYNAMICS - LEAST CLEAR, NOTHING, JUST NEED MORE PRACTICE - WANTS, MORE PROBLEMS NEXT TIME
#4	- TAKEAWAY, LEARNING TO SOLVE DYNAMIC PROBLEMS - I WOULD LIKE SOME MORE PRACTICE **USING REFERENCES RELATIVE TO THE ORIGIN**, AS I AM NOT ENTIRELY CLEAR ABOUT THAT
#5	- GOOD REVIEW OF SOME BASIC DYNAMIC PROBLEMS - MAYBE **DESCRIBE THE CONCEPT A LITTLE BIT MORE BEFORE SOLVING PROBLEMS**

the amount of effort they decide to put into their responses.

Extra Office Hours

An additional session of office hours was held to clarify issues from the second class period and to assist students with their lab reports.

Assessment of Lab Assignment

I assessed whether students attempted to apply concepts taught in my module on their lab report designed by the instructor.

Results

Observations and Reflections from Class 1

During the brainstorming session, students were engaged and responsive but many ideas did not flow into the subsequent lecture material. The prompts could be improved to direct more specific questions. Light blue lecture sections (shown in fig. 1) prompted multiple questions from students in addition to my directed questions. This may have been due to setting an interactive tone with the brainstorming activity. The small class size may also have simply been more conducive to participation. Students showed increased confidence due to student interaction during the think-pair-share activity. While they did not get far on their own, students wrote down basic concepts to start the problem. Once paired, they were able to combine their ideas to create general governing equations for the problem. I often heard students saying "I thought about using this approach but I wasn't sure,"

indicating a willingness to show uncertainty with their peers. The minute reflection papers were by far the best part of the class while also requiring the least effort from both myself and the students. I received specific feedback on how to improve for next time and saw students thinking about their own learning and limitations (Figure 2).

Observations and Reflections from Class 2

This class would have greatly benefited by not covering the literature model at the start of class. While this paper (Case & Raibert, 1984) was intended to demonstrate the numerical behaviors of a hopping model, the details of the specific study took too much time to cover without directly adding to the students' ability to solve their own leg problem. Students asked fewer and fewer questions as this discussion went on and were mainly confused about varied concepts that were not established in the last class.

The discussion of cerebral palsy yielded strong participation as students were very curious about the different physical manifestations of the disease. One student specifically asked about a person they had known who had a constant toe walk. Ideally, this portion of lecture would have also tied into a disease model of walking but in this case served as motivation to learn about normal walking mechanics. The clarifications from the previous class successfully addressed students' concerns. Students were receptive to my clarifications to their previous questions and asked new questions that referred to more advanced details of the problems, or to other hypothetical prob-

lems, demonstrating a higher level of understanding.

Going into the culminating dynamics problem, I realized that I did not have enough time to set up the problem and work through it with the students. Unfortunately, I needed this time to carefully frame the problem as being numerically solvable and not analytically solvable like the previous problems. Because I was not able to convey this clearly, students were able to discuss basic aspects of the problem including force balances and estimation of parameters but were quickly overwhelmed and lost focus, stating that they were unsure what we were "solving for." The think-pair-share activity yielded no progress as the students did not have a sufficient starting point from which to continue the problem.

Observations from Office Hours

Office hours took place after a hopping lab at Ithaca College, where students were able to hop on a force plate and collect both body height trajectories and forces/moments from the force plate. I related the data the students had acquired to the culminating dynamics problem, showing which unknowns could be filled by the time-series data. Students appreciated seeing the value in the overwhelming amount of data they had collected. However, the main area where students wanted help was in how to use numerical methods to calculate integrals and derivatives based on their data, including how to preprocess their data (e.g. resampling) for these calculations. Such topics are generally covered early on in engineering curricula and are often assumed as prior knowledge for later courses.

Minute Reflection Papers

The minute reflection papers provided me with rapid feedback on my teaching. It was also a clear moment when students were revisiting the material from the lecture and thinking about their own learning. The most useful feedback fell into two categories: specific advice on how to alter lecture strategies and concepts the students needed more help understanding (Figure 2).

Student Integration of Concepts

Students were able to use the topics introduced in the first three background problems. However, they did not attempt the portion of the assignment requiring the full model. They did demonstrate the ability to apply the additional materials covered in office hours, mainly numerical calculation of integrals and derivatives, to solve the remainder of the questions in their lab report.

Discussion

The main challenge for presenting this material is the short time in which it needs to be presented in the course and the variety of other materials covered by the course that do not overlap with walking mechanics. The first class period went well because the majority of time was focused on dynamics principles with each problem building in additional concepts. Students demonstrated that they were able to confidently address the problem on general motion and ask follow-up questions that led to further understanding in the next lecture.

In my assessment, the shortcomings of class two arose from having too much material to cover given the constraints of this module within the course. This had the secondary effect of also taking time away for me, as the instructor, to prep course materials, effectively making it harder for me to implement the essential parts of the class well with my own time constraints. Given that this was the first time the course was being offered, and my first time teaching the material, the extra literature and disease discussion could have been removed. However, in the future I feel that the concepts from the literature on visualizing hopping model behaviors could simply be taught in the context of the culminating dynamics problem, without the need to examine a different model. Further, the discussion of disease models of walking could remain but be moved to the end of class as a lower priority to understanding normal walking mechanics. The review from the previous class could still be used as an engaging part of the lecture, but at the very start of class, while the think-pair-share activity could continue to break up the lecture after half way through class time. An improved version of the second class period that is directed specifically at preparing students for their hopping lab will be able to provide a more complete experience for students in later years.

Lastly, I would strongly recommend the use of reflection papers in any engineering classroom for students to give feedback to instructors in a way that benefits themselves, the instructor, as well as future students. This feedback could be sought from students at regular intervals during a semester or ad hoc by the instructor with minimal preplanning or preparation.

Acknowledgements

cation) as awarded to the Department of Biomedical Engineering at Cornell University.

References

Case, O., & Raibert, M.H. (1984). Hopping in Legged Systems-Modeling and Simulation for the Two-Dimensional, (3), 451–463.

Clyne, A.M., & Billiar, K.L. (2016). Problem-Based Learning in Biomechanics: Advantages, Challenges, and Implementation Strategies. Journal of Biomechanical Engineering, 138(7), 70804.

Huston, R.L. (1999). What I Learned in 25 Years of Teaching Introductory Biomechanics. International Journal of Engineering Education, 15(4), 240–242.

Knudson, D. (2016). Research on Learning Introductory Biomechanics. International Conference on Biomecanics in Sports.

Hibbeler, R.C. (2009). Engineering Mechanics: Statics and Dynamics (12th ed.).

Savinainen, A., Nieminen, P., Mäkynen, A., & Viiri, J. (2013). Teaching and evaluation materials utilizing multiple representations in mechanics. Physics Education, 48(3), 372.

Scott, C.C., Volz, T., & Athanasiou, K.A. (2006). Learning How to Teach Continuum Biomechanics: See One, Do One, Teach One. International Journal of Engineering Education, 22(1), 183–187.

Shakerin, S. (2006). The Art of Estimation. Internal Journal of Engineering Education, 22(2), 273–278.

Teaching Scholar Bootcamp: A Professional Development Technique to Increase Self-Efficacy and Confidence in Graduate Student Instructors

Larry L. Bowman[1,2,*], Kelly J. Culhane[2,3], Ahrume Julie Park[2,4], Samantha Lin[5], and Kaury Kucera[2]

[1] Department of Ecology and Evolutionary Biology, Yale University
[2] Center for Teaching and Learning, Yale University
[3] Department of Molecular Biophysics and Biochemistry, Yale University
[4] Department of Cell Biology, Yale University
[5] Department of Molecular, Cellular, and Developmental Biology, Yale University
[*] Corresponding Author: larry.bowman@yale.edu

Abstract

Many studies examine the effects of graduate student and post-doctoral teaching development on student outcomes, and even more studies examine faculty development programs' effect on student outcomes. We aim to uncover the effects graduate student and post-doctoral teaching development programs have on the self-efficacy and confidence of the individuals undergoing development, the teaching assistants themselves. How does teaching training effect the teaching assistant, and is it effective for long-term implementation of teaching techniques? To understand the impact of teaching assistant development programs, we held a short, intensive bootcamp that covered inclusive teaching and diversity topics for graduate students and post-doctoral teaching scholars. We then observed the graduate students and post-doctoral teaching scholars in their classrooms throughout the semester noting inclusive teaching and active learning techniques used within the classroom, even if the strategies used were not explicitly taught in the bootcamp. We tested whether priming teaching assistants with the knowledge and vocabulary to talk about diversity, inclusivity, and active learning in the classroom would increase their confidence and self-efficacy in implementing these strategies into their teaching and in their classrooms. Our results show that trends in short-term outcomes are measurable but that priming alone and one-time pedagogical training sessions are not enough to sustain long-term implementation of these strategies. This study supports the need for continuous and possibly long-term, cohort-style teaching development for graduate teaching assistants and post-doctoral teaching scholars.

Introduction

At many universities, there are centers for teaching and learning or schools of education that develop graduate students in pedagogical theory and practice for the explicit purpose of preparing them to teach undergraduates while studying at the same institution. However, many more universities lack the structure, resources, or staff to support formal programs in pedagogical development for a large body of burgeoning young educators [1]. To address this issue, we developed a short-form, low-cost, and low-effort bootcamp-style pedagogical training session and assessed its ability to change the pedagogical practices, effectiveness, and confidence of graduate student teaching assistants (GTAs) and post-doctoral teaching assistants (PTAs). We believe this method of pedagogical and professional training will benefit new instructors both in the classroom and on the job market, leveraged as professional development in leadership and inclusivity and diversity strategies in organizations.

This style of pedagogical workshop was developed out of an increasing literature that intentional directed study in pedagogy helps prepare graduate students for teaching duties [2] and that providing graduate students with the correct vocabulary to talk about their professional development increase self-efficacy in teaching [3]. However, while most research focuses on semester-long or half-semester courses this study aims to see if similar results can be gained from a primer or bootcamp-style three-hour workshop.

Our perspective for undertaking this project was informed by the idea that priming students' previous knowledge of material is a commonly used review technique that instructors use as a connective tool for new learning experiences. We want to utilize this priming effect for graduate student and post-doctoral teaching development and determine whether priming during teaching assistant development classes can have lasting effects continuing throughout the semester.

Here we focus primarily on two specific opportunities in the classroom that have been show to increase student outcomes: active learning [4-8] and inclusiv-

Table 1.Overview of Teaching Scholar Bootcamp schedule and topics covered.

Table 1. Overview of Teaching Scholar Bootcamp Schedule and Topics Covered

Time	Topic	Specific Elements
20 min	Introductions	Icebreakers; Inclusive strategies from day 1
10 min	Theory & Pedagogy	Overview of literature; learning & learning communities
60 min	Microteaching 1	Active learning techniques; Model & practice
30 min	Lesson Planning	Classroom strategies for efficiency & effectiveness
60min	Microteaching 2	Inclusivity and diversity strategies; Model & practice

ity and diversity strategies [9-11].

Finally, our focus in this study is on GTA/PTA professional and educational development outcomes. Most current literature implements GTA/PTA interventions and looks to student outcomes as evidence for effectiveness. We aim to explicitly measure the implementation of our techniques by the GTA/PTAs themselves under the assumption that using these techniques in the classroom leads to subsequent increases in student outcomes as is supported by the literature [8,9].

The main objective of this research was to determine if GTA/PTAs that were trained to use current, evidence-based knowledge and language of active learning techniques and inclusive teaching strategies were more likely to implement those strategies in their classrooms and if they were more likely to continue to implement those strategies over the course of the semester.

Methods

Bootcamp Description

The bootcamp and subsequent research were undertaken at Yale University in Spring 2017 as part of the foundational biology course sequence. Graduate Teaching Assistants and Post-Doctoral Teaching Assistants apply to become Teaching Scholars, where they develop and deliver curriculum, for one of four six-week-long modules covering different topics in foundational biology. Though Teaching Scholars develop and deliver curriculum, we henceforth will refer to them as GTAs and PTAs, respectively, for continuity with the literature. GTAs and PTAs for the second foundational module in Spring 2017 were offered the opportunity to opt-in to participating in the bootcamp. Nine out of ten GTAs and PTAs teaching foundational biology modules took advantage of this professional development opportunity.

Foundational biology modules at Yale University are structured in a cohort design where GTAs and PTAs meet weekly with the professors and course coordinators to discuss upcoming lesson plans, challenges and opportunities, and logistics, attendance, and grading for the module. Our intervention aimed to augment this already important and cultivated learning community by priming GTAs and PTAs with language and concepts focused around active learning and inclusivity and diversity in the classroom. GTAs and PTAs receive various forms of professional and pedagogical development at Yale university varying by department from no formal training to extensive multi-week workshop classes.

Our intervention consisted of a three-hour bootcamp style primer on how to implement active learning and inclusivity/diversity strategies into the introductory biology classroom. Our bootcamp intervention took place one week before GTA/PTAs began teaching their modules. We arranged the content of the bootcamp around modeling the strategies we were teaching (Table 1). For example, the opening exercise was an icebreaker designed to learn everyone's name and an interesting fact about them, incorporating both active learning and the inclusivity strategy of knowing names and personal facts. The content of the bootcamp consisted of a short lecture on learning communities, the cognitive aspects of how we learn, and why we chose active learning and inclusivity/diversity strategies as focal areas for the bootcamp. We provided GTAs and PTAs with lists of evidence-based active learning techniques [5] and strategies to increase inclusivity and recognize diversity in the classroom [9]. We, as workshop facilitators, modeled some of these strategies through microteaching before asking GTAs and PTAs to plan microteaching lessons [12] around one of the strategies we taught them. We provided 30 minutes of lesson-planning time for GTAs and PTAs to work in

pairs to co-facilitate using one of these techniques to teach any concept, pertinent or not to introductory biology.

Pairs of GTAs and PTAs taught for five minutes and then were given three minutes for constructive feedback from the entire group with the goal of providing a framework of how the cohort could dialogue about, give, and receive critical feedback regarding their teaching after the bootcamp. GTAs and PTAs completed two microteaching exercises, one incorporating an active learning technique [5] and one incorporating an inclusivity/diversity strategy [9] with different partners with the explicit intent of increasing rapport between GTAs, PTAs, and the course coordinator.

The stated goals of the bootcamp were to 1) incorporate techniques for active learning and inclusivity in their classrooms and 2) discuss and leverage their professional development in a meaningful way on the job market/interview process.

Participants

The GTAs and PTAs participating in this bootcamp were voluntary opt-ins for extra professional and pedagogical development. Seven participants were graduate students; two participants were postdoctoral scholars/fellows. Of those who responded to opt-in polling, 86% had some teaching experience prior to our bootcamp intervention. However, 57% had never received any type of pedagogical training, and of the 43% that had some form of pedagogical training, 100% had attended fewer than three such events. Explicit demographics were not collected, but perceived gender composition of Teaching Scholars was 33% male, 67% female.

Classroom Observations

In the introductory Biology module series, each GTA or PTA leads a 50-minute section of 10-20 students, where he, she, or they present, review, or expound upon the course's content material. For the module affected by the intervention, GTAs and PTAs led built-in active learning sections composed of review, problems sets or homework problems, and enrichment exercises.

We observed each GTA or PTA three times over the course of the six-week module. To avoid perceived effects of pre-or post-exam anomalies, we observed GTA and PTAs during the first, second, and fourth weeks of the module, avoiding the midterm and final exam sections.

McDougal Graduate Teaching Fellows at the Yale Center for Teaching and Learning, who have formal training in peer-observation and feedback, conducted

observations in these sections using a modified CO-PUS [13] which included a map of the classroom, perceived demographics of undergraduates, and a list of concluding questions in order to capture many aspects of the classroom. The results reported in this study are from the map of the classroom portion of the observation that included perceived demographics of each student, how many times student spoke, asked, or answered a question in the whole-group setting, and how many times the GTA/PTA moved around the classroom. This form also captured the number and type of questions asked by the GTA/PTA throughout the section. Questions were categorized as implicit, explicit, or open-ended [14].

We aggregated data to observe overall trends of the bootcamp on the entire cohort, and we looked at individual GTA/PTA responses to elucidate whether there was a strong buy-in effect from the GTA/PTAs. We performed all statistical analyses in JMP Pro 13 [15].

Metrics for Active Learning and Inclusivity Assessment

We used GTA/PTA movement in the classroom as a proxy for GTA/PTA confidence in implementing active learning techniques and teaching. Because active learning techniques are structured into the foundational biology curriculum, we measured GTA/PTA movement as evidence that GTAs/PTAs were not only assigning active learning activities, but were themselves engaging with students in the classroom individually and in small groups. Additionally, we used number of questions asked by students and asked by GTA/PTAs as a measure of active learning with the rationale that more exchanges between the instructor and students in the classroom leads to deeper learning [16].

We used the number of times the GTA/PTAs asked implicit versus explicit or open-ended questions as a proxy for inclusive-teaching strategies. We defined implicit questions as questions that engage the students' thought process and make student metacognition apparent, e.g., "How did you come to that answer and why is it better than another?" Explicit questions are those that test only specific content knowledge, e.g., "What is the next step in this process?"; open-ended questions are those that are broad, non-specific, usually addressed to the entire class or are checks for understanding, e.g., "Does everyone understand the cell life cycle?" To further understand the nature of inclusivity over time in these classrooms, we used the number of times perceived minority gender and race students spoke in the classroom.
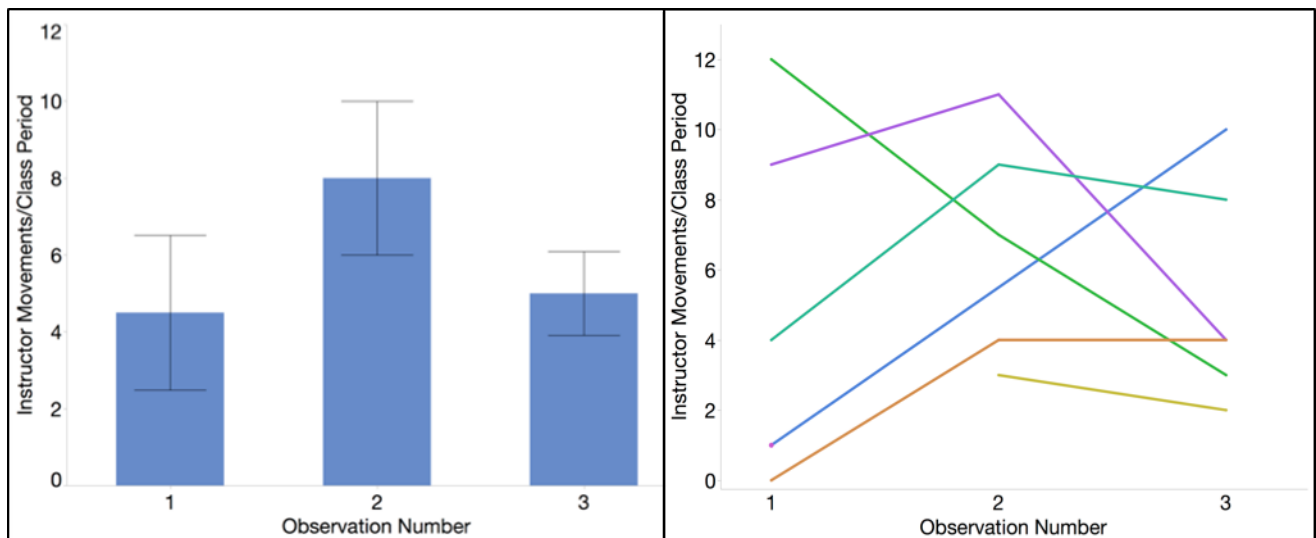
Figure 1: A. Mean changes in instructor movements in the classroom throughout the semester. Observation number corresponds to weeks 1, 2, and 4 respectively during a six-week module of an introductory biology course. Error bars represent +/-1SE from mean, N=9 observations per week; N_{total}=27. B. Changes in instructor movements in the classroom throughout the semester by individual GTA/PTA. Each color represents a different GTA/PTA. Observations correspond to same weeks as A. N=6 individuals, 1 observation/individual/week; N_{total}=17.

Results

Measures of Active Learning Techniques

GTA/PTA movements.

Graduate and post-doctoral teaching assistants exhibited initial increases in movements in the classroom from the first observation to the second observation, followed by a subsequent decline in the final observation (Fig. 1A). Fifty percent of GTA/PTAs ended the semester module with more movements in the classroom than when they started, while fifty percent showed decreases over the course of the semester module (Fig. 1B).

GTA/PTA-posed questions.

Graduate and post-doctoral teaching assistants did not exhibit any change in number of questions posed between observations one and two and asked significantly more questions in the final observation (Fig. 2A). The majority of GTA/PTAs asked more questions as the semester progressed, though one-third of GTA/PTAs asked fewer questions at the end of the semester than they did at the beginning (Fig 2B).

Measures of Inclusivity and Diversity Strategies

Types of GTA/PTA questions.

Graduate and post-doctoral teaching assistants were observed to ask significantly more explicit than implicit questions in their classrooms in all observations (Fig. 3). They also asked significantly fewer implicit questions and significantly more explicit ques-

tions as the semester progressed. There was no significant change in open-ended questions posed by the instructors across the semester (Fig. 3).

Individually, only one-third of GTA/PTAs asked more implicit questions at the end of the semester than they did at the beginning (Fig. 4A). Meanwhile, 89% of GTA/PTAs asked more explicit and open-ended questions at the end of the semester than they did at the beginning of the semester (Fig. 4B).

Student-posed questions.

There was no trend among all classes in number of questions asked by students across the semester that corresponded to minority race (defined as perceived traditionally disadvantaged race) or minority gender (defined as perceived traditionally disadvantaged gender or non-binary gender) (Fig. 5).

There was a highly significant trend in percentage of questions asked by traditionally disadvantaged gender students depending on the gender composition of the classroom (Fig. 6A). When corrected for gender composition, the majority gender present in the classroom was highly correlated with asking the majority of the questions in traditionally disadvantaged genders and vice versa, i.e., in majority male-presenting classrooms, male-presenting students asked the majority of questions; in majority female-and non-binary-presenting classrooms, female-and non-binary-presenting students asked the majority of questions. Meanwhile, there was a significant trend in percentage of questions asked by stu-
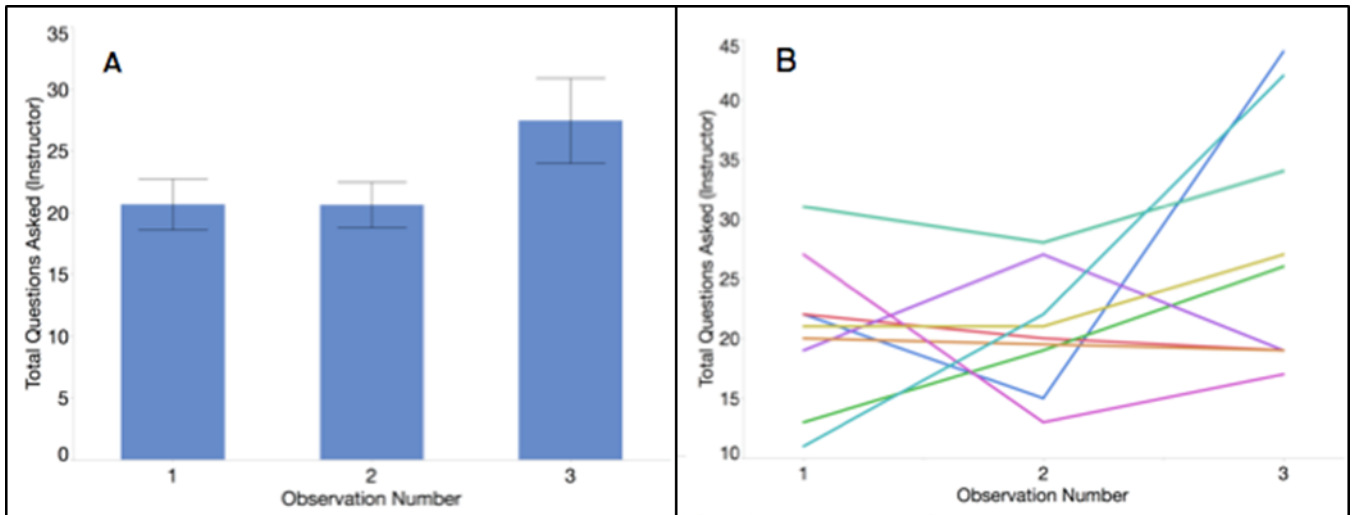
Figure 2: A. Mean changes in total number of GTA/PTA-posed questions throughout the semester. Observation number corresponds to weeks 1, 2, and 4 respectively during a six-week module of an introductory biology course. Error bars represent +/-1SE from mean, N=9 observations per week; N_{total}=27. B. Changes in total number of GTA/PTA-posed questions throughout the semester by individual GTA/PTA. Each color represents a different GTA/PTA. Observations correspond to same weeks as A. N=9 individuals, 1 observation/individual/week; N_{total}=27.
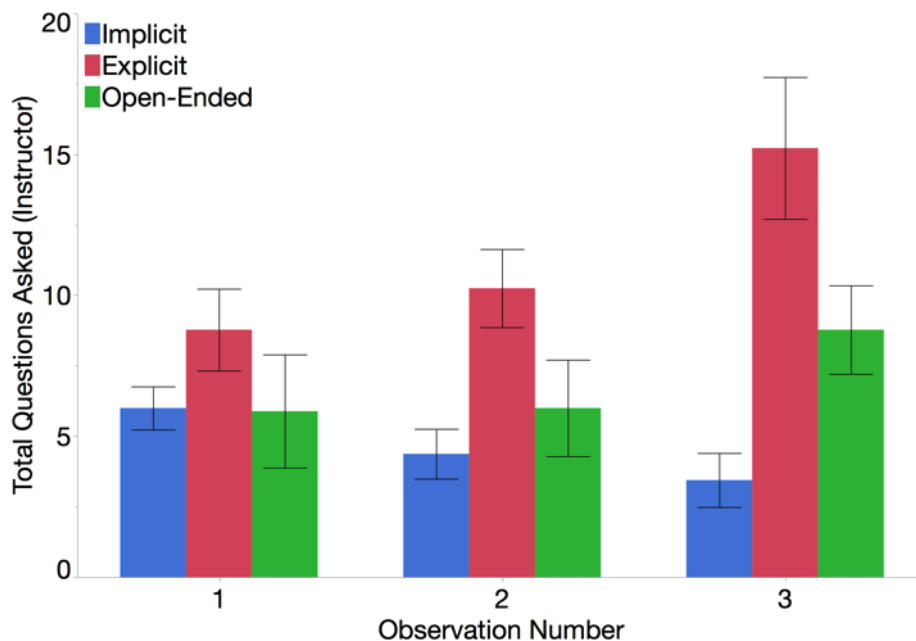
Figure 3: Mean changes in total number of GTA/PTA-posed questions throughout the semester by question type: implicit (blue), explicit (red), and open-ended (green). Observation number corresponds to weeks 1, 2, and 4 respectively during a six-week module of an introductory biology course. Error bars represent +/-1SE from mean, N=9 observations per week; N_{total}=27.
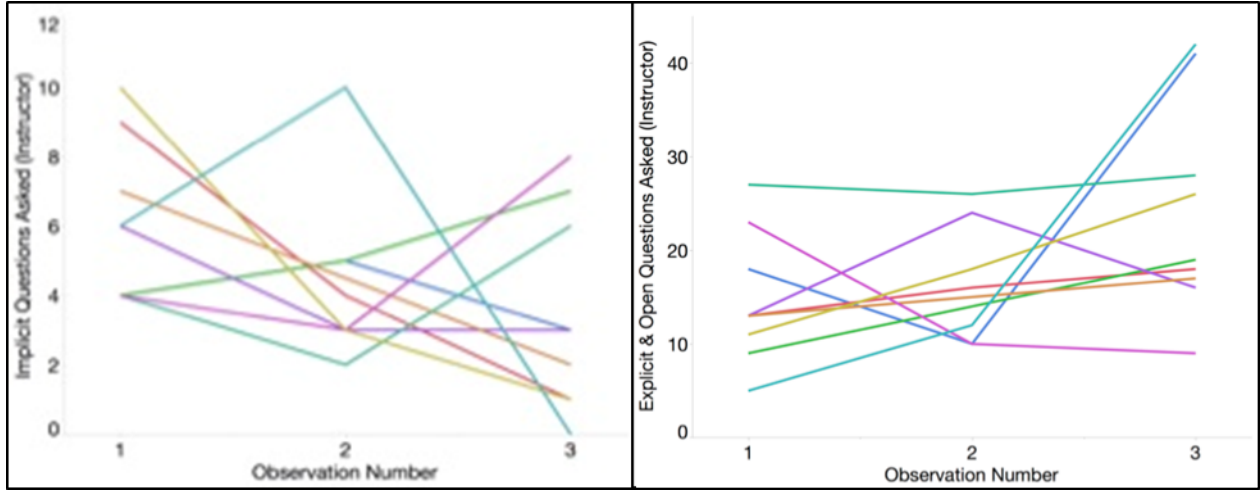
Figure 4: A. Changes in total number of GTA/PTA-posed implicit questions throughout the semester by individual GTA/PTA. Observation number corresponds to weeks 1, 2, and 4 respectively during a six-week module of an introductory biology course. Each color represents a different GTA/PTA. N=9 observations per week; N_{total}=27. B. Changes in total number of GTA/PTA-posed explicit and open-ended questions throughout the semester by individual GTA/PTA. Each color represents a different GTA/PTA. Observations correspond to same weeks as A. N=9 individuals, 1 observation/individual/week; N_{total}=27.
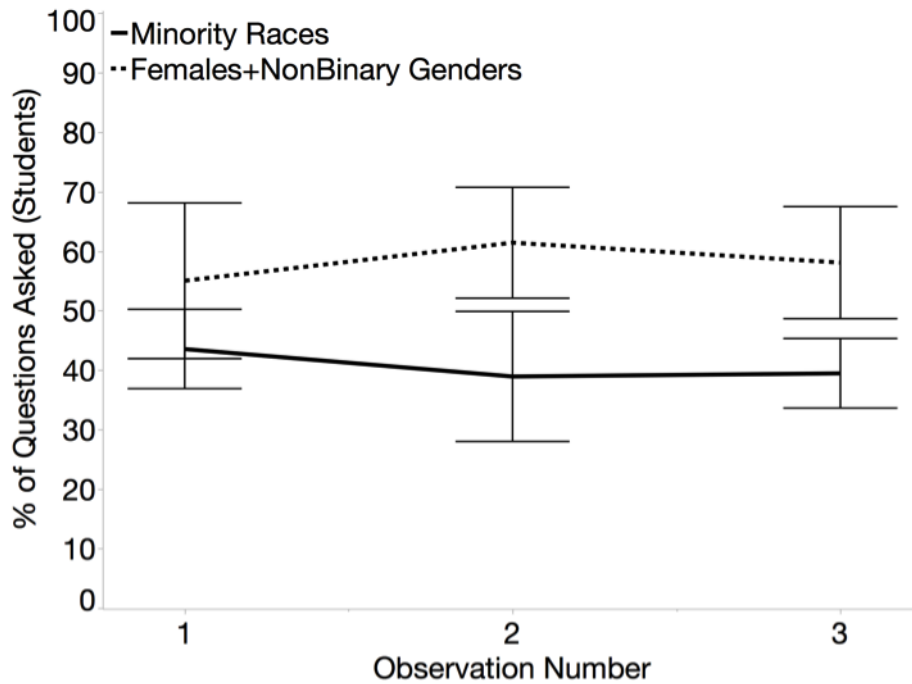
Figure 5: Mean changes in percentage of student-posed questions throughout the semester by minority races (solid) and genders (dashed). Minority races were defined as those perceived as traditionally disadvantaged or non-White. Genders were defined as those perceived as traditionally disadvantaged (female) or non-binary genders. Observation number corresponds to weeks 1, 2, and 4 respectively during a six-week module of an introductory biology course. Error bars represent +/-1SE from mean, N=9 observations per week; N_{total}=27.
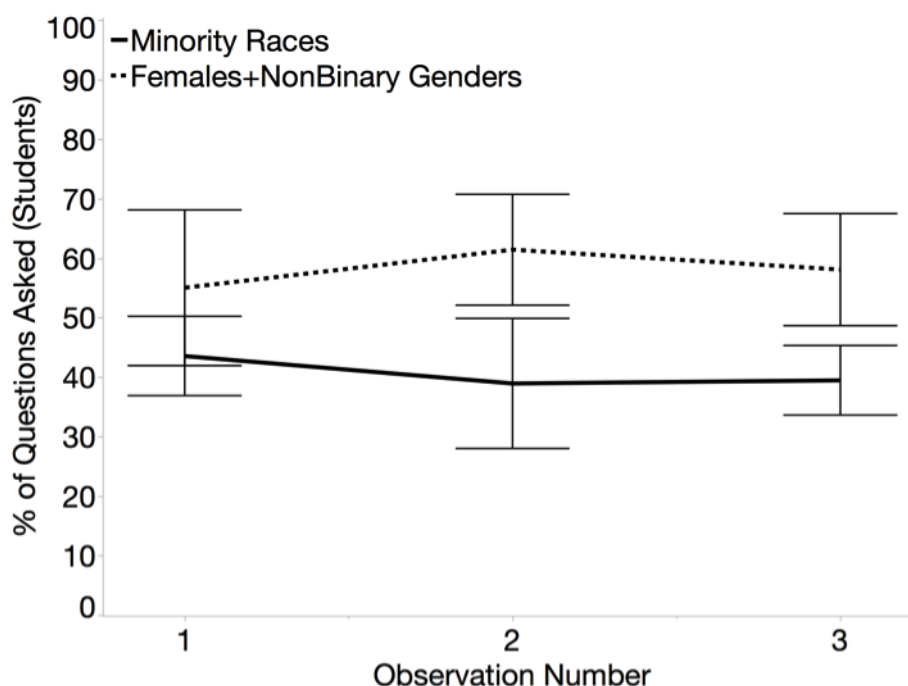
Figure 6: A. Bivariate fit of gender ratio and percentage of minority gender student-posed questions. Polynomial fit R^2=0.90; ANOVA P¡0.001; N_{total}=21. B. Bivariate fit of diversity ratio and percentage of minority race student-posed questions. Linear fit R^2=0.57; ANOVA P=0.0029; N_{total}=13.

dents of color depending on the racial composition of the class, but one that can be accounted for by the racial composition of the class, i.e., when there are more students of color, more students of color ask questions, but not in a manner inconsistent with students of majority races (Fig. 6B). If students of color asked a consistently low (or high) percentage of questions regardless of the composition of the class, we would expect to see no relationship between composition and percentage of questions asked, i.e., a flat line or no trend. Here, the trend is explained by more students of color are in the class and proportionally ask more questions, thereby, race composition has no effect when considering percentage of questions asked.

Discussion

The Teaching Scholar Bootcamp was conceived as a conceptual framework for departments and universities to provide educational professional development to their graduate and post-doctoral teaching assistants that was relatively short in nature and cheap to produce. We wanted to keep the time and effort commitment intentionally low for participants and facilitators to allow institutions without formal staff to facilitate such professional development to replicate our process and gain many of the benefits of a more structured or formal pedagogical development course. We do believe that our bootcamp achieved our objectives of 1) incorporating techniques for active learning

and inclusivity in introductory biology classrooms and 2) discussing and leveraging their educational professional development in a meaningful way on the job market/interview process.

With specific regard to our first objective of seeing active learning and inclusivity strategies incorporated into section design in this course, the specific effects of our intervention are present but their sustainability seems to wane over the course of the semester for both active learning and inclusivity strategies.

Active Learning Techniques

The dominant trend we observed in classrooms was one of increased movements in the classroom followed by a subsequent return to previous amounts of movement (Fig. 1A, B). A movement was recorded when GTA/PTAs moved and guided students during active learning techniques in the classroom. Generally, as the semester progressed, GTA/PTAs did not decrease the amount of active learning techniques in their classrooms, but we did see a marked decrease in GTA/PTA moving and guiding students. Individual trends varied with some GTA/PTAs continuously increasing moving and guiding throughout the semester. This is important to note because the bootcamp intervention may not be sustainable for all individuals but appears to have lasting effects for some individuals. We believe the sustainability of these effects could be increased by continuing conversations around active learning in the classroom

throughout the semester in a more formal learning community for sustained support and strategies to use in the classroom.

In terms of confidence in active learning techniques, we saw significant increases in number of GTA/PTA-posed questions across the semester (Fig 2A, B), suggesting that GTA/PTAs become more comfortable employing active learning techniques as they practice them in the classroom. Individual trends vary, but the vast majority of GTA/PTAs showed increases in the number of questions they asked across the semester, suggesting that incorporating the importance of active and engaging lecturing and teaching into the bootcamp helped GTA/PTAs' confidence in including these techniques in their classrooms.

Inclusivity Strategies

The dominant trend we saw for GTA/PTAs incorporating inclusive teaching techniques was a decline in the use of inclusive strategies over the semester. As a measure of inclusivity, we binned questions into implicit, explicit, and open-ended, with implicit questions being our measure of inclusivity because they particularly value individual thinking and metacognition in the student and between the student and instructor. As the semester progressed, GTA/PTAs asked significantly fewer implicit questions (Fig. 3), despite an overall increase in number of questions posed by the instructor. Individual trends using these strategies were mostly negative with the majority of GTA/PTAs asking fewer implicit questions that when they began the semester (Fig. 4A, B). We also see a significant increase in number of explicit questions posed by the instructor as the semester progresses (Fig. 3).

These results suggest that the effects of the bootcamp intervention wane after several weeks, at least in terms of this measure of inclusivity. However, considering that classrooms did not show marked decreases in our other measure of inclusivity in the classroom — questions asked by minority genders and races — it also might be possible that inclusivity strategies may be more effective and necessary in the beginning of a semester to build rapport and community among students. Graduate and post-doctoral teaching assistants may not ask as many community building or metacognition revealing questions toward the end of the semester because they have a rapport with individual students and understand their metacognition or because there is rapport among the learning community created within the classroom at that point in the semester. It is therefore worth considering that our intervention had a very positive im-

pact in terms of inclusivity in these classrooms, despite the fact that we saw a decrease in inclusive questions posed by the instructors. However, we do concede that continuing these questions throughout the semester would be more ideal, and therefore recommend a continued conversation of how to increase diversity of opinions and metacognition in our students throughout the semester and not just to create an inviting environment in the beginning of the semester.

Interestingly, we saw no increase or decrease in the number of questions posed by traditionally disadvantaged races or genders in any of the classrooms (Fig 5). However, there is a highly significant effect of gender composition in the classroom (Fig. 6A). Whereas race composition has no effect in our classrooms of which students are more likely to ask questions (Fig. 6B), gender composition dictates very strongly which genders will ask questions. For example, if the class is majority men, nearly all questions ($>$90%) are asked by men. Meanwhile, if the class is majority women or non-binary genders, nearly all questions are asked by women and non-binary genders (polynomial fit R^2=0.90; ANOVA P <0.001). When corrected for class demographic composition, the effects of race are not present, but the effects of gender remain significant. Though disconcerting, especially considering the retention of women and non-binary individuals in STEM fields, these results highlight an important area for future research and thought when designing and planning courses in introductory STEM classes. Future research would entail constructing smaller, single-gender recitation sections that are part of a larger mixed gender class/lecture. Overall student performance increases would further support single-gender interventions as part of larger mixed gender classes and would suggest that multiple gender environments of learning are necessary for all genders.

Though innovative in design and implementation, our findings have mixed outcomes. The Teaching Scholar bootcamp clearly has many positive effects for those who took advantage of it, and we especially recommend this type of professional development for institutions that may not offer formal pedagogical development programs as a low-cost and low-effort alternative to ensure GTA/PTA success in the classroom. However, we are also cognizant that many of our results from this bootcamp disappear several weeks after the intervention. Therefore, in order to make these strategies and techniques habitual in the teaching of GTA/PTAs, we suggest a concerted, long-term learning community or cohort structure to provide GTA/PTAs with more consis-

tent priming and updates about active learning and inclusive teaching techniques. We believe that following a bootcamp intervention with informal or formal conversations around these topics would increase the sustainability of these techniques in classrooms. In final interviews and feedback, all GTA/PTAs asked for more conversations and the opportunity to troubleshoot with colleagues as an important next step in developing this pedagogical development tool.

Acknowledgments

References

1. Brownell SE, Tanner KD. 2012. Barriers to faculty pedagogical change: Lack of training, time, incentives, and...tensions with professional identity? CBE Life Sci Educ 11(4):339-46.

2. Marbach-Ad G, Schaefer KL, Kumi BC, Friedman LA, Thompson KV, Doyle MP. 2012. Development and evaluation of a prep course for chemistry graduate teaching assistants at a research university. J Chem Educ 89(7):865-72.

3. Wyse SA, Long TM, Ebert-May D. 2014. Teaching assistant professional development in biology: Designed for and driven by multidimensional data. CBE Life Sci Educ 13(2):212-23.

4. Faust JL, Paulson DR. 1998. Active learning in the college classroom. Journal on Excellence in College Teaching 9(2):3-24.

5. Foust RA. 2008. Learning strategies, motivation, and self-reported academic outcomes of students enrolled in web-based coursework. Wayne State University.

6. Meyers C, Jones TB. 1993. Promoting active learning. strategies for the college classroom. ERIC.

7. Paulson DR. 1999. Active learning and cooperative learning in the organic chemistry lecture class. J Chem Educ 76(8):1136.

8. Prince M. 2004. Does active learning work? A review of the research. J Eng Educ 93(3):223-31.

9. Tanner KD. 2013. Structure matters: Twenty-one teaching strategies to promote student engagement and cultivate classroom equity. CBE Life Sci Educ 12(3):322-31.

10. Tanner K, Chatman LS, Allen D. 2003. Approaches to cell biology teaching: Cooperative learning in the science classroom–beyond students working in groups. Cell Biol Educ 2(1):1-5.

11. Tanner K, Allen D. 2004. Approaches to biology teaching and learning: Learning styles and the problem of instructional selection–engaging all students in science courses. Cell Biol Educ 3(4):197-201.

12. Borg WR. 1970. The minicourse. A microteaching approach to teacher education.

13. Smith MK, Jones FH, 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(4):618-27.

14. Khishfe R, Abd-El-Khalick F. 2002. Influence of explicit and reflective versus implicit inquiry-oriented instruction on sixth graders' views of nature of science. Journal of Research in Science Teaching 39(7):551-78.

15. SAS Institute Inc. 2016. JMP® Pro 13 [computer program]. Version 13.0.0 (64-bit). Cary, NC.

16. Chin C, Brown DE. 2000. Learning in science: A comparison of deep and surface approaches. Journal of Research in Science Teaching 37(2):109-38.

Teaching for the First Time: The Graduate Student Teaching Experience

Elizabeth Case[1,*] and Reece Kearney[2]

[1] Department of Earth and Environmental Sciences, Columbia University
[2] Sibley School of Mechanical and Aerospace Engineering, Cornell University
[*] Corresponding Author: case@ldeo.columbia.edu

Introduction

Graduate teaching assistants (GTAs) often begin their teaching duties underprepared. They may have little or no experience with the course material or training in the classroom [1]. Moreover, support for new GTAs from professors and departments varies significantly. Extensive literature exists on various mentorship, training, and evaluative programs [e.g. 1, 2, 20, among many others].

There has been less research to understand the diverse experiences of GTAs (but see, for example, [3] for study on GTA motivation), whose performance in the classroom is both part of their academic training and tied to undergraduate learning. Additionally, teaching science has been linked to better research performance [4]. By better understanding GTA experiences – potential causes of stress, changes in confidence levels, the impact of faculty support, and enthusiasm for teaching, among others – universities can provide more tailored programs to improve teaching and thereby improve undergraduate student retention [5-10]. This study had two goals: to gain a qualitative and quantitative understanding of the first semester teaching assistant experience, and to explore the impact of small group sessions on feelings of confidence and support for new GTAs.

The small group sessions outlined in this study provided a setting to discuss difficult experiences, practice and hone lessons, share effective classroom practices, and connect with other students interested in the act of teaching. Similar programs of peer support are shown to positively affect teacher development [1]. This pilot study proves instructive in understanding and improving the first semester teaching experience, which in turn may impact undergraduate student learning.

Study Background

All new College of Engineering GTAs at Cornell University are required to complete one-and-a-half days of training, which is administered by the Office of Engineering Learning Initiatives. Eight graduate students with prior teaching experience are hired to teach learning styles, interactive teaching techniques, grading best practices, and classroom management [14]. In the 2016-2017 academic year, authors Case and Kearney were two of these eight graduate students.

The training additionally includes sessions on diversity, inclusion and the Family Education Rights and Privacy Act. To complete training, all new students are required to go through the "Microteaching" process. They prepare a five-minute lesson, which is presented, videotaped, and critiqued in a group of four to five new students and one of the teaching specialists. In Case and Kearney's experience, these Microteaching sessions provide a basis for small, connected communities of teaching assistants.

Case and Kearney's experience was part of the motivation for the study, and should not be excluded from discussion. Both had difficult experiences teaching as graduate students for the first time, attributed mainly to two factors: lack of faculty support, and unfamiliarity with the subject material. This informed the study design.

Methodology

All study participants were graduate students in masters of engineering, masters of science, and doctoral programs in Cornell University's College of Engineering who were teaching in Spring 2017 for their first semester at Cornell. Participants may have taught elsewhere; see the "Participant Experience" section for detailed background. They were recruited from the teaching assistant training programing detailed above. The study was voluntary and self-selective (IRB Protocol ID #1701006895).

The study was split into two main groups: the "survey group" (N=20) was sent five surveys online through Qualtrics (Qualtrics, Provo, Utah) throughout the semester and otherwise had no contact with the authors. The "supported group" (N=7) was sent the same surveys, and each participant was interviewed at the beginning and end of the semester. These participants met with each other and Case and Kearney four times, about once a month.

Three of the surveys measured a variety of aspects related to teaching: personal feelings of the participant about teaching, likeliness of a participant's career involving teaching, prior experience teaching,

Table 1: Descriptions of survey instruments utilized in this study.

Survey #	Recipients & # times distributed	Information collected from survey
#1 *throughout the semester*	Survey – 1 Supported – 1	Stress, excitement, and confidence about teaching; time spent on teaching-related duties; successes and challenges during teaching
#2 *pre-semester*	Survey – 3 Supported – 4	Interest in teaching; likeliness of teaching as a future career; prior teaching experience; anxiety; knowledge of the course content; anticipated difficulties and successes
#3 *post-semester*	Survey – 1 Supported – 1	Similar to survey #2, but administered at the end of the semester
#4 *post-semester*	Survey – 0 Supported – 1	Effectiveness of "supported group" meetings

and prior knowledge of the subject. The surveys also collected qualitative data about difficulties and successes each month, which was used to frame the breadth of experiences of first semester teaching assistants. The fourth survey was given only to the "supported group", and measured the usefulness of the monthly meetings. All quantitative questions were measured on a Likert-like scale, with 1 indicating low and 7 high levels of excitement, stress, confidence, or other variables. See Table 1 for an outline of the surveys, and Appendix 1 for the complete surveys.

In the monthly group meetings, Case and Kearney led a series of activities that were tied into the teaching experience. See Table 2 for the activity and related teaching aspect.

Qualitative Analysis

Audio from individual interviews with "supported group" participants was recorded and transcribed and the small group sessions were recorded and notes were taken. Directly after the third and fourth group sessions, Case and Kearney made written records of the type and quality of interactions between participants, along with their content. The audio from group sessions were transcribed fully or in part excepting casual banter.

From these transcripts, Case and Kearney primarily used open coding individually to identify common themes among the interviews and meeting conversations. The three themes that emerged from Case and Kearney's coding are discussed in the results and discussion section. A priori coding was also used insofar as questions asked on each survey were related to the phenomena investigated in the study. These included internal and external stressors, confidence factors, and excitement. Coding was based off of methods discussed in [15] and [16].

Qualitative data from the surveys was categorized as indicating faculty or peer support, and were also grouped under the three common themes that emerged from the "supported group" interviews and meetings.

While almost all of the qualitative data was collected through formal interviews, small group meetings, and surveys, a minority was collected through informal conversations with participants outside of the study's formal scope but within permission granted by the IRB. These informal conversations provided a way for participants to discuss their teaching experiences without the pressure of a group setting or audio recording.

This information was used to estimate the quality of support GTAs received and along with survey answers, provided a baseline for the types of interactions, difficulties, and successes commonly faced by GTAs. Three themes emerged from the interviews and group sessions and qualitative data on the surveys, which are discussed in the results section.

Table 2: Description of activities utilized to highlight aspects of the teaching experience of graduate students during monthly supported group meetings.

Activity	Description	Relation to teaching
Two Truths and A Lie	Participants share two true personal details and one false personal detail with the group	How much personal information do you share with students?
The Zebra Puzzle	Participants have to match sets of cards based on clues, but each person only gets a subset of clues and can only share them verbally	Group dynamics: who dominated the decision making? What information might you be missing when talking with students?
Stranded at Sea	Participants rank the most important items from a list to take with them on a lifeboat; they repeat the activity as a group, and compare it to expert rankings	Group dynamics: was the decision making a group process? Were the scores closer to the expert ranking as a group or as an individual?

Quantitative Methods

All quantitative data was analysed in Python using the matplotlib, NumPy, Seaborn, pandas, and SciPy packages [17-19]; the code and Jupyter Notebooks (Continuum Analytics, Inc.) are available by request from the first author.

Filled boxplots show the interquartile range (IQR); the whiskers extend to show the rest of the distribution with the exception of outliers, which were all data points beyond 1.5 times the interquartile range (H-spread), the Turkey convention.

The statistical significance of the linear regressions was determined by computing the Wald statistic followed by a t-test on the sample size minus two, as per the SciPy statistical package for linear regression. All p-values are two-sided.

Results

Case Studies of Each "Supported Group" Participant

Much can be learned about the diversity of GTA experience through the personal stories of each "supported group" participant. To preserve their privacy, the participants will be referred to by a number with their gender preserved and any specific details generalized to prevent identification.

Participant 1

Participant 1, a master's of engineering student, had little teaching experience when he started this semester, and no formal teaching experience within a higher education institution. Along with Participant 2, he was a teaching assistant for a senior design class that required students to draw upon their entire undergraduate curriculum to successfully complete their project. His job was primarily to give feedback and to mentor students through the process, as opposed to giving a recitation or grading assignments. Like Participant 2, he was on the debate team as an undergraduate, so he felt comfortable speaking in front of others. While he had enjoyed mentoring in the past, his primary reason for choosing an assistantship was to help fund his education. His only teaching concern was encountering unfamiliar subjects since the course was so broad.

During the semester, Participants 1 and 2 both approached their work in a similar fashion: they went above and beyond their duties as teaching assistants to mentor their student groups, as exemplified their professor, a visiting faculty member whose only duties were to the course. This shows how an exemplary mentor might foster an enjoyment for teaching – these two participants never had to struggle to provide for their students; their depth of their care formed, perhaps, because of the example they were given. Participant 1's teaching experience was wholly positive. He, like Participant 2, cited the commitment and genuine interest of his professor as major factors in the ease and enjoyment of his semester as a GTA. He said he felt more confident in his teaching ability as the semester progressed and earned credibility among his students. Formal teaching will likely

not be a part of his future career.

Participant 2

Participant 2, also a master's of engineering student, came in with substantial prior teaching experience. He founded and was still running two education organizations in his home country that focused on soft-skill development like public speaking, stating that he found an aptitude for teaching as the head of the debate team at his undergraduate institution.

In the group meetings, Participant 2 often talked about how rewarding it was to work for a professor supportive of his goals and his students learning, and to guide the students through their projects. He often held office hours once a day during the week, which exceeded the scope of his teaching contract. One defining moment of Participant 2's semester was working through a grading conflict with his professor, who Participant 2 said felt guilty about assigning a poor grade after students complained. That the faculty member was open to Participant 2's suggestion exemplifies the respect shared between a successful GTA-faculty working relationship.

Overall, Participant 2 reported his teaching experience "to be bliss", specifically citing the behavior of the course professor as a major contributing factor. His main success was learning that even if he did not have an answer, he could still guide students to one.

Participant 1 and Participant 2 may have been teaching the same class, but did so for different reasons and with different levels of experience. Despite this, both had similar experiences, as reflected in surveys and interview responses.

Participant 3

Participant 3 was a PhD student with only tutoring experience. He expressed significant anxiety about teaching, especially speaking in front of a classroom, which he wanted to overcome because he was considering a career in academia.

Participant 3 was assigned to teach a newly-created graduate-level course co-administered by two professors, one of whom would give lectures remotely. Reviewing the syllabus, Participant 3 was excited for the class content because it included some of his research interests. He anticipated that his duties would be to conduct review sessions, hold office hours, and correct homework.

Early in the semester, Participant 3 expressed concerns that the pace of lectures was slow-going. After about two weeks, the course was severely behind schedule: all planned exams were cancelled, homework assignments were rarely evaluated or even assigned, and lectures were weeks behind. Participant

3 reported being frustrated at how "off the rails" the class had become despite his efforts to touch base with the professors to coordinate homework solutions, exam expectations, and other course details. While this chaos meant a reduced workload for Participant 3, he was frustrated because he lost out on the opportunity to improve his teaching.

Participant 3 took what he could from this negative teaching experience. He reported that the most valuable lessons he learned were about what not to do and the extent of coordination required to plan a new course. Participant 3 expressed that the experience had not deterred him from seeking a career as a professor, but rather inspired him to be a better one.

Participant 4

Participant 4 was a doctoral student with little formal teaching experience aside from tutoring and mentoring other students. He was looking forward to teaching and developing his ability to explain a problem or concept in several different ways. With a background in high school theatre and experience presenting at conferences, he had confidence in his ability to speak in public or in front of a classroom. Participant 4 was one of a large team of GTAs for a laboratory class for third-year students. He felt confident in his knowledge of the mathematics necessary for the class, but less so about the particular applications in the lab.

He reported feeling very supported by fellow GTAs and his professor, especially for her organization of the course; the instructor would distribute past years' rubrics to doctoral GTAs who would then go over and make corrections to them before distributing them to the master's GTAs who handled a bulk of the grading. In addition, the instructor divided up responsibilities so that some GTAs had more lab hours while others had more office hours or grading responsibilities. This clear division of responsibilities, Participant 4 reported, was "perfectly equitable" and went a long way to reduce his stress levels. He also mentioned that he wished that he could receive some critical feedback from his students so that he could improve his own skills, but that the limited feedback he received from his students' mid-semester evaluations was positive.

Participant 4 was satisfied with his experience as a TA, reporting that he became better at answering questions that he didn't immediately know the answer to, and that he learned how to help students calm down when they were frustrated. The support from his teaching team continued consistently throughout the semester.

Participant 5

Participant 5 was a master's of science student who specifically sought out a teaching assistant position; it was not required of a student in her program to do this. She had no formal teaching experience, but Participant 5 had tutored at various times and worked as a resident assistant during her undergraduate studies. She reported that she was comfortable giving presentations in front of people, having worked in industry before graduate school, where she had to prepare formal presentations regularly. She mentioned that she was excited to try to make a difference to her students.

Participant 5 taught the same class that Participant 4 did, a junior-level laboratory class. Though she had studied the course material during her undergraduate career, she reported feeling somewhat nervous about using the lab equipment for the class. Like Participant 4, Participant 5 reported that she was very well supported by her professor and the rest of her teaching team throughout the semester.

Participant 5 was open in sharing her difficulties with the group to collect feedback on a variety of issues. Among these, she asked for advice on formulating rubrics and grading, as well as reacting to some critical feedback from her students.

Participant 5 felt her semester of teaching was successful. She accomplished her stated goal of becoming more familiar with and confident using the laboratory equipment and felt that she had grown in her ability to lead students to the right answer rather than giving them the information. She was also satisfied at the progress that her students made, noting that the quality of their lab reports had increased from the beginning of the semester to the end. She was highly favorably evaluated by her students, and she will be working as a teaching assistant again over the summer term.

Participant 6

Participant 6, a master's of engineering student, came to Cornell University straight from her undergraduate institution where she had three semesters of experience as an undergraduate teaching assistant. These experiences were largely based on grading homework assignments; she had little contact with students. Her major concern was unfamiliarity with the course material – her undergraduate degree was in a different engineering discipline than her master's program, and she wasn't confident she could play the role of both student and instructor in learning and teaching the material. This concern was compounded because she learned which course she would be teaching only one week before classes started.

She felt supported by her professor, who she said was extremely organized and made the duties and assignments clear for students and teaching assistants. She worked with three other teaching assistants for a laboratory-focused class of 30 undergraduates.

Participant 6 attended every subgroup meeting and participated regularly in the conversations. She had an interest in teaching, but it was unlikely her career would include teaching.

Participant 6 attributed a good semester of teaching to the support she had from a relatively small TA to student ratio (four assistants for 30 students), to the presence of a head GTA who had taught the class previously, and to the organization of her professor. She continued to be concerned about her level of mastery of the subject, and said she wished she had known sooner what she would be teaching so she could have had time to prepare for a course with unfamiliar subject material.

Participant 7

Participant 7, a first-year doctoral student, sits apart from the other supported group participants for several reasons. First, he was removed from initial teaching assignment after poor mid-semester evaluations and transferred to a different assignment. Second, he reported actions that could be construed as abuse from his professor at the outset. Third, he faced similar turmoil in his research career. And fourth, he never attended a group meeting. His case may show the fallibility of the first interview, and the value of creating a rapport with teaching assistants to better understand their full experience, rather than the one they hope, or perceive, to have.

On paper, Participant 7 was one of the best-trained teachers, having served as a lecturer at a university in Pakistan prior to enrolling in his graduate studies at Cornell. He was confident and well-spoken, describing himself "a very verbose person", but blunt about his mild disdain toward his previous students in his home country and his current situation as a mediator between "horny young students vs. paranoid old people". He spoke highly of the role of a teacher, stating "literally this is how knowledge is transferred from generation to another" and "a good teacher can replace a dozen books"; he also said he owed his previous success to good teaching. All in all, he seemed genuinely enthusiastic about having the opportunity to teach a subject he was familiar and comfortable with.

During the semester, Participant 7 attended none of the monthly meetings, but authors Case and Kearney learned about his progress through informal conversations with him. Participant 7 became frustrated

with his students' lack of initiative: he complained they only wanted easy answers and they lacked interest in learning. Regardless of the truth of this statement, he was clearly offended by this perceived lack of willingness by the students to put the work in to learn the material. He also said he felt disrespected by his professor. By mid-semester, things had clearly boiled over. He was moved to a different teaching post. Case and Kearney had never heard of a student being removed from a teaching post, but clearly Participant 7 was not well-received as a teacher by his students. At the same time, his research adviser asked him to find a different adviser.

Themes of Challenges and Experiences of GTAs

These themes emerged from coding the surveys from both groups and transcripts from the supported group. Several themes in the diverse experiences of first time graduate teaching assistants emerged through the data. Here, three themes will be discussed: 1) The impact of prior knowledge of course content on confidence, 2) The need for faculty support of the GTA, and 3) The helpfulness of the small group meetings implemented during this study.

Prior knowledge of course content.

The most common anticipated difficulty for new graduate teaching assistants was having sufficient mastery of the course material, with 9 of 15 people who provided an answer to this question in Survey 2 indicating this was an issue. Graduate students enter their programs with diverse backgrounds, and may pursue a degree different than their undergraduate concentration. For this reason, many GTAs reported they had partial or no overlap in course content with classes they had taken in the past. As one GTA stated, "The majority of the course material is new for me, so I will have to become proficient with the material in a short amount of time." In other instances, GTAs were concerned about subject mastery because of the number of years between taking the course and teaching it. This concern manifested differently for GTAs: some were concerned that they would be asked a question that they did not know how to answer, while others worried it would inhibit their ability to provide lucid explanations if they did not deeply understand the concepts.

This concern lessened only slightly over the course of the semester. On the post-semester survey 6 of 12 respondents still indicated that this was a major concern, but some were more focused on being confident of the material rather than knowing it. Self-reported successes often conveyed pride at explaining course material they felt they understood well and conveyed effectively to their students. Keeping up with the pace of the material was also a commonly reported difficulty. Many GTAs learned to deal with this more effectively by seeking support from their teaching team, especially other teaching assistants. In particular, multiple teaching assistants mentioned using Slack, a digital messaging service, to communicate with other TAs in their course. This trend was consistent between the small group and the survey group.

When asked about the greatest teaching challenge during their semester, fewer GTAs reported concerns about knowing the course content, but it was still the most common difficulty reported. Some GTAs mentioned that having confidence in their knowledge was challenging but helpful for teaching their students.

Faculty support.

Though the extent and quality of faculty support was not the primary focus when designing this study, it became clear in conversations with small group participants that faculty influence significantly impacts the GTA experience and may be an interesting subject for further study. The discussion here is based primarily on qualitative data gathered during small group meetings and one-on-one interviews.

Faculty attitude and management can positively or negatively influence the working experience for a teaching assistant. Participants 1, 2, 4, and 5 spoke throughout the semester about how helpful, respectful and organized their professors were. These GTAs rarely reported difficulties in their teaching and were more positive about their teaching experience overall.

Conversely, Participants 3 and 7 had unsupportive faculty members in charge of their classes.

At first, Participant 3 reported that he felt supported by his professor despite the faculty member not being on campus: there was a well-organized syllabus for the new course and the professor was responsive to emails. This changed early in the semester when Participant 3 attempted to contact the faculty member in order to discuss solutions for the first homework assignment. The professor was unresponsive, and larger problems soon followed. The pace of lectures fell behind, homework that was on the syllabus was never assigned, and nearly 50% of the students dropped the class. Participant 3 expressed constant dissatisfaction with the state of the class. He was concerned that the students would have no fair way to be evaluated. He was also disappointed that he had not been able to improve himself as a teacher. When asked about his GTA experience at the end of the semester, Participant 3 replied that it was much different than he expected and that the
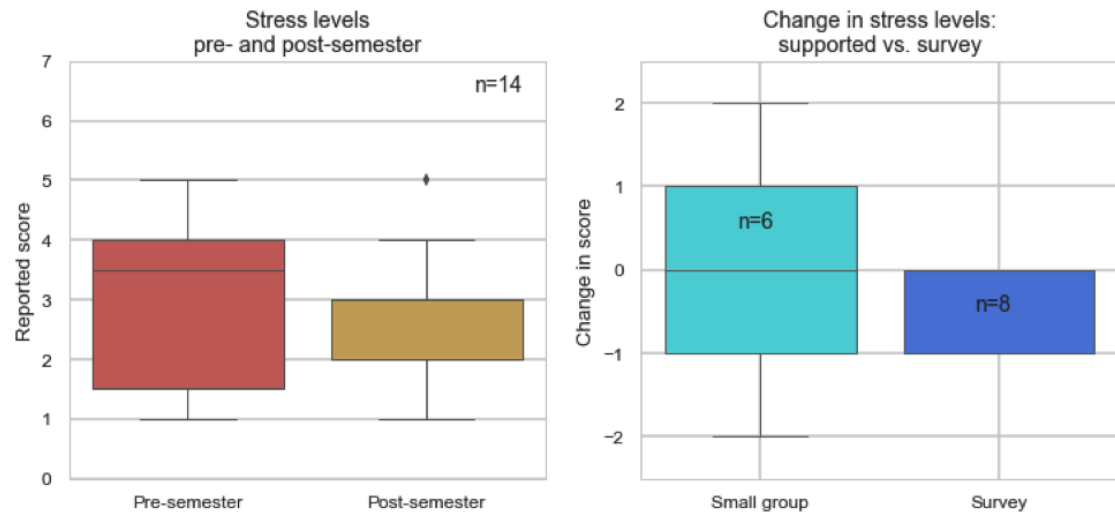
Figure 1: Stress over time and between groups as indicated by pre- and post-semester surveys (Surveys 2 and 3).
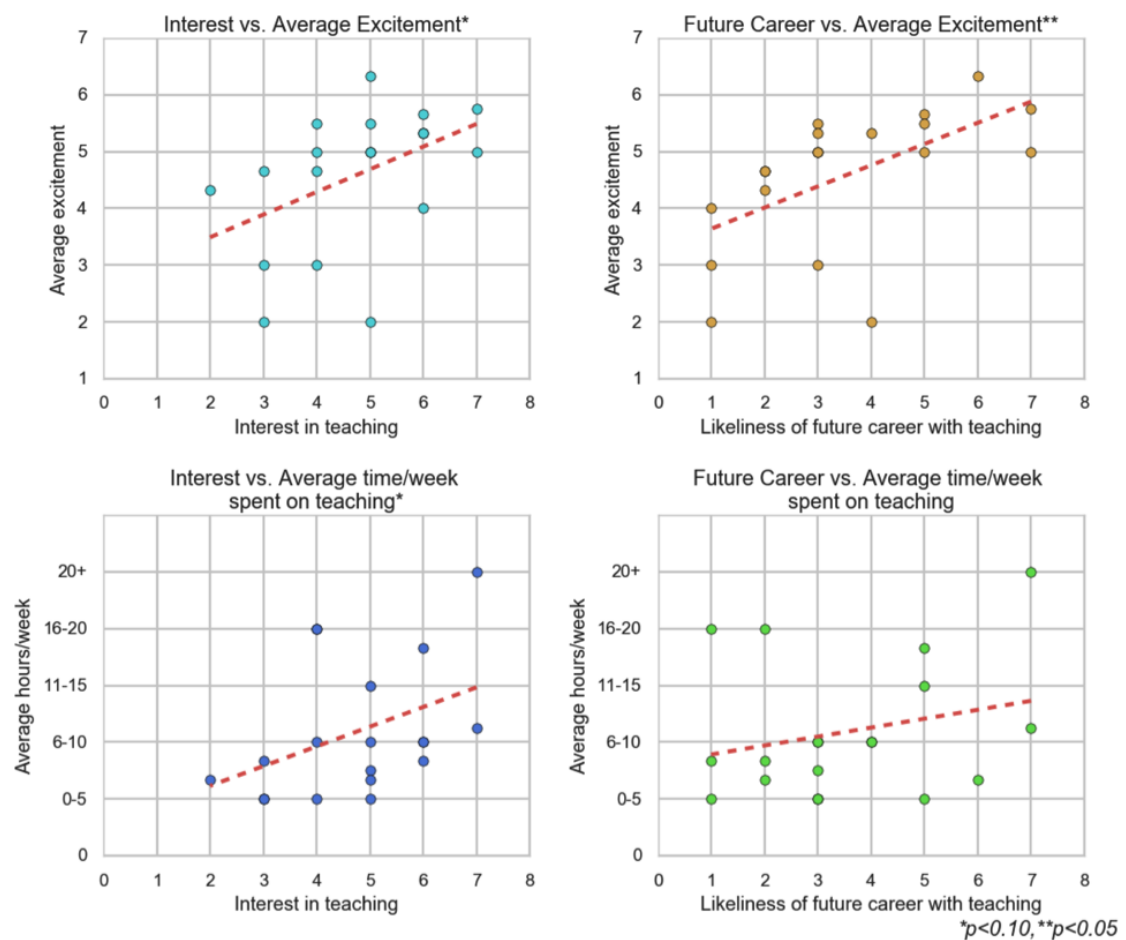

Figure 2: Regression trends between excitement for and time spent on teaching versus interest in teaching and likeliness of teaching as part of a future career.

disorganization overshadowed everything else.

Participant 7 already felt disrespected by his professor by the first interview. The participant said his professor had threatened to fire him and his fellow GTAs after they experienced technical issues with the "iClicker" in class. He felt his professor expected immediate answers to all emails and micromanaged in a way that was aggressive and rude. Perhaps if Participant 7 had felt more supported by his professor, he would have gone to him with his concerns about his students.

Participant 6's experience suggests there may be value in having a designated mediator between new GTAs and faculty members. While she reported positive interactions with her professor, she mentioned one of the most impactful aspects of her experience was having a head GTA who had already taught the class, and so knew how to run the laboratory equipment, the potential points of difficulty for students, and the material for the course.

Helpfulness of small group meetings.

The GTAs who attended the small group meetings were given the opportunity to discuss their challenges, successes, and failures with their peers, in addition to participating in interactive activities. During these meetings, discussion would often depart from the planned topics when the GTAs brought up topics on which they wanted advice. This was encouraged by Case and Kearney.

Those GTAs with positive teaching experiences said that the group meetings were not explicitly helpful, but were enjoyable. GTAs who had more difficulty in their first teaching experiences reported that the small group meetings were both enjoyable and helpful. Participant 3, who had a negative teaching experience, thought that the sessions were useful; in particular he pointed out the benefit of being able to discuss common problems.

Participant 5, who had been much better supported by her teaching team, reported that she still found value in the meetings because of the freedom to talk about personal subjects like how she felt receiving negative feedback from her students. Both Participant 5 and Participant 6 said they found value in the activities. Participant 5 provided the caveat that while she thought the lessons offered by the activities were valuable, she did not incorporate those lessons into her teaching. She suggested a direct challenge to the participants to apply a particular lesson to improve their teaching might have been more effective.

Quantitative Results

Each participant was asked to fill out a survey about once a month about their feelings toward teaching and recent experiences, as outlined in the methodology section. Overall, the spread in responses about confidence, stress, and excitement decreased throughout the semester, but the means stayed the same for both the "supported group" and "survey group", suggesting that group meetings had negligible impact on overall teaching experience for the number of participants considered in this study.

Figure 1 shows stress levels as measured by Surveys 2 and 3 as well as the change in stress levels from the beginning to the end of the semester between the "supported group" and the "survey group". The variability in stress decreased throughout the semester for both groups, but average stress did not change significantly throughout the semester for either group. Likely, the change in stress levels in the "supported group" was due to the wide variety of experiences noted for each participant. Some anticipated teaching to be less stressful than it ultimately was, others more, and this may have been tied with faculty support or the timing of the surveys.

Figure 2 shows scatterplots and linear regressions of the data collected in Survey 1 throughout the semester. Students were significantly more likely to be excited about teaching if they either had a stated interest for it ($p < 0.10$), or if teaching was likely to play a role in their future careers ($p < 0.05$). Students were more likely to spend more hours per week on teaching-related work if they stated an interest for it ($p < 0.10$).

Figure 3 shows box plots of the results of Survey 1 over the course of the semester. The results of the surveys given to both groups throughout the semester showed changes in spread, determined by the change in the standard deviation, but not means, determined by paired student t-testing, throughout the semester. Time spent on teaching decreased except for select few outliers. The spread of confidence decreased, but the mean did not change significantly, perhaps indicating that student settled into their roles but still faced some challenges during the semester. This might be expected, as teaching is not a skill learned in a matter of months. Fewer GTAs were excessively stressed by teaching by the end of the semester; perhaps expectations solidified and became routine as teaching roles were clarified and no longer new.

Figure 4 shows confidence in teaching versus prior knowledge as scatterplots. Graduate teaching assistants' prior knowledge was not correlated with their overall confidence in presenting information, as shown in the pre-semester survey data depicted
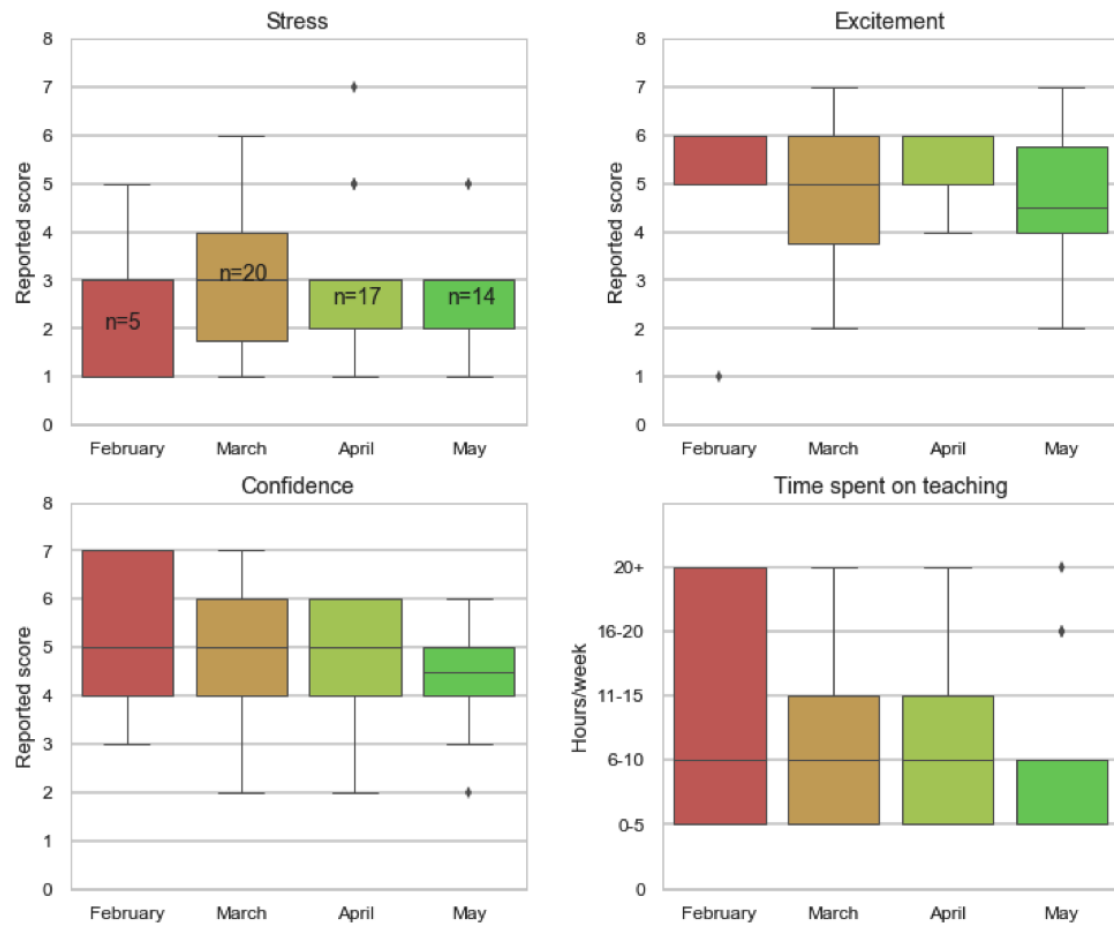
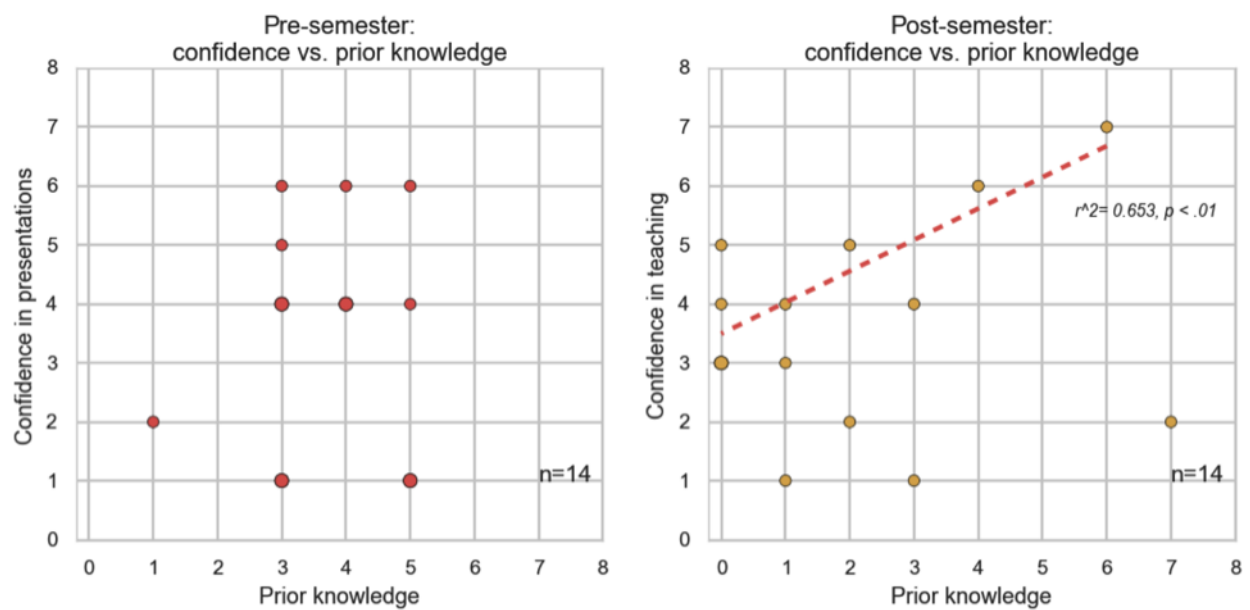Figure 3: Survey 1 results throughout the semester.

Figure 4: Confidence in teaching.

in Figure 2. Post-semester, two diverging trends emerge: for those with some confidence in their teaching abilities, i.e. they rated their confidence as 3 or higher, prior knowledge was strongly correlated with their confidence in the classroom ($p<0.01$). For those with very little confidence in their teaching, most of whom had middling to no prior knowledge, prior knowledge had no real effect on their confidence. The GTA training program might identify these GTAs to find out the specific barriers between their experiences and their confidence in the classroom.

Discussion

While the group meetings did not have a significant effect on the confidence, stress, or excitement levels of the participants, the experience of Participant 3 suggests that for those having a difficult teaching experience meetings provide a way to see what a positive teaching experience might look like. Additionally, many training programs have been implemented that effectively provide a small community through weekly seminars that teach graduate students pedagogical techniques (e.g. the University of South Carolina [11] and the University of Maryland, College Park) [12]. Participants in the "supported group" who already had teaching experience suggested in their post-semester interviews that more teaching-focused activities, like case study analysis, be incorporated. While that was not the intention of this study, it shows that longer pedagogical training programs are of interest to a subset of GTAs.

The findings of this study, especially when comparing good first teaching experiences to poor ones, point to a number of suggestions for ensuring the former. While much has been written about how to train graduate students to be better teachers or what graduate students might do to improve their teaching, faculty members and departments can take concrete steps to better ensure positive teaching experiences.

Faculty members might take note that organization, not personality, relatability, or other factors, was the most common trait discussed by the "supported group" participants. When GTAs understood how work was divided, what their duties were, and the dates and content of homework and exams, they could plan their teaching, grading, or office hours around those factors. Mentoring handbooks, like a short, three-page example from the University of Delaware, [13] also describe organization as a major factor.

Some tangible activities that were mentioned by participants as particularly helpful included hiring one experienced GTA who had already taught in the course, unambiguous grading rubrics, multiple tiers of teaching assistants (Ph.D., master's, and undergraduates) with varying duties and a clear workflow, explicit contracts outlining duties, and weekly or monthly meetings with TAs and the faculty member(s).

Departments might consider informing graduate students of their teaching duties at least one month ahead of time, so that they might have a chance to learn unfamiliar course material. They should also take into account a student's undergraduate and graduate coursework or research focus when assigning teaching duties. It is obviously better for everyone if the teacher already understands the material he or she is teaching.

Departments or colleges might consider adding course-specific information to teaching contracts so that GTAs better understand their duties. These would be prepared in collaboration with faculty members, and could be a requirement for faculty members to hire teaching assistants. See [13] for example contracts. Departments or colleges might consider incentivizing these best management practices, and also track graduate student evaluation of their teaching experience and faculty member mentorship to ensure that new teaching assistants are assigned to faculty members who have positively impacted their previous mentees.

The authors would like to acknowledge that numerous variables confound the results of this study. Some departments may have been overrepresented in our sample, some may support their students more than others, extent of faculty support was not captured nor were student evaluations of their teaching assistants to tie these metrics with performance. Surveys did not assess language proficiency, the lead time to notification of teaching duty, or if GTAs had any input on or choice with regards to their appointments. These may all be important and should be investigated in future work.

Acknowledgments

References

1. Boyle, P., & Boice, B. (1998). Systematic Mentoring for New Faculty Teachers

and Graduate Teaching Assistants. Innovative Higher Education, 22(3), 157–179. http://doi.org/10.1023/A:1025183225886

2. Park, C. (2004). The graduate teaching assistant (GTA): lessons from North American experience. Teaching in Higher Education, 9(3), 349–361. http://doi.org/10.1080/1356251042000216660

3. Kajfez, R.L., & Matusovich, H.M. (2017). Competence, Autonomy, and Relatedness as Motivators of Graduate Teaching Assistants. Journal of Engineering Education, 106(2), 245–272. http://doi.org/10.1002/jee.20167

4. Feldon, D.F., Peugh, J., Timmerman, B.E., Maher, M.A., Hurst, M., Strickland, D., Stiegelmeyer, C. (2011). Graduate Students' Teaching Experiences Improve Their Methodological Research Skills. Science, 333(6045), 1037–1039. http://doi.org/10.1126/science.1204109

5. Nora, A. (2016). Determinants of Retention among Chicano College Students: A Structural Model http://www.jstor.org/stable/40195773

6. Lamport, M. (1993). Student-faculty informal interaction and the effect on college student outcomes: A review of the literature. Adolescence, 28(112), 971–90.

7. Campbell, T.A., & Campbell, D.E. (1997). Faculty/student mentor program: Effects on academic performance and retention. Research in Higher Education, 38(6), 727–742. http://doi.org/10.1023/A:1024911904627

8. Schulte, L.E., & Thompson, F. (n.d.). Undergraduate Faculty and Student Perceptions of the Ethical Climate and its Importance in Retention. College Student Journal, 1–15.

9. Fike, D.S., & Fike, R. (2008). Predictors of First-Year Student Retention in the Community College.

10. Schreiner, L.A., Noel, P., Anderson, E.C., & Cantwell, L. (2011). The Impact of Faculty and Staff on High-Risk College Student Persistence. Journal of College Student Development, 52(3), 321–338. http://doi.org/10.1353/csd.2011.0044

11. University of South Carolina TA Training Pamphlet. (2016). Retrieved from https://sc.edu/about/offices_and_divisions/cte/graduate_teaching_assistants/docs/ta_training_announcement_sp17_v2.pdf

12. Marbach-Ad, G., Schaefer, K.L., Kumi, B.C., Friedman, L.A., Thompson, K.V., & Doyle, M.P. (2012). Development and evaluation of a prep course for chemistry graduate teaching assistants at a research university. Journal of Chemical Education, 89(7), 865–872. http://doi.org/10.1021/ed200563b

13. University of Delaware Faculty Mentoring Guide. (n.d.). Retrieved from http://ctal.udel.edu/files/2014/11/factaguide0812-1ft071e.pdf

14. College of Engineering. "Training for New TAs." Cornell University. Accessed July 2017. https://www.engineering.cornell.edu/academics/teaching/ta/training.cfm

15. Williams, K. (2015). "Doing Research to Improve Teaching and Learning." Routledge, Taylor & Francis Group.

16. Curry, L. (2015). "Fundamentals of Qualitative Research Analysis." Yale University. https://www.youtube.com/watch?v=wbdN_sLWl88&list=PLqHnHG5X2PXCsCMyN3_EzugAF7GKN2poQ

17. Hunter, J.D. (2007). "Matplotlib: a 2D graphics environment." Computing in Science & Engineering, 9(3), 90-95. IEEE Computer Science. https://doi.org/10.5281/zenodo.573577

18. Jones, E., Oliphant, T., Peterson, Pearu, et. al. (2001). "SciPy: Open source scientific tools for Python." www.scipy.org

19. Waskom, M, et. al. (2017)."Seaborn." https://zenodo.org/record/824567

20. Gaia, A.C., Corts, D.P., Tatum, H.E., & Allen, J. "The GTA Mentoring Program: An Interdisciplinary Approach to Developing Future Faculty as Teacher-Scholars." College Teaching, 51(2), 61-65.

STEM Faculty Reflect on Career Influences External to the Classroom and Their Impact on Instruction

Mary L. Centrella

Department of Entomology, Cornell University
mlc344@cornell.edu

Abstract

This study seeks to understand the personal and professional journey two faculty instructors in STEM fields undergo and the ways these factors influence and impact their teaching experience. We are specifically interested in contextual factors, outside of the classroom, that impact faculty teaching. We begin to address our question through interviews, syllabus evaluations, and course observations of two faculty members representing two very distinct situations. One is an adjunct at a liberal arts college and the other is a tenured professor at a research-one institution. Our analysis of the interviews, observations, and syllabi suggest that both of our focal faculty members experience stress and pressure from both personal and institutional external factors that affect their teaching experience. Our study suggests that these pressures external to teaching can influence faculty teaching. Additionally, we found that personal and institutional external factors worked in concert to impact faculty teaching experience. Future work on teaching should take into account the external and contextual pressures that faculty face outside of the classroom. Our results indicate that the institution can have a large impact on its faculty in both positive and negative ways. Efforts in pedagogy should include not only what is going on in the classroom but also the influences faculty are facing outside the classroom which we show here have an impact on faculty experience.

Introduction

I (MC) am currently a STEM graduate student interested in pursuing a career as a STEM instructor in higher education. Surrounded by STEM faculty in my current position, I became curious about the various factors, both detrimental and beneficial, that affect their ability to teach. My co-author and I were interested in the journey STEM faculty at different institutions have taken to become instructors. We wondered about the ways their personal interests, role or job title, and institutional type may have affected their experience as teachers. We not only wanted to ask faculty about their personal perceptions regarding their teaching experience through interviews, but also wanted to assess their teaching practice itself, as observed in the classroom. Though education studies often focus on the faculty experience inside classroom, there is less known about how factors external to the classroom, both personal and institutional, affect instructors. Our research has helped us better understand what influence external factors may exert on faculty teaching science, technology, engineering, and mathematics (STEM). This study sheds light on the journeys STEM faculty take to become instructors and on what inspires and what deters them from excelling in the classroom. We explore the experiences of faculty in both research one and liberal arts institutions and faculty with more tenuous adjunct positions as well as the experiences of faculty with tenure. This study is particularly relevant now, as the career pathway to full professor is not as clear-cut as it once was. The insights we gain from this study may help inform institutions and faculty about what we can change to improve STEM faculty experiences and, ultimately, teaching quality.

Literature Review

STEM Instruction and Student Attitudes

We know that teachers can have direct and lasting impacts on their students. Good teaching positively impacts STEM student perceptions of the subject, as evidenced by a study of introductory calculus courses across 123 institutions (Sonnertet al. 2014). In this study it was shown that good teaching practices, such as clear explanations and listening carefully to student comments and questions, had a positive effect on student attitudes towards the subject material itself, in this case mathematics. In STEM, shifting attitudes towards the subject material can have much more serious effects than in other fields. Fewer STEM applicants are produced than are demanded. Thus, the stability of the STEM fields themselves can be greatly improved by positively influencing student perception of the subject.

Although research is being done on teaching inside of the classroom, we often ignore the full context in which faculty teach. Teachers spend most of
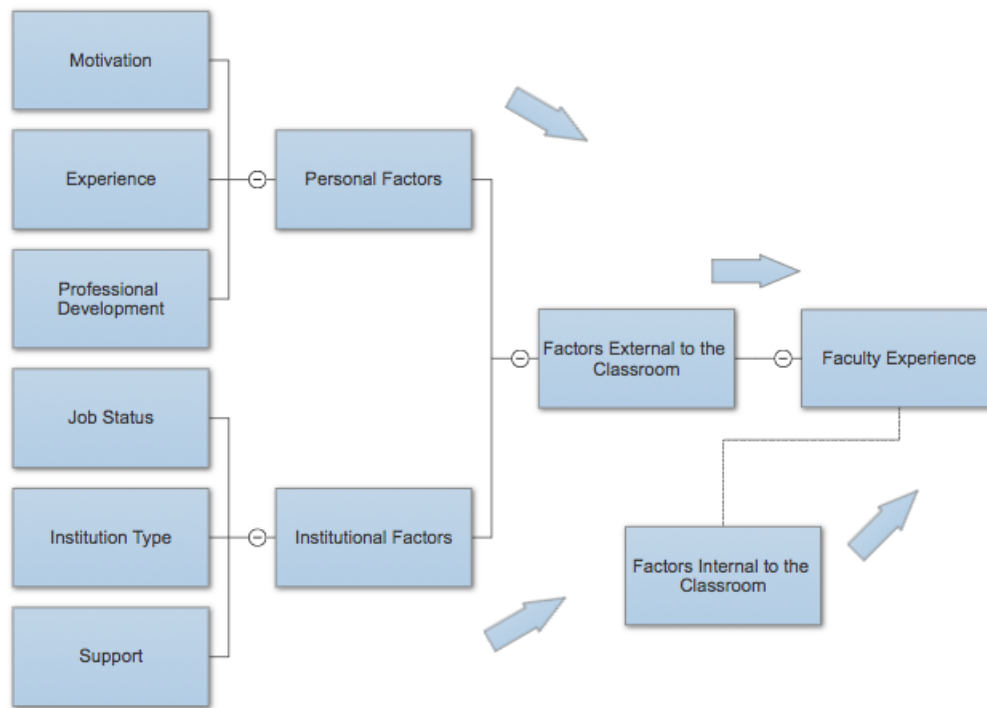
Figure 1: Schematic of the factors external to the classroom that could influence faculty instruction experience and our hypothesized interactions between these factors. Relationships are marked with lines and influences with arrows.

their day outside of the classroom, paid by jobs with demands and responsibilities that are not associated with the class in question. In order to understand the experience of STEM higher education teachers we cannot ignore factors external to the classroom that influence their teaching. In fact, in a subject where faculty teaching can make or break student interest it is imperative to understand factors outside of the classroom that may influence a faculty member's ability to teach well.

Factors that can Contribute to Faculty Instruction

The ability of an individual to teach effectively is dependent on the context of their personal life and structures of the institution in which they teach. Here we highlight six contextual factors that influence faculty experiences and, likely, teaching quality. Personal factors that may influence teaching include motivation for teaching, experience in teaching, and professional development in teaching. Institutional factors include institutional support, job status, and institution type. As the literature review below will demonstrate, these factors have been shown to impact faculty in the past, and also emerged in our study (Figure 1).

Personal factors.

1. Motivation: Motivation is essential to keep faculty teaching and to help them teach well. The types of motivations instructors have may be important to their success as instructors. To understand why instructors leave education, Hellsten and Prytula (2011) interviewed 279 entry-level teachers in Saskatchewan about their motivations. They found that loss of teachers was controlled in part by instructor motivations changing over time. Motivations upon entry to a position were often more idealistic than the motivations teachers developed once established. Entry motivations included: wanting to make a difference in people's lives, working with students (in this case children or youth), and the opportunity to teach subjects of interest. As time passed, teachers became more motivated by aspects of the position itself, such as: salary and benefits, independence (having one's own classroom in this case), and high professional quality of life. Additionally, it seems intuitive that highly motivated faculty members would be "better" teachers. A report from Oregon State University corroborates this idea. A program instituted there in the 1980s offered instructional guidance to faculty, along with a monetary award. The motivations of this program resulted in more efficient and effective instruction by faculty (Osterman, 1984).

While motivation can be influenced by many factors, it is clearly an important, personal element driving instructors and their ability to share knowledge with their students.

2. Experience: Experience as an instructor can improve teaching quality. Experienced adjunct faculty have higher motivation scores and enjoy teaching more than their younger counterparts. Additionally, adjunct faculty with more teaching experience are closer to their colleagues, more satisfied with their compensation, and more supportive of their institution's core values (Dickens, 2011). We presume that increased motivation and improved connections to the institution would likely impact the students being taught. In a long-term study of student satisfaction at a state school, faculty longevity was the main element driving student satisfaction. Students preferred more experienced instructors over better trained, but less experienced instructors. Having experienced faculty was more important to students than was the student's particular degree program (Kaneet al. 2015). Clearly, experience with teaching is important to both instructors and their students.

3. Professional Development: Some universities are developing more specialized, career-focused degree programs, as college students are becoming more diverse and selective, and the mediums and technology available to disseminate information are increasing. Interestingly, STEM professors often approach teaching with wildly varying levels of experience. Many STEM faculty take up positions with a severe lack of training in pedagogy (McKee & Tew 2013), while others have a much more developed experience with teaching.

Recently, opportunities for higher-education instructor "training" have appeared in universities in the form of instructor development programs (such as workshops) in pedagogy and teaching. Teacher development impacts student perceptions about the quality of instruction they are receiving. A study conducted in two New Mexico institutions spanning 145 students revealed that Radiological Technology students' perception of the quality of teaching increased with increasing professional development participation of their faculty (Elliott & Oliver 2016). It is not just student perceptions that change with faculty development – these development programs improve teaching quality itself. An observational study of eight STEM faculty teaching Anatomy showed that after participation in a two-year development program centered on student-based learning faculty unanimously exhibit a shift from traditional, instructor-based teaching to a reformed, student-based teaching practice. All participants reported feeling more confident implementing student-based lessons than they had before the training and seven out of eight felt more confident designing those lessons on their own (Mattheis 2014). Finally, research shows that instructor development does not just improve instructional quality, it also improves learning. In a study which offered swim instructors training on teaching pupils with autism overall instructor skill climbed far above the base-line but, more importantly, their pupils' skills improved in conjunction with their training. Child compliance increased with instructor development, and students' swimming skills improved overall, suggesting that not only does instructor training improve teaching, but also these improvements positively affect student learning (Jull & Mirenda 2016).

Faculty development most likely has a big impact on STEM instructors. Development improves faculty teaching, student perceptions of the subject and their instructors' abilities, and, most importantly, it improves student learning. Faculty development could be a powerful factor in STEM field instructor experience, where many instructors have less training in pedagogy.

Institutional factors.

1. Support: Institutional and colleague support are immensely important to teaching. Elliot and Oliver (2016) surveyed STEM faculty who participated in professional development about the efficacy of these programs. The results from this study reveal three aspects of professional development programs that make them more effective:(i) assessment of the program's efficacy by implementers, (ii) increased availability of programs at the institution and participant encouragement from the institution and, (iii) overall institutional support of faculty development in teaching (Elliott & Oliver 2016). Though assessment techniques for faculty development programs can be improved and development resources can be increased, these efforts may have little bearing on faculty if the university and collegiate climate is alienating faculty from pedagogical development. A recent review of faculty development literature reported that studies rarely address the position of the university offering the development services or how the institutional context and setting impact faculty development (Amundsen & Wilson 2012).

The institution is not the only source of teacher support. Though often hard to disentangle from the institution, colleagues can offer teaching support or discouragement, which can be a powerful thing. In a study tracking faculty being trained as online in-

structors over a three-year period faculty cited the level of support they had from other faculty and their instructional designer as one of the top three factors that contributed to their development success (Kraglund-gauthier et al. 2010). In the same study faculty participant reflections on these development programs "illustrate the critical importance of support and collaboration for novice online instructors". In a study of pedagogy development in Anatomy professors across private, public, and community institutions, six of eight participating faculty members reported their collaborations with other faculty as the best part of the two-year project (Mattheis 2014). The climate created by faculty peers is a very important one, especially when teaching is undergoing a shift towards student-based learning, a dramatic change from what most faculty experienced as learners.

A recent study on faculty development stresses that changing teachers' beliefs, not just their practice, proved to be extremely important in helping them reform their teaching style to be more student-based and student-driven (Mattheis 2014). This sort of dramatic shift would be difficult to make if faculty did not experience teaching and professional development support from their university or their peers.

2. Job Status: Job status, either full or part time, adjunct or tenure-track, can be either a source of stress or support for faculty. We presume that job status could play a large role in teaching efficacy. A study conducted in Ohio state schools revealed that first-semester students were less likely to continue into their second year of college when taught by adjunct versus tenure-track professors (Bettinger & Long 2005). However, Fike and Fike conducted a study with over 1,000 mathematics students found that faculty employment status had little impact on student grades. In the late 1980's, research was conducted to examine the link between job status and instruction. The authors found that faculty that were part-time employees were just as effective instructors as their full-time colleagues (Fike & Fike 1987). However, it was also found that increased faculty education did have a positive impact on student grades; students had higher grades when they were taught by faculty who had graduate degrees . It should be noted that grades are a poor and inconsistent measure of student learning as they are highly correlated with the individual faculty assigning them. Still, better grades in similar courses can at least coarsely indicate student improvement. More recently, this link between job status and student grades was re-examined. The research was conducted on entry-level mathematics students and showed the opposite result. Authors

found that student grades were higher when taught by lower-rank faculty (Guptaet al. 2006). It appears that job status and teacher education is important to student success but the effect may vary with context. Our investigations into the factor of job status on faculty teaching may shed light on this contentious relationship.

3. Institution Type: The type of institution where instruction occurs could impact instruction quality. To some extent we may expect smaller institutions to produce more quality instruction because they are more conducive to the formation of more close-knit relationships between faculty and professional development staff-members and effective launching of programs to improve teaching quality (OECD, 2010). However, larger institutions could have more diverse opportunities and initiatives for instructor training and quality. More specifically, it appears that an institution's focus can have a large impact on instruction (OECD, 2010). In a review of learning initiatives across a variety of international institutions, business and economics schools showed highest support of teaching and learning policies (OECD, 2010). Graduate institutions had limited support for programs, but enhanced management of those programs. Postgraduate institutions had more initiatives supported at the institutional level, instead of the undergraduate level (OECD, 2010). Multi-disciplinary schools tended to be less involved in pedagogy policy, but showed initiative in developing their own programs, specific to their schools (OECD, 2010). Program type led to variable initiatives in teaching. It should follow that these differences in programs will also affect the faculty experience. We will explore here how institution type can play a role in faculty experience and teaching style.

Based on this background literature and our own experiences, the goal of this study is to better understand how both personal factors (e.g., motivation, experience, and professional development) and institutional factors (e.g., job status, institution type, and support) may influence STEM faculty teaching in higher education. An improved understanding may help us improve instruction quality for faculty and the students they teach (Figure 1).

Methods

Faculty Participants

Two faculty participants were investigated as case studies. These faculty members were chosen because they are both instructors of introductory courses in the same STEM subject for mostly non-major students, so their classes and syllabi were comparable.

Table 1: Interview questions related to personal or external factors influencing instruction asked of each participant. Additional follow-up questions were posed during the interviews that are not written above.

| Motivation | Personal External Factors | |
	Experience	Professional Development
-What drew you to teaching or being a professor? Was this always a goal for you? -What is your favorite aspect of teaching? What is your favorite class to teach? -What do you most struggle with as a teacher and how have you approached this problem so far? -If you are teaching your students well, what do you think they should be able to do that they couldn't before taking your class? -What kinds of careers and goals do you think the top students in your class should pursue? -What are your goals and aspirations as a teacher? What are some things that are going well? What are aspects of your teaching practice that you want to develop?	-How do you use student feedback? -How do you approach grading students? What do you hope to accomplish with the feedback you give and the grades you set? Do you find your feedback and grades help guide your students? What is your biggest challenge with student interactions? How much time do you spend preparing to teach and how do you prepare? How do you develop syllabi, assignments, and tests?	-How and where and when did you learn to teach? -How much of that learning can be attributed to programs, peers, or experience?

However, we also chose them because of their differences. They work in two distinct institutions: a liberal arts college and a research-one (R1) university, and they are at different positions in their institutions. One faculty member represents an early-career professor in an adjunct position and the other represents a later career, tenured instructor with a more secure position. Faculty participants were made aware of the study and volunteered to participate. Faculty members were observed and interviewed between November 2015 and May 2016. Audio recordings of the interviews were taken, and salient parts of the interviews were transcribed for analysis. Any identifying information (such as name or location) in the interviews, course observations notes, and syllabi were removed and the participations are referred to using assigned pseudonyms. Cornell University's Institutional Review Board determined the research exempt from review.

Faculty participant 1 - "Greg".

Greg is a senior, tenured faculty member at an R1 institution. He has 10 years of teaching experience. He has developed three courses at his current institution and was teaching a self-developed introductory STEM course at the time of the research.

Faculty participant 2 - "John".

John is a temporary adjunct professor at a liberal arts college. He has five years of teaching experience. At the time of the research he was teaching a laboratory section for another instructor's course as well as teaching a self-developed introductory STEM course for non-majors in the same field as Greg's course.

Interviews

Each participant was interviewed for one to three hours total, in one or two separate sessions, using semi-structured interview methods (Edwards & Holland 2013). Interviews focused on faculty members': (i) motivation to teach, (ii) support and responsibilities, (iii) journey to become an instructor, and (iv) current teaching environment. Questions targeted the main factors we hypothesized could be affecting instruction and organized under these categories (Tables 1 and 2). Faculty were questioned with the same set of questions but follow-up questions and interview trajectories varied based on responses.

Electronically recorded faculty interview responses were coded and evaluated for themes using Bryman's framework for qualitative analysis (Bryman & Burgess 1997). In addition, transcribed excerpts of the interviews were analyzed for word-content and

Table 2: Interview questions related to institutional factors influencing instruction asked of each participant. Additional follow-up questions were posed during the interviews that are not written above.

	Institutional	
Job Status	**Institution Type**	**Support**
-What is your ultimate ideal career goal and is it different from what you are doing now? If so, how? -Do you have tenure? What is the nature of your job security? Does your job security influence your teaching?	-How are you judged and evaluated as a teacher? Does that influence the way that you teach? Did you design the course or aspects of it? What is the nature of the classes you teach? Do they have discussions and labs? Is there a curriculum you must adhere to? Did you face constraints when designing the class or assignments? How big is the class and what are the demographics like? Is this typical of classes you teach?	-How many courses do you typically teach in a semester of typical academic year? What are your other responsibilities and how much time do you spend on them? Do you find you're spending more time on one area? What kinds of resources are offered to you through the department to assist with your teaching? Do you have TA's? What kinds of tasks do you assign TA's? -How emphasized was teaching and how much was teaching discussed with you during your degree program? How important and discussed is teaching among your colleagues currently?

frequency of word use using Voyant Tools software.

Course Observations

Written notes were taken by the author during class observations, but audio and video recordings were not taken. These written notes focused on: 1) Preparation: organization of material, depth of material, logic of lecture, and depth of lecture, 2) Practice: presentation of material, appearance of slides, use of media, teaching styles, and instructor style, 3) Student response: reactions to material in class, attendance, student behavior, student-student and student-instructor interactions, and 4) Interactions: how much time is spent in student-student, instructor-student interactions, or alternative instruction such as mixed media or experiential learning lessons.

Notes were analyzed using qualitative research theory and coding methods, informed by Strauss and Charmaz (Charmaz 2006; Strauss 1987). Class observations are very much subject to the judgements of the first author, who is a less experienced instructor than both participants. The notes taken from the course observations are important to this study because they often correspond with the themes from the interviews (see Discussion), but the author would like to note that the judgements made in the classrooms, especially in the "preparation" and "practice" categories, felt uncomfortable to make and are not reliable assessments of teaching quality, mostly due to a lack of experience or authority on the subject by

the author.

Syllabus Analyses

The syllabi for the courses each of the faculty created were analyzed and themes were extracted from values emphasized in the syllabus. Additionally, syllabi were evaluated for their consistency with themes from the interviews and class observation notes. Syllabi analysis was informed by the same methods used for course observations (Charmaz 2006; Strauss 1987).

Overall Analysis

Faculty perceptions and comments from the interviews were matched with the course observations and syllabi and evaluated for consistency and over-arching themes (Charmaz 2006; Strauss 1987). A new theory for how external factors can affect faculty experience was developed using grounded theory techniques (Bryman & Burgess 1997).

Results

Interviews

As a result of the interviews, a number of major themes emerged for each faculty member (Tables 3 and 4).

Participant 1 - Greg.

An excerpt from Greg's interview was analyzed. The interviewer's questions were removed so that only Greg's answers were analyzed. The excerpt was

Table 3: Themes emerging from interview with Greg.

| Question Theme	Emerging Themes	Quotes from Transcription
Motivation	-Research -Impact on students and science through writing textbooks and papers	"Give me more time to write stuff that's meaningful. There's a shallowness to this stuff that we have to write that just takes so much time and its half-life is so short."
Experience	-Class organization	
Professional Development	-Introspection and Reflection	"I don't have a whole lot of control over what they [students] end up doing" "I like what you're giving me [referring to questions about his teaching in the interview]...I don't tend to think like this about teaching "
Job Status	-Freedom to design courses	"I feel very secure in my job – that's the beauty of having tenure."
Institution Type	-Research 1 institution	"I don't think we expect our faculty to be award-winning teachers [but we do expect them to be]...reasonable teachers"
Support	-Administrative obligations and duties	"I just don't love it [administrative duties] – I'd really much rather be in the lab or in the field or teaching"
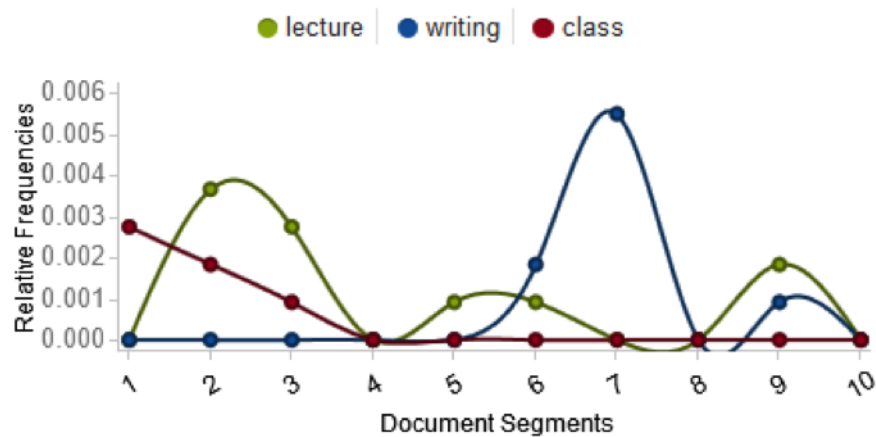
Figure 2: Relative frequencies of the 3 most commonly used words in Greg's interview excerpt as found in each segment of the interview document.

1,095 words long, with 334 unique words. The top three most commonly used words found in the transcribed interview excerpt were quantified, ignoring any commonly used vernacular. These top three thematically significant words were: lecture (11 uses), writing (9), and class (6) (Figure 2). As we will explore later, writing is an important way that Greg feels he can make an impact. Interestingly, Greg's use of the word "writing" appears apart from sections of the interviews where he uses the words "lecture" or "class".

Participant 2 - John.

An excerpt from John's interview was analyzed. The interviewer questions were removed so that only John's words were included in analysis. The excerpt was 1,153 words long, with 344 unique words. The top three most commonly found words in the transcribed interview excerpt were quantified, ignoring any commonly used vernacular. The top three thematically significant words were: pressure (14 uses), teaching (12 uses), and research (9 uses) (Figure 3). Interestingly, the word "pressure" was less frequently used by the interviewee in association with just "teaching" on its own. "Pressure" was more prevalent in the interviews in sections where the word "research" was used or where "teaching" and "research" were used together.

Course Observations

Participant 1 - Greg.

Watching Greg teach, I noticed his strength was in organization and presentation of material. He created clear expectations for what he expected students to know and remember. He summarized material well in study guides and reviews for students and revisited previous lectures using review questions. The questions he asked of students were deeper thought and learning questions, not just rote memorization but questions that made students synthesize what they had learned and apply their knowledge to other scenarios or phenomena. Greg had a phenomenal use of scientific examples and case studies to bolster his arguments in lecture. These deeper learning questions and the command of scientific examples we see in the classroom complement Greg's emphasis on research during the interviews and his self-purported interest in teaching students how to become researchers.

Greg's self-reported lack of pedagogical training might influence his presentation style. At times, he appeared nervous, said "um, and uh" often, rushed towards the end of the lecture, and used jargon. Use of jargon relates once again back to his training as a scientist at an R1 institution and his interest in

research but could potentially be off-putting to non-major students without this knowledge.

One area of the lesson that was slightly underdeveloped was the broader context or framework for the material being presented. Greg's efforts were focused on the science itself and less on the broader context of the concepts he was presenting. This struck me as connected to insight from the interview. Greg mentions that he doesn't think he has too much control over what students end up doing in the future and that he does not often dwell on his teaching trajectory. In class, Greg is teaching science in a fun and interesting way, but he teaches science for the scientifically-minded. In the observation, Greg did not often use analogies from outside of the field or explain broader impacts and thus did not spend much effort trying to appeal to non-STEM-oriented students. This way of teaching science could also be influenced by the university where Greg works. Research institutions emphasize research over teaching and Greg could simply be reflecting this focus on research in his teaching style.

On the single day of class that was observed, attendance was low – about one third of enrolled students were in class. In response, Greg announced that he would begin giving participation points for attendance. Students appeared bored at times, talking with neighbors, but their attention was also captured at times during the lecture especially when Greg was playing relevant videos or snuck jokes into the lecture. His visual presentation and use of relevant media and examples was masterful and forced students to pay attention. This relates back to interviews, where Greg says he has freedom to design his own course and has "fun" doing it.

Despite his lack of pedagogical training, Greg varied his lecture such that he spent about one fourth of the lecture using videos to show his point, and had an additional five-minute period where he informally quizzed students to review material they had already learned. Still, the class was lecture-based and there was no time, at least in the class I attended, for student-student interactions.

Participant 2 - John.

One of John's strengths in the classroom was his presentation style. He seemed at ease in the classroom context and moved around the room often. It could be that this ease in the classroom stemmed from John's pedagogical background and workshops he had taken. John also expressed a passion for teaching in the interviews that might influence his presentation style. John's use of analogies was strong and he presented the broader context and motivations

Table 4: Themes emerging from interview with John.

| Question Theme	Emerging Themes	Quotes from Transcription
Motivation	-Enthusiasm for teaching	"The one student out of 90 that really walks away from your class loving it or feeling engaged – it makes it all worthwhile and so it's my enthusiasm for teaching and sharing my interest in biology and getting people to really love biology that keeps me going."
Experience	-Student-based teaching	"In the [first institution where he taught] it really was sort of Baptism by fire…it was so overwhelming…I really almost quit."
Professional Development	-Pedagogy informs teaching through workshops and courses he has enrolled in	
Job Status	-Adjunct-related anxiety -Freedom to design course and assignments	"What it [applying for a job as a STEM faculty member] comes down to, after your qualifications, is this kind of amorphous thing that they call 'the fit'. And to get that, it's like winning the lottery. I haven't found that fit yet."
Institution Type	-Liberal arts college	"How are you supposed to publish if you just have an adjunct position and you're just supposed to just be teaching, you have no infrastructure to do research here – it's tough, right…These are the kinds of challenges that face adjuncts and that's the kind of background anxiety and stress I'm talking about that's always there." "There is great collegiality among the faculty about talking about your teaching approach and your teaching methods. Yes. Even to me as an adjunct, whatever, there's no difference there. They're very happy to discuss if you're having issues in your class, if you feel something isn't working. It's fantastic."
Support	-Incentive to invest in career	"I basically come in, teach my courses, and leave…There's no incentive for me to do anything more… In an adjunct position, loyalty and extra work is not rewarded at all." "There is almost an attitude [among fellow adjuncts] of 'well, pardon me, it's hard to really care when this place doesn't really care about me'."
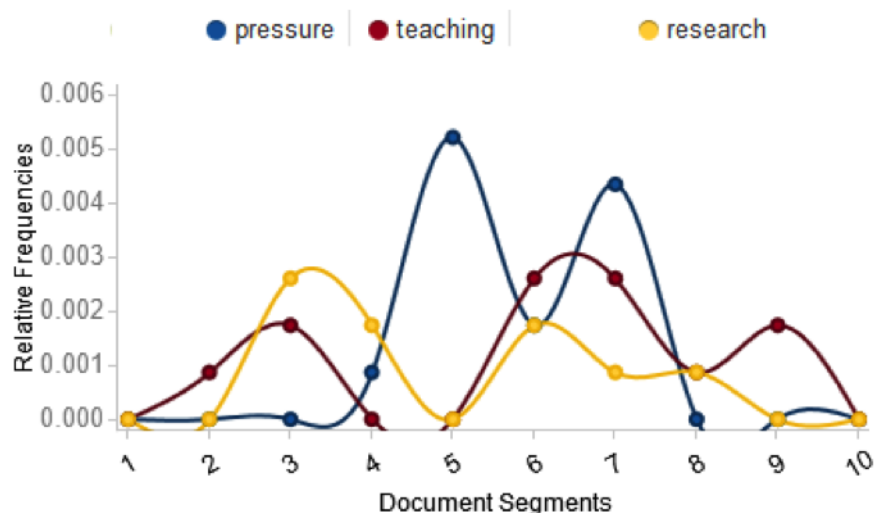
Figure 3: Relative frequencies of the 3 most commonly used words in John's interview excerpt as found in each segment of the interview document.

behind the more specific examples he showed to students in class. He summarized complex material to students using little jargon. This relates to John's interviews, where he showed a very student-centric view of teaching. Also, John was working at a liberal arts school where critical thinking may be championed over content, so this might be a reflection of the institutional factors John is a part of. Sometimes John's summary of material lacked a space for student interaction. He was telling students what scientific authors found without presenting raw data for students to interpret themselves. Perhaps because of the summary, John presented many sides to each argument and was able to show many examples this way including evidence contradictory to the concepts he was teaching. This follows what John said in interviews about wanting to help students become critical thinkers and critical consumers of science by exposing them to many different arguments and opinions.

John used jokes and colorful stories to captivate his students. He had an obvious rapport with students, as they seemed at ease asking him question and chatting with him after class. Indeed, in interviews, John said his motivations for teaching were student-centered. Attendance was high. Students walked in and out during class, chatted, or worked on non-related material on their computers. John's lecture slides were often wordy and the pictures were small, which may have made them more difficult for students to digest.

Despite his pedagogical training, John's lecture style was much like Greg's, He spent one fourth of the class on related videos and a five-minute portion of the class on review using iClicker questions. John's review questions appeared shallow, focusing on content instead of deeper learning. This does not seem to correspond with John's goal of creating critical thinkers but might be related to John's limited experience of only five years of teaching, or to his self-reported anxiety as an adjunct. There were no instructor-solicited student-student interactions during the two class sessions I observed.

Overall course observations.

The course observations were quite consistent with what was discussed in the interviews, and we got a sense of how faculty motivation influences the way they teach. Both courses were excellently taught and both instructors appeared passionate, whether about their subject material or about teaching. Despite reported differences in pedagogical training both instructors had mostly lecture-based courses and both were implementing strategies to keep material manageable for students, whether through summarization or review questions and organization of material.

Syllabus Analyses

Participant 1 - Greg.

In the syllabus analysis for Greg's course we found that half of the course was an introduction to the subject and its relation to scientific concepts, such as ecology and evolution. The other half of the course was dedicated to the subject's interaction with and impacts on human society. The basic course was mostly lectures except for a few videos and out-of-class readings from a single textbook. The discussion section (three-credit option), supported mainly by the three TA's, offered a wealth of activities including interactive and experiential learning opportunities such as museum visits and laboratory experiences as well

as assistance with development of the term paper. Along with an understanding of the course content and the impacts the subject has on humans, goals of the course included strengthening scientific reasoning and communication. Concerning assessment, the overall course grade was based on two exams that comprised half of the grade and a mixture of in-class and out of class quizzes, attendance, and writing assignments for the other half. Writing assignments involved learning the skills to properly read and comprehend a scientific article. Once students understood the article they were expected to write a popular press article that summarized the scientific content they had gleaned. Media was used via a course website which housed online quizzes. This same website also served as a sounding board where TAs would post current-event articles relating to course content. Attendance, absences, and student disability policy were briefly discussed. There was no diversity statement or mention of special circumstances besides illness that might impede student learning. Overall, this syllabus is strongly focused on the course content and the instructor's enthusiasm for the material and dedication to teaching scientific thought and principles.

Participant 2 - John.

In our analysis of John's course syllabus, we found that about half of the course is dedicated to introducing students to the course content and the second half was focused on the interactions the subject has to human culture. While most learning objectives focused on mastery of the material there were also objectives centered on interaction with human society. One learning objective centered around trying to get students to propose novel solutions to problems plaguing human society using the content covered in the course. Less than half of the course assessment was based on four exams while the rest was based on in-class electronic iClicker questions and two group projects, one where students develop their own test questions and another poster presentation on a special topic related to the course. This assignment shows that student-student interactions might occur in John's course outside of class-time. Inside of class most of the time was spent on lectures with some videos and some in-class exam reviews. Course readings were expected to be done outside of class and came from multiple sources including textbooks, popular books, and popular press articles. Multimedia was used in video, iClickers, and the course website. John was the sole instructor for the course with no TA. Attendance was emphasized as incredibly important and we saw high attendance in the observed classes. Religious exceptions, health emer-

gencies, and disability policies were discussed. There was emphasis on resources available to struggling students and the instructor appointed himself as a source of support if students felt overwhelmed. Extracurricular activities were included as an excusable absence from class. Interestingly, the instructor mentioned his responsibility as instructor, as well as the responsibilities of students. "This syllabus constitutes a contract between you (the student) and me (the instructor)" and mentioned course evaluations as a strict requirement. This syllabus emphasizes the course material's relationship to humans and the instructor's interest in his students – both their learning and well-being.

Overall Analysis

The interviews, course observations, and syllabi were analyzed for common themes about the teaching experience and the factors that affect teaching. The major themes identified during this analysis were: (i) institution type might influence personal external factors that affect teaching and (ii) institutional stress impedes instruction. These themes and their importance to teaching will be the subject of the discussion.

Discussion

Interactions between Institutional and Personal Factors

Interviews.

Though our hypothesis was originally that institutional and personal factors had independent effects on teachers' experiences (Figure 1) our analysis strongly suggests that there is an interaction between the personal and institutional drivers affecting teaching (Figure 7). For example, Greg is personally motivated by research and by shaping students to become researchers. This fact cannot be separated from Greg's institutional influences. Greg teaches at a R1 institution which has deep interests and goals surrounding research. Greg is candid about his personal professional development. He has little experience in pedagogy and professional development surrounding teaching. He feels that he doesn't "have a whole lot of control over what they [students] are doing" in the class and during the interview he said that he doesn't "tend to think like this about teaching". Greg has been teaching for over 10 years and has learned his practice mainly through experiences in the classroom. He says he outlines lectures the way he outlines grant proposals. This mirrors his institutional standing. Greg's lack of training follows the typical stereotype of a R1 professor. The institution puts more emphasis on research than teaching, there is little motivation for professors to seek professional development, and faculty are expected to

just "figure things out" as they teach. Greg says it best himself about his institution: "I don't think we expect our faculty to be award-winning teachers [but we do expect them to be] ... reasonable teachers". It should be noted that despite his lack of formal training Greg is thriving as an instructor, has developed three courses, and clearly cares about the careers of the students he trains. His training tends to focus on helping students to become researchers which is in line with both his university and his personal motivations.

John's teaching experience shows a similar link between personal and institutional factors. John is personally motivated by his enthusiasm for teaching. This personal motivation is impossible to extricate from his institutional situation: a liberal arts college where teaching is the main priority. Additionally, one of John's specific personal goals for his students is for them to become critical thinkers. This goal is also expressed by many liberal arts institutions. John is dedicated to pedagogy, has attended professional development workshops, and has learned to improve his teaching using student feedback and focusing on student growth. This model of student-based learning and emphasis on pedagogy and the practice of teaching are values emphasized by his institution type.

Course observations and syllabi.

This interaction between personal and institutional motivations and goals is also evident in the course observations and syllabi. Greg emphasizes his personal dedication to research in both his syllabus and in the observations. He uses a plethora of scientific studies to back up his arguments in class. His use of jargon is at times excessive, especially for a non-majors course. His syllabus is focused on the course content and the major class assignment is to read a scientific article and make it understandable to the public. This sort of assignment champions research. Students end up learning how to read scientific articles, but they are doing so through research. The course objectives include strengthening scientific reasoning and communication. The students enter the syllabus language as observers, but they are not the main focus of the course. The subject material drives the course. This emphasis on the research and the course content could easily stem from a R1 institution's influence.

In turn, the liberal arts institutional philosophy is strongly evident in John's syllabus and his class observations. In interviews, John says his personal goal is for his students to become critical learners and John also cites his interest in student-driven teaching. During his class John presents topics from a variety of different viewpoints. He also uses analogies to make the subject more relatable to students and tells colorful stories that captivate his audience. These actions in class are also linked to John's personal goals. In John's syllabus we see an emphasis on readings from a variety of different sources and articles, an observation in line with John's purported goal of developing critical thinkers of his students. We should note that this goal of shaping students into critical thinkers is one we would typically associate with the liberal arts college. It is possible, indeed probable, that John's location and surrounding institutional influences in part dictate the personal goals he has as a teacher. John's syllabus corroborates this idea. The syllabus emphasizes the students as active players in the classroom: "This syllabus constitutes a contract between you (the student) and me (the instructor)" and focuses on student feedback to his teaching. One typically associates student-based learning and teaching with liberal arts colleges.

Our analysis suggests we must change our framework from one where the personal and institutional factors affect teachers in isolation (Figure 4a), to one where they interact with one another to affect teaching (Figure 4b). This interaction between the institution and the personal is important to understand when we work to improve faculty teaching. We must take into account the institutions the faculty work in for effective training and pedagogy studies in the future.

The literature is consistent with our findings. Factors influencing teaching most likely do not exist in isolation but are influenced by each other. For instance, a study on faculty motivations in higher education found that job status factors such as compensations and benefits feed into personal motivation to teach. Also, factors such as job design, work environment, interaction with colleagues and independence in decision-making influence faculty motivation (Rasheed et al. 2010). Studies show that instruction quality can be improved by department-wide efforts that offer professional development to increase faculty motivation and change instruction styles (Hativa 1995; Osterman 1984). Also, in Dickens' (2011) work on adjunct instruction we see that simply being a more experienced instructor can improve perception of institutional support and can also increase motivation. In these cases institutional factors that are aspects of the job interplay in a very definite way with personal factors affecting faculty teaching. The elements leading to quality teaching are often interwoven and interdependent.

Though we can now be confident that these two factors interact with one another, our study cannot
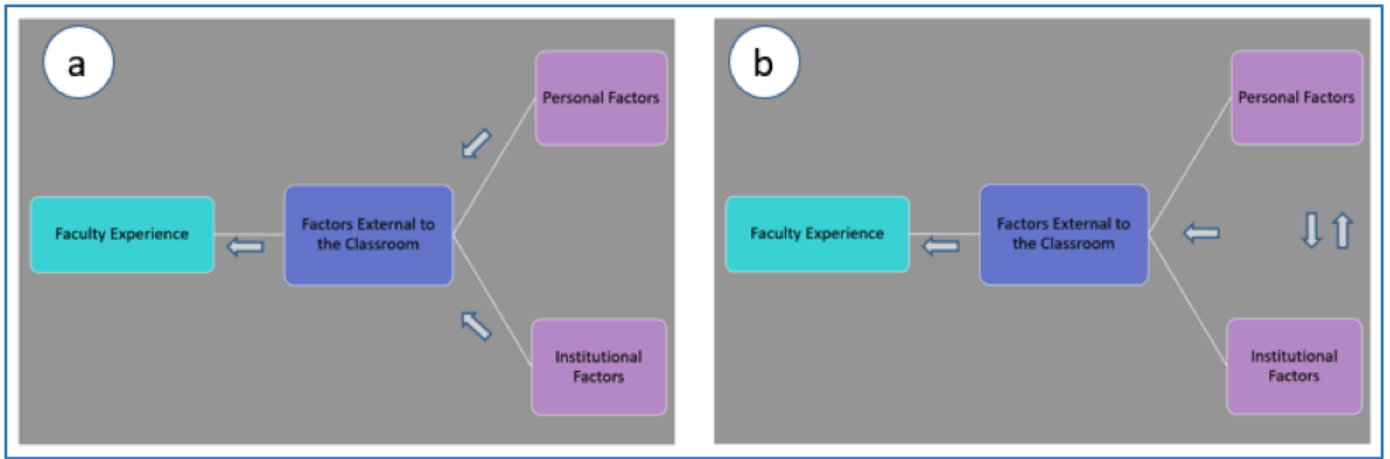
Figure 4: Schematics of (a) our original hypothesis for how factors external to the classroom affect teaching in isolation, and (b) the new way that both factors interact to affect teaching, informed by our results.
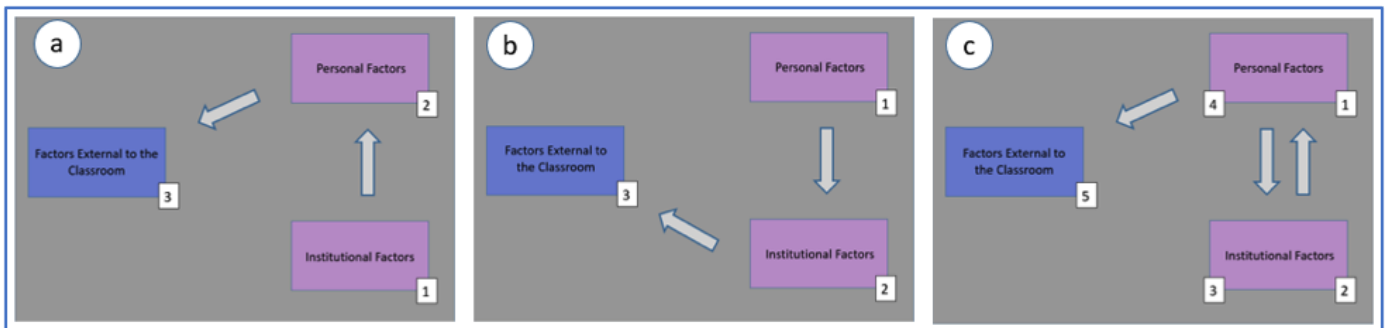
Figure 5: Three ways the institutional and personal factors surrounding teaching could interact.

address the mechanism for that interaction. There are three possibilities we can see for how personal and institutional external factors could dictate teaching: 1) Institution Influences Personal Factors: Institutional factors are driving personal factors like motivation and professional development. An instructor's personal goals do not necessarily align with the institution where they teach. However, once an instructor joins a university the personal factors that influence their teaching change and become shaped by their institution (Fig 5a). 2. Personal Factors Drive Institution Selection: In this case, personal factors affecting teaching are perhaps pre-determined by the instructor, and the instructor is drawn towards or simply has better chances to be hired at an institution that matches those goals (Figure 5b). 3. Mixed Effect: In these cases, a mixture of the two previous scenarios occurs. The instructor starts out with personal factors that drive them to select or be selected by a like-minded institution. However, once they have joined the institution their personal factors become shaped even more to be in line with institutional goals (Figure 5c).

Future studies should address which of these three scenarios are occurring. Understanding which came first, the professors' personal goals, or the institutional goals, or both will invariably help us understand the system of higher education in North America and what role teachers play inside of that system.

Institutional Stress Impacts Faculty Teaching Experience

Stress stemming from their institutions had a large impact on faculty experience, making it harder for faculty to achieve their personal goals. Often during Greg's interview the idea of making an impact came up. He recognizes that he can make this impact through teaching and through writing. Greg is currently writing a book in his field. University-related stress and responsibilities can interfere with Greg's goals. Despite the security of his tenured position there was a significant amount of pressure on Greg caused by administrative duties associated with his position. He expresses frustration that his administrative duties often distract from his teaching and writing pursuits. "I just don't love it [administrative duties] – I'd really much rather be in the lab or in the field or teaching". Regarding the impact he can make by writing, Greg says: "Give me more time to write

stuff that's meaningful. There's a shallowness to this stuff that we have to write that just takes so much time and its half-life is so short." Clearly, aspects of Greg's job, driven by the institution, are causing him stress that affects his ability to make an impact and to teach.

John's goal is to help develop his students into critical thinkers and spark their interest in his field through student-driven teaching. However, the instability and low salary associated with his job status were subjects of much stress, one of his more frequently used words during the interview. John's position is solely a lecturer position and he therefore is given no teaching assistants and relatively low funding. He has little opportunity for any expensive experiential learning activities and when he brings in examples for the class they are from his private collection instead of a university collection. This could make active learning and experiential learning activities much more challenging to conduct and hinder John's goals. In the interviews John expressed a great deal of frustration and "anxiety" surrounding his adjunct position and his hope to find a more permanent position in the future. Speaking about his job search, John says "What it comes down to, after your qualifications, is this kind of amorphous thing that they call "the fit". And to get that, it's like winning the lottery. I haven't found that fit yet." He also indicates that the pressure he feels from his job status affects his motivation as an instructor and faculty member. "I basically come in, teach my courses, and leave... There's no incentive for me to do anything more... In an adjunct position, loyalty and extra work is not rewarded at all." This lack of incentive coupled with increasing rejections from potential positions can wear adjunct faculty thin. John states: "There is almost an attitude [among fellow adjuncts] of 'well, pardon me, it's hard to really care when this place doesn't really care about me'." The stress John is facing that is caused by job status and institutional factors he cannot control is detrimental to his ability to teach. It causes him stress and also degrades his incentive to be exceptional at his job.

While stress from the job itself could hinder Greg and John's personal goals, their jobs also, and perhaps more insidiously, intervened in their professional development. If we assume that professional development in pedagogy improves teaching then the climate Greg is in could be doing both Greg and the students he teaches a disservice. The R1 status of the university placed more emphasis on research than teaching, making it so that Greg is not often evaluated on his teaching (only once that he can recall in 10 years!). Greg's teaching skills are mostly self-taught because his university does not focus on pedagogy or professional development in teaching. It should be noted that although we can show that the climate of the university is less conducive to teaching development workshops and programs we have no evidence that this effects teaching quality. In fact, in our course observations we saw that overall Greg is an excellent instructor with command of the classroom and a clear vision for his students. John's position also stymies his professional development as a STEM professorial candidate. Science professors applying to any institution must not only be good teachers but are also expected to have publications and current research projects. As John says: "How are you supposed to publish [scientific research] if you just have an adjunct position and you're just supposed to just be teaching, you have no infrastructure to do research here – it's tough, right... These are the kinds of challenges that face adjuncts and that's the kind of background anxiety and stress I'm talking about that's always there." In interviews the word "stress" came up most frequently when the words "teaching" and "research" were also prevalent, or when "research" was prevalent. "Teaching" alone did not seem to spark usage of the word stress as frequently. This association of the word "stress" with research and with the intersection between teaching and research might indicate John's job search, keeping his research current while looking for teaching positions. Surely this "background anxiety" is not making teaching any easier for John.

Institutional stresses pose more of a threat than just a pesky distraction from teaching. In fact, in a study conducted with new faculty the most important factor governing self-perception of success was professional balance. Other factors examined, such as expectations, collegiality, location, and personal balance, had greater effects on indirect measures of success such as heath, stress, and job and satisfaction, but the factor that new faculty used most to judge their own success was professional balance. Helping faculty to balance institutional demands with their personal life is essential to help them feel successful and, thus, to teach most effectively (Stupnisky et al. 2015).

Institutional Factors can Positively Impact Teaching Experiences

It should be noted that in our analysis there were also aspects of the institution climate that positively affected faculty. Indeed, if stresses posed by the institution or job can impact faculty so negatively it follows that positive influences of the institution could have positive effects on faculty teaching experience.

Despite administrative responsibilities and pressure

Greg's position as a tenured faculty member offers him security and freedom. He says: "I feel very secure in my job – that's the beauty of having tenure". His job status allows him the freedom to develop course material while feeling "creative" and having "fun". In his syllabus we see that Greg designed a discussion component to his course that includes experiential and active learning activities, such as museum visits and laboratory sessions. The institution in part made this possible as they fund such activities and also fund three teaching assistants for Greg's course. Greg also uses the freedom his position offers to develop new classes. One course Greg has recently developed offers undergraduates the opportunity to study abroad, investigating Greg's field in more depth. John's adjunct position offers him the flexibility to design his own lectures and relieves him of many of the institutional administrative duties he would otherwise be obligated to perform. Despite the obvious drawbacks we discussed earlier, as an adjunct, John receives support from fellow faculty members and feels they regard him as an equal, even if he is only in a temporary position. The institutional climate for John, at a liberal arts college, is one in which teaching is an integral value. John's instruction is evaluated three times per semester. John attends regular professional development workshops and feels support regarding teaching from his colleagues: "There is great collegiality among the faculty about talking about your teaching approach and your teaching methods. Yes. Even to me as an adjunct, whatever, there's no difference there. They're very happy to discuss if you're having issues in your class, if you feel something isn't working. It's fantastic." This kind of atmosphere could help John develop as a teacher and perhaps offset some of the development in research that he loses while serving in his position.

It is clear that both types of positions and institutions come with drawbacks and advantages. However, it's important that we continue to study those things that can be changed, such as what goals the university champions, what duties they expect of faculty, or the longevity of the positions they offer. Our research shows that the institution can have strong impacts, both positive and negative, on faculty teaching experience. Further research in these areas could help us identify what institutions can do to best support faculty in their roles as teachers.

Conclusion

We sought to understand how personal factors and institutional factors may affect STEM instruction using two case studies in higher education. While there is a plethora of complex factors important to teaching, we found that the institutional factors seem to be what are most driving both self-reported and observed teaching. Institutional factors such as job status and institution support of teaching can both support teaching experience and contribute to stress that hinders faculty instruction. More investigation on this topic is necessary to understand how these stresses can be mitigated. We must work on ways to help faculty find professional balance and help faculty to develop personal factors such as motivation to be more compelling teachers. We found an interaction between institutional and personal external factors on faculty teaching experience. Further research should explore this interaction further so we can understand how the process of professor recruitment is controlled and what role the institution has in shaping its instructors or what role instructors have in shaping the institution. Our current higher education climate should be wary of adjunct positions as these positions appear to cause undue stress to faculty. We should also be wary of saddling tenured faculty with excessive administrative duties which can also negatively affect their ability to be effective instructors. We can conclude that external factors outside the classroom, both institutional and personal, can drive faculty course development, faculty teaching practices, and the personal experience of faculty members. While many factors impact teaching in higher education STEM classes, it appears that job status, institution type, and institution support of teaching contribute significantly to faculty teaching experience and stresses. Institutions can be doing more to support their faculty. We can help improve faculty experience and teaching by changing the kind of positions, support, and university climate offered by the institution.

References

Amundsen, C., & Wilson, M. (2012). Are We Asking the Right Questions?: A Conceptual Review of the Educational Development Literature in Higher Education. Review of Educational Research, 82(1), 90–126. http://doi.org/10.3102/0034654312438409

Bettinger, E., & Long, B.T. (2005). Help or hinder? Adjunct professors and student outcomes. What's Happening to Public Higher Education.

Bryman, A., & Burgess, R.G. (1997). Analyzing qualitative data. (A. Bryman & R. G. Burgess, Eds.)User-friendly handbook for mixed method evaluations. New York: Taylor & Francis Group.

Charmaz, K. (2006). Constructing Grounded Theory: A Practical Guide through Quali-

tative Analysis. (D. Silverman, Ed.)Journal of Chemical Information and Modeling (Vol. 53). Thousand Oaks: SAGE Publications. http://doi.org/10.1017/CBO9781107415324.004

Dickens, D. (2011). Measuring adjunct instructor job satisfaction by using Herzberg's motivation-hygiene theory. LAMAR UNIVERSITY - BEAUMONT.

Edwards, R., & Holland, J. (2013). What is Qualitative Interviewing? (G. Crow, Ed.). New York: Bloomsbury Publishing Plc.

Elliott, R.W., & Oliver, D.E. (2016). Linking faculty development to community college student achievement: a mixed methods approach. Community College Journal of Research and Practice, 40(2), 85–99. http://doi.org/10.1080/10668926.2014.961590

Fike, B.D.S., & Fike, R. (1987). Does Faculty Employment Status Impact Developmental Mathematics Outcomes?

Gupta, S., Harris, D.E., Carrier, N.M., & Caron, P. (2006). Predictors of Student Success in Entry-Level Undergraduate Mathematics Courses. College Student Journal, 40(1), 97–108.

Hativa, N. (1995). The department-wide approach to improving faculty instruction in higher education: A qualitative evaluation. Research in Higher Education, 36(4), 377–413. http://doi.org/10.1007/BF02207904

Hellsten, L.A.M., & Prytula, M.P. (2011). Why teaching? Motivations influencing beginning teachers' choice of profession and teaching practice. Research in Higher Education Journal, 13, 1–19.

Jull, S., & Mirenda, P. (2016). Effects of a Staff Training Program on Community Instructors Ability to Teach Swimming Skills to Children With Autism. Journal of Positive Behavior Interventions, 18(1), 29–40. http://doi.org/10.1177/1098300715576797

Kane, R.T., Shaw, M., Pang, S., Salley, W., & Snider, J.B. (2015). Relationships among Faculty Training, Faculty Degree, Faculty Longevity, and Student Satisfaction in Online Higher Education. Online Journal of Distance Learning Administration, 18(4).

Kraglund-gauthier, W.L., Chareka, O., Orr, A.M., & Foran, A. (2010). Teacher Education in Online Classrooms: An Inquiry into Instructors' Lived Experiences. The Canadian Journal for the Scholarship of Teaching and Learning, 1(2). http://doi.org/10.5206/cjsotl-rcacea.2010.2.4

Mattheis, A. (2014). Fostering improved anatomy and physiology instructor pedagogy. Advances in Physiology Education, 38(4), 321 – 329. http://doi.org/10.1152/advan.00061.2014

OECD. (2010). Learning Our Lesson: Review of Teaching Quality. Institutional Management in Higher Education. http://doi.org/10.1787/9789264079281-en

Osterman, D.N. (1984). Motivating Faculty to Pursue Excellence In Teaching Motivating Faculty to Pursue Excellence In Teaching. To Improve the Academy, (60).

Rasheed, M.I., Aslam, H.D., & Sarwar, S. (2010). Motivational Issues for Teachers in Higher Education: A Critical Case of IUB. Journal of Management Research, 2(2), 1–23.

Sonnert, G., Sadler, P.M., Sadler, S.M., & Bressoud, D.M. (2014). The impact of instructor pedagogy on college calculus students' attitude toward mathematics. International Journal of Mathematical Education in Science and Technology, 46(3), 370–387. http://doi.org/10.1080/0020739X.2014.979898

Strauss, A.L. (1987). Qualitative Analysis for Social Scientists, 319.

Stupnisky, R.H., Weaver-Hightower, M.B., & Kartoshkina, Y. (2015). Exploring and Testing the Predictors of New Faculty Success: A Mixed Methods Study. Studies in Higher Education, 40(2), 368–390. http://doi.org/10.1080/03075079.2013.842220

Using Superheroes in Disease-Centered Physiology Instruction for Biomedical Engineering Undergraduates

Aaron E. Chiou

Meinig School of Biomedical Engineering, Cornell University
aec267@cornell.edu

Abstract

Biomedical engineering (BME) undergraduates need to be able to apply physiological knowledge to design solutions for health-related problems. However, physiology is often taught to BME students in a traditional lecture-based format that rarely introduces broader applications of course material. To investigate whether BME students are receptive to, or benefit from, disease-centered physiology instruction, a sophomore-level BME physiology class of 37 students was taught a module on the musculoskeletal system with focus on altered physiology applications such as disease and superheroes. Classroom observation, student surveys, and quiz performance data suggested that students had a positive experience during the module and performed well on assessments given a week following instruction, as compared to assessments after traditional instruction.

Introduction

Physiology is a core component of the undergraduate biomedical engineering (BME) curriculum; however, an optimal method to teach the subject in a manner suited for engineering students has yet to be clearly defined. With the rise of undergraduate BME programs, the role of physiology instruction has increasingly shifted from teaching the details of all organ systems, as is the case in biological and medical programs, to providing a solid framework for students to approach new physiological problems (Troy & Linsenmeier, 2003).

Biomedical engineers would benefit from learning physiology with design applications in mind, developing a mindset for thinking through physiology problems rather than memorizing factual details. This includes identifying a problem, diagnosing the underlying issue, and designing a solution to address that issue. In the absence of context, the material in a physiology course will likely fail to generate student interest or motivation to learn. My undergraduate physiology course was designed for life sciences majors and focused on normal physiological systems while barely touching on relevant diseases, leaving me bored and thinking there was little reason to attend class. Surveys indicate that BME undergraduates are most motivated and learn best when their courses are taught with a broader application context (Dee, Nauman, Livesay, & Rice, 2002). My experience supports this statement, seeing as the lack of real-world applications in my physiology course was never able to engage me.

Cornell University has recently begun offering an undergraduate BME major with application- and concept-driven courses at its core. One year ago, as a Graduate Assistance in Areas of National Need (GAANN) Fellow and soon-to-be Teaching Assistant, I attended planning meetings for the new BME physiology course, Physiology of Human Health and Disease (BME 2010). The idea was to teach physiology in homeostasis and disease in parallel, with each organ system taught in a modular format. We were driven to create a course that helped BME undergraduates develop a framework to approach physiological problems, rather than teaching a class that emphasized memorizing facts.

However, the implementation of the course in its first year diverged from its original design, as it defaulted to a traditionally taught, lecture-based course focusing on details of physiological systems rather than broader applications. My GAANN Fellowship provided an opportunity to develop and teach my own module in the course, all while receiving training in current pedagogical practice. I taught the musculoskeletal system last year, attempting to refer more often to current engineering approaches that address conditions including osteoarthritis and muscular dystrophy. The second time the course was offered, I returned to teach the musculoskeletal system module as a guest lecturer. This time I adjusted my lectures to break away from the traditional model of teaching physiology, and took this as a chance to incorporate more active learning strategies and greater focus on disease applications and to test run a module's worth of the original vision of the course. In doing this, I sought to answer the question of whether the inclusion of concrete applications into otherwise standard physiology lectures would keep BME students moti-

vated and help them to learn.

To determine whether my disease-centered module had the intended effect, I combined insights from observation, student-reported survey data, and student quiz performance. My objectives were to determine: 1) how receptive students were to a module of disease-centered instruction in the middle of a traditionally taught physiology course, and 2) how effective disease-centered instruction was in helping students learn. I hypothesized that students would adapt readily to this structure of instruction, and that focusing on altered physiology applications such as disease processes and superheroes would improve conceptual understanding and lead to elevated student performance on assessment questions related to material from the disease-focused module.

Literature Review

In recent years, instructors have reported instructional methods meant to improve student learning and development of critical thinking skills in physiology. For example, a flipped physiology classroom in which students worked in teams of four to six for laboratory exercises, case studies, and team assessments led to 85.5% of students earning higher exam scores and final course grades than predicted by their major GPA (DeRuisseau, 2016). Compared to the traditional classroom approach, the flipped classroom allowed for an increased proportion of exam questions at the application level of Bloom's taxonomy. Interestingly, the benefits reaped from a flipped classroom approach may largely be the result of active learning strategies, as a comparison of active learning in flipped and non-flipped approaches resulted in similar student perceptions and performance on low- and high-level assessment questions regardless of class structure (Jensen, Kummer, & Godoy, 2015). A recent meta-analysis of 225 published and unpublished studies further supports this result, finding that active learning interventions have increased average exam scores by 6% across STEM disciplines and class sizes, with greatest benefit in small classes of under 50 students (Freeman et al., 2014). Active learning strategies that compose even 10% of total class time can yield a similar effect size. Together, these studies suggest that even minor incorporation of active learning strategies could yield improvement in student learning outcomes in physiology instruction.

To engage and motivate BME students to learn, incorporation of real-world applications has also proven effective. Surveys indicate that BME undergraduates are most motivated and learn best when their courses are taught within a broader application context (Dee et al., 2002). When introduced at the beginning of a modular lecture, human pathologies hold the attention of a class and provide a frame of reference for students to understand how disease manifests from altered homeostatic processes (Guilford, 2005; Matsuki et al., 2009). This approach is easily adaptable to time-constrained courses in basic science, leading to increased course ratings without requiring extensive changes to a traditional lecture format (Walters, 1999). Furthermore, students have been shown to learn more and develop innovative thinking when "challenged" to understand a patient's medical issue compared to traditional instruction (Martin, Rivale, & Diller, 2007). Students report that contextual, disease-based lectures not only stimulate their interest and motivation to learn, but make it easier to draw connections between basic science and their research (Matsuki et al., 2009). Another approach science educators have taken to motivate student learning is incorporating discussion of superheroes from popular culture, such as Batman and Captain America, to explore physiology concepts from a familiar frame of reference. While this approach is unconventional and less commonly described, students are described as more engaged when instructors use superheroes to discuss the limits of human physiology and physical performance (Brown, Smith, McAllister, & Joe, 2017; Zehr, 2011). However, there are no studies to-date that examine the use of superheroes to teach physiology to undergraduate BME students.

Overall, pedagogical techniques in the physiology classroom such as active learning strategies, disease-based lectures, and discussion of superheroes have shown promise for promoting innovative, higher-level thinking, as well as enhancing student motivation and learning. It remains to be seen whether these strategies are effective for BME students learning physiology, and whether these approaches encourage students to develop a framework to understand how altered states of physiology can be explained in terms of perturbed homeostasis.

Methods

To evaluate student response to a module focused on altered physiological states, I analyzed qualitative and quantitative data from observations, student surveys, and quizzes.

Course Details

I was a guest lecturer for three 50-minute class sessions in a physiology course for sophomore biomedical engineering majors, Physiology of Human Health and Disease (BME 2010), at Cornell University. A total of 37 students were enrolled in the course. The course was primarily taught in a traditional lecture format,
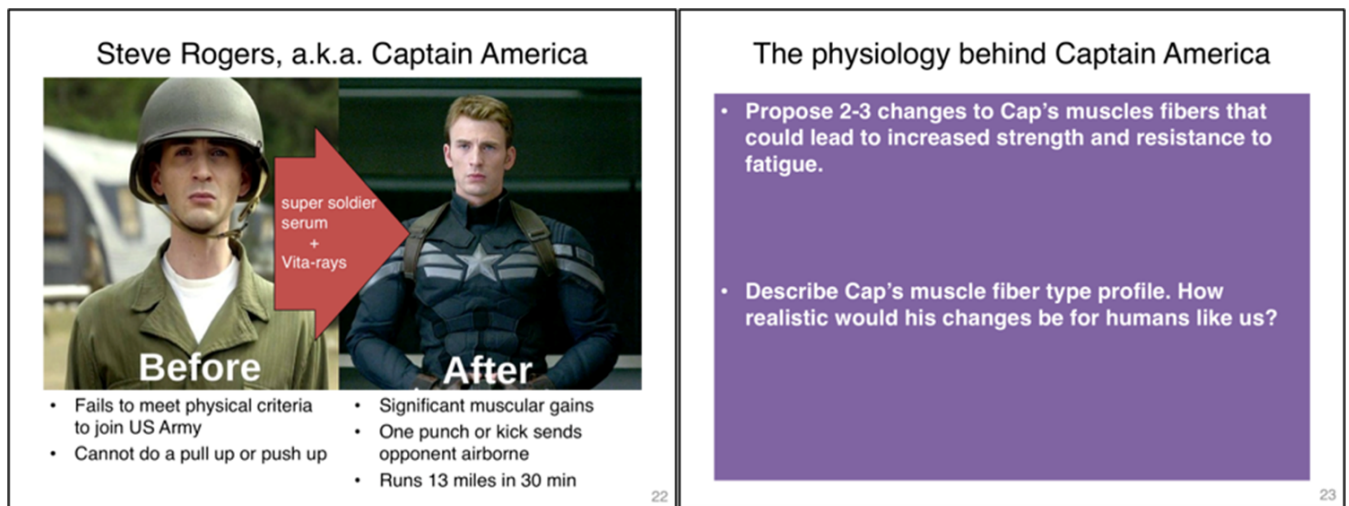
Figure 1. Example slides from lecture on muscle physiology centered on Captain America.

and I taught my module on the musculoskeletal system six weeks into the semester. To facilitate discussion during each lecture, I used broader applications such as disease (osteoarthritis, muscular dystrophy) and superheroes (Captain America). Diseases were introduced at the beginning of class sessions, with slides and think-pair-share questions relating back to diseases. Captain America was introduced in the middle of the third lecture, with a comparison of his physical capability and muscularity before and after his superhero transformation (Figure 1) (Brown et al., 2017). After this, two questions considering Captain America's muscle physiology were discussed in think-pair-share style.

Qualitative Methods

My own observations and those of two observers of the final lecture of the student reception to my teaching are included here. Students were also administered an anonymous survey at the end of the last lecture, which included the following open-ended questions, the responses to which were coded for themes: 1) What did you think was most effective about this style of teaching?, and 2) What did you think could have been better about this style of teaching?

Quantitative Methods

Likert-scale survey questions were used to quantitatively assess student reception to my disease-centered module. The Likert scales provided two negative (1-2), one neutral (3), and two positive (4-5) response options. The following questions were included: 1) How do you rate your overall learning experience during the musculoskeletal module?, 2) How engaged were you in the altered physiology-centered approach (more time devoted to discussion and applications) compared to traditional lecture?, and 3)

How effective was the altered physiology-centered approach for your learning of the musculoskeletal system? Responses to each Likert-scale survey question were compared to an expected value of 3 using a Wilcoxon signed-rank test, with a p-value of less than 0.05 indicating a response statistically significant different from neutral.

To evaluate student learning and the effectiveness of the disease-centered module, quiz performance was evaluated for the disease-centered module compared to a traditionally taught (control) module following it. The quiz was given one week after the end of the disease-centered module. Multiple choice questions and open-ended short answer questions were graded by BME 2010 teaching assistants and quizzes with students' personal information removed were provided to me for analysis. Possible student scores were full (1), partial (0.5), and no (0) credit. Percentage of points earned out of those available was calculated per quiz question and per module category (disease-centered or control). A paired t-test was used to determine whether there was a statistically significant difference ($p<0.05$) in student performance between the disease-centered and control modules.

Results

In general, observations indicated that students were more engaged and interested during activities and discussion compared to lecture. During lecture-heavy portions, students had difficulty paying attention, and sometimes, even staying awake. However, activities and discussions appeared to increase student engagement, and Captain America seemed to improve attention and engagement even more so than discussion of diseases did.

Survey results indicated that students were posi-
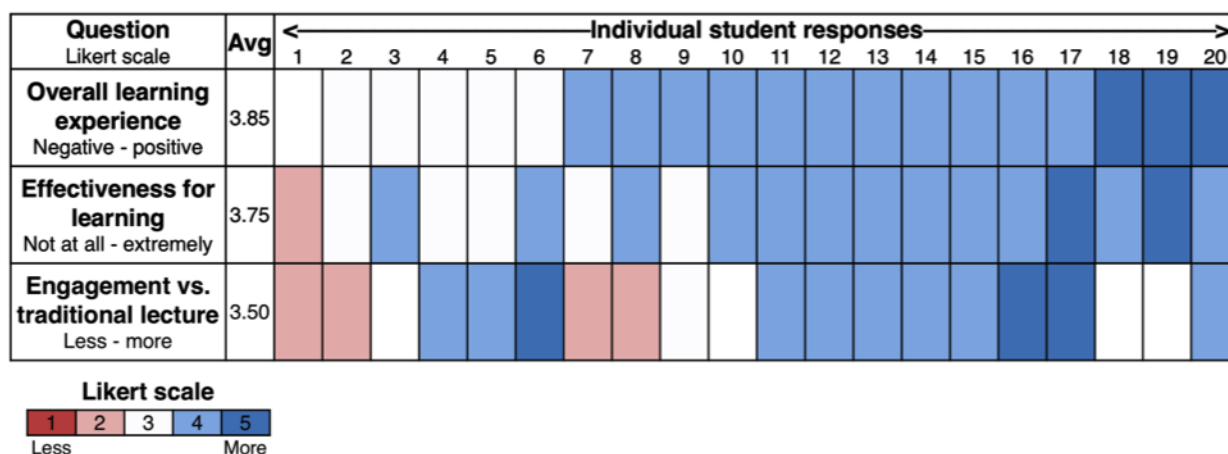
Figure 2. Heat map of student responses to Likert scale questions on feedback survey. Each column represents an individual student's responses. Warmer colors represent lower numbers (negative) and cooler colors represent higher numbers (positive). 'Avg' column indicates mean rating.

Question Likert scale	Avg
Overall learning experience Negative - positive	3.85
Effectiveness for learning Not at all - extremely	3.75
Engagement vs. traditional lecture Less - more	3.50

Likert scale: 1 2 3 4 5 (Less — More)
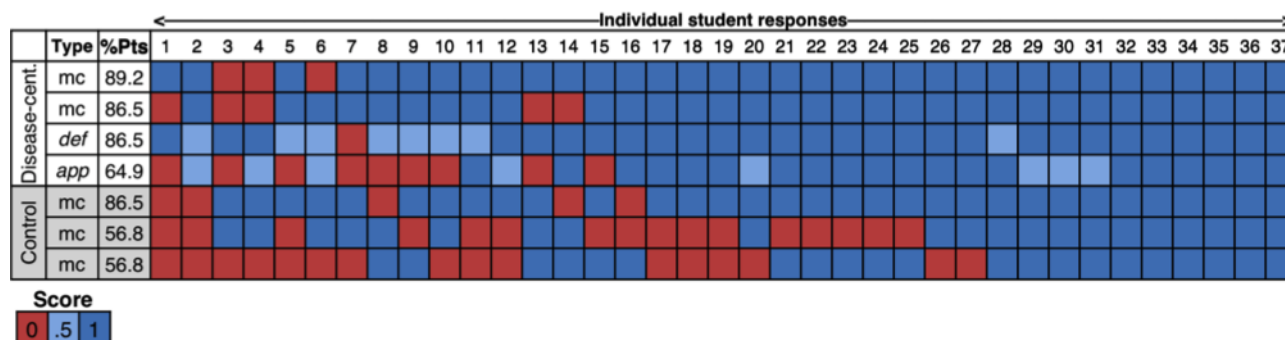
Figure 3. Heat map of student responses to quiz questions one week following disease-centered module. Each column represents an individual student's responses. Red represents no credit, light blue represents partial credit, and blue represents full credit. 'Type' key: 'mc' – multiple choice, 'def' – definition, 'app' – application. '%Pts' indicates percentage of available points earned by students per quiz question.

	Type	%Pts
Disease-cent.	mc	89.2
	mc	86.5
	def	86.5
	app	64.9
Control	mc	86.5
	mc	56.8
	mc	56.8

Score: 0 .5 1

tive about learning but experienced varied engagement during the disease-centered module (Figure 2). Of the 20 students (54% of those enrolled) who responded, the mean scores of the Likert scale questions tended toward positive. Most students had a generally positive experience (p<0.001), with all respondents rating their overall learning experience between neutral and extremely positive. Similarly, students viewed the disease-centered module as effective for their learning (p = 0.), aligning with their perception of the overall learning experience. However, students reported a greater frequency of negative responses when rating their degree of engagement in the disease-centered lectures compared to traditionally taught lectures. While most students felt similarly or more engaged during disease-centered lectures when compared to traditionally taught lectures, a few students reported feeling less engaged despite neutral or positive learning experiences and effectiveness ratings (Figure 2, students 2, 7, 8). Student engagement was marginally statistically different from neutral (p = 0.0536).

Open-ended survey questions provided greater insight into the student survey responses. Coding for themes revealed that students felt the most effective aspects of the disease-centered instruction were its applicability to the real world and fostering of critical thinking skills, and that they appreciated the discussion of Captain America. However, students desired more interaction, greater motivation to pay attention and stay engaged, and more application-based content. For example, several students mentioned that the course instructor had been providing lecture notes with blanks to be filled in, which kept them motivated to pay attention in class, which I was not aware of when I began teaching the module. Students also gave comments more directed at my teaching itself, with some praising the presentation style and others taking issue with the pace, clarity, or delivery style.

Results of a quiz that covered concepts from the

disease-centered module and the following module that was taught using traditional lecture style suggested that the superhero- and disease-centered approach of teaching can improve student performance (Figure 3). Multiple choice and an open-ended definition question from the disease-centered module yielded similarly high student performance, with over 85% of possible points being earned. On the other hand, students earned only about 65% of possible points from an open-ended application question that tested concepts like those in the Captain America muscle fiber discussion. While this lower performance could be attributed to the question's higher-level ranking on Bloom's taxonomy, the overall percentage of points earned for that question still exceeds that for two out of three of the control module's multiple choice questions. Overall, quiz data indicates that students scored better on quiz questions from the disease-centered module (p = 0.0056), despite needing higher-ordered thinking and more time elapsed between relevant lectures and assessment.

Discussion

The purpose of this study was to determine how receptive undergraduate BME students were to a module of superhero-based and disease-centered instruction in the middle of a traditionally taught physiology course (BME 2010), and whether this method of instruction would improve student learning. In comparison to traditional physiology instruction, which first focuses on normal physiological systems and then introduces relevant diseases later, the disease-centered approach of teaching combined with the use of superheroes appeared to provide a positive and effective learning experience for biomedical engineering students. In addition, students appreciated the context and opportunities for critical thinking afforded by the disease-centered instruction method and the focus on superheroes. However, students reported feeling mixed levels of engagement, perhaps due to unanticipated lack of "fill in the blank" lecture notes that the course instructor had been providing. Notably, the inclusion of superheroes in this module may have fostered more student engagement than the disease-centered aspects, as indicated by observations and survey responses. Ultimately, student performance on quiz questions related to the altered physiology-focused module was better compared to those related to a traditionally taught module.

Based on this study, I believe that adapting traditionally taught physiology lectures to center more on applications such as disease and superheroes has the potential to improve learning outcomes for BME undergraduates. While this research may not be generalizable to all BME classrooms, it underscores the importance of informing BME undergraduate education with existing research on teaching in the life sciences and engineering. Positive student reception of superhero-centered discussion warrants further research on the benefits of using superheroes in engineering education. Lastly, it may be valuable to consider how students define and view their own engagement in the classroom, what motivates it, and how it affects their learning.

References

Brown, S. P., Smith, J. W., McAllister, M., & Joe, L. (2017). Superhero physiology: the case for Captain America. Advances in Physiology Education, 41(1), 16–24. https://doi.org/10.1152/advan.00106.2016

Dee, K. C., Nauman, E. A., Livesay, G. A., & Rice, J. (2002). Research Report: Learning Styles of Biomedical Engineering Students. Annals of Biomedical Engineering, 30(8), 1100–1106. https://doi.org/10.1114/1.1512677

DeRuisseau, L. R. (2016). The flipped classroom allows for more class time devoted to critical thinking. Advances in Physiology Education, 40(4), 522–528. https://doi.org/10.1152/advan.00033.2016

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. Proceedings of the National Academy of Sciences, 111(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Guilford, W. H. (2005). "Shrink wrapping" lectures: Teaching cell and molecular biology within the context of human pathologies. Cell Biology Education, 4(2), 138–142. https://doi.org/10.1187/cbe.04-10-0054

Jensen, J. L., Kummer, T. A., & Godoy, P. D. D. M. (2015). Improvements from a Flipped Classroom May Simply Be the Fruits of Active Learning. CBE-Life Sciences Education, 14, 1–12. https://doi.org/10.1187/10.1187/cbe.14-08-0129

Martin, T., Rivale, S. D., & Diller, K. R. (2007). Comparison of student learning in challenge-based and traditional instruction in biomedical engineering. Annals of Biomedical Engineering, 35(8), 1312–1323. https://doi.org/10.1007/s10439-007-9297-7

Matsuki, N., Takeda, M., Yamano, M., Imai, Y., Ishikawa, T., & Yamaguchi, T. (2009). Effects of unique biomedical education programs for engineers: REDEEM and ESTEEM projects. Advances in Physiology Education, 33(2), 91–7. https://doi.org/10.1152/advan.90120.2008

Troy, J. B., & Linsenmeier, R. A. (2003). Optimizing the delivery of content in physiology instruction. IEEE Engineering in Medicine and Biology Magazine, 22(4), 80–87. https://doi.org/10.1109/MEMB.2003.1237506

Walters, M. R. (1999). Case-Stimulated Learning Within Endocrine Physiology Lectures. Advances in Physiology Education, 21(1), 74–78.

Zehr, E. P. (2011). From Claude Bernard to the Batcave and beyond: using Batman as a hook for physiology education. Advances in Physiology Education, 35(1), 1 LP-4. https://doi.org/10.1152/advan.00120.2010

Evaluating Efficacy of a Novel Two-Stage Homework System

Ashley Earle[1,*], Derek Holyoak[1], Shivem Shah[1]and Kimberly Williams[2]

[1] Meinig School of Biomedical Engineering, Cornell University
[2] Center for Teaching Innovation, Cornell University
[*] Corresponding Author: ak978@cornell.edu

Introduction

Engineering is often seen, especially by students, as synonymous to creative problem solving. Engineers are expected to find innovative solutions to problems with respect to constraints and ever changing variables (Jonassen et al., 2006). The development of these problem-solving skills and overall engineering competence and expertise is contingent upon learning foundational concepts (Streveler et al., 2008). Within a problem, foundational concepts allow people to organize and define the overall framework, extract meaningful information and quantities, form and troubleshoot approaches logically, as well as to identify components of successful approaches to form novel procedures (Rittle-Johnson et al., 2001).

Within engineering, it is safe to say that homework is considered a necessary part of mastering almost all engineering competencies. As recently graduated engineering students, we have experienced and felt the effects of poorly executed, and poorly designed, homework. The engineering curriculum is quite demanding which often leads to students being forced to make choices on how they spend their time. Do you work out all the problems from scratch on your own, or do you google the question and see what help you can find? Do you work alone, with a group, or even take a divide and conquer approach to get your work done? Do you learn as you go or figure you'll learn it in time for the exam? Do you receive feedback on your performance? If you do receive it, is it more than a grade? Do you get it while you still remember your thought process? Do you receive if before the exam on the material? Do you only focus on math, or do you also have homework focused on the concepts? Do your exams test both?

The list of questions could continue, but the gold-standard for engineering homework design is answerable. At most institutions, engineering homework is primarily math focused and feedback is limited, if existent and is often given a very distally from the assignment. This distinctly goes against most of the recommendations that learning required frequent and fast feedback and also does not provide a good incentive to do more than just get the homework done (Martinez et. al., 2015). On the other side of the coin, students, often overwhelmed by their coursework, will do whatever it takes to get the work done and get the grade, and worry about fully unpacking the material when the exam comes. This needs no explanation, as "cramming" is well known to be an ineffective method of long term learning and leads to the persistence of misconceptions after the exam has passed. Although these experiences do not describe every engineering student or every academic institution, these practices are prevalent enough to pose a major hindrance to students that often relegates homework from a true and deep learning experience to busywork.

This paper explores the role of homework in learning and specifically attempts to evaluate a novel method of homework design that is aimed at improving feedback, conceptual learning, correction of misunderstandings, and student anxiety over work load – three of the main issues that we believe are detrimental to the effectiveness of homework as a learning tool. Our method is called a two-stage homework system. In this system, students first submit an initial draft of the homework assignment in which they are assessed on their thought process and conceptual knowledge. After submission, students will receive feedback on their conceptual draft and subsequently submit a second draft with their final conceptual and quantitative answers. This allows students to think through a concept, reflect upon their misconceptions, and correct them with the guidance of the instructor's feedback. This tool will be evaluated in the context of a sophomore level Thermodynamics class, a subject that has traditionally been difficult for students to learn due to the large conceptual payload.

Evolution of (and Issues With) Homework

Homework is broadly defined as "tasks assigned to students by school teachers to be carried out during non-school hours" (Cooper, 1989). The role of homework is multifaceted and should provide students with a structure to tackle appropriate challenges, practice material discussed in class, get feedback on performance, and experience success in learning or mastering a concept (Wankat, 2001). This is extremely important as students will usually assume that they

are correct until they see grade deductions for incorrect answers and the correct solutions posted. In theory, excellent homework performance should lead to excellent exam scores. However, studies show that homework grades are an inconsistent predictor of exam grade (Bas, 2017; Fernandez et al., 2006). These disparate results have fueled an age-old debate of whether homework is actually helpful for student learning.

At the onset of the 20th century the US at-large was against homework (Watkins and Stevens, 2013). However, the initial success of the Soviet Union in launching Sputnik I during the Cold War changed opinions, bringing a "back to basics" attitude to the education system that heavily focused on STEM fields and homework. The nationwide attitude and policy was reversed again in the 60s and 70s to allow for more student freedom and to encourage students to make their own academic choices. The pendulum swung back and forth through various movements such as A Nation at Risk (1980s), Goals 2000 (1990s), No Child Left Behind (2000s), and most recently, Common Core (2010s)(Bas, 2017; Watkins and Stevens, 2013). With each movement came ample studies supporting their hypothesis; there is still no consensus on homework's utility.

Reading between the lines of these articles brings a very important point to the foreground. Homework is most useful when it closes the feedback loop and when students engage in it as intended (Nicol and Macfarlane-Dick, 2016) It must provide appropriate challenges, time to practice, regular feedback, and allow individual students to experience success, i.e. master the material, or the impact it has on students will vary based on how the individual student approaches homework. In a study based on social cognitive theory, it was found that self-efficacy, or the belief of an individual that he/she can perform at the required level, was the primary predictor of correlation between homework and final course grade (Bembenutty, 2012). In fact, self-efficacy, and not intrinsic motivation or help seeking behaviors, was the primary predictor of whether students were satisfied with their homework completion and what they learned from homework. Although this study did not correct for class size, homework type, or instructor differences, it provides a useful examination of the question, "Does homework help or hurt?"

If a student does not believe that they can perform at the appropriate level, how will that impact the way that they do homework under normal circumstances? It seems from the literature, and experience, that students will either value their grade and turn to cheating or decide they do not care and turn in poorly done assignments, or no assignment at all. What complicates this is that many engineering students don't view copying homework as cheating. Almost all students will say that copying someone else's work on an exam is cheating, but only 70% of students feel that it is cheating to copy from another student on a homework assignment, 50% that it is cheating to copy an assignment handed down from a more senior student, and only 20% that it is cheating to copy something directly from the textbook (Passow, 2006). These results, although not definite, represent the views of nearly 700 students at 11 different campuses. This general attitude explains how students working together in study groups would not consider it cheating to "work together" even if one or two people did most, if not all, of the work (Fernandez, 2006; Wankat, 2001). Interestingly, the number of students "cheating" on homework decreased from almost 30% during the second year down to almost 0% at the 4th or 5th years, when compared to their first year in college. It may be that these students realized that cheating on homework negatively impacted their learning and they changed their behavior (Williams, 2006).

Another study looked at a Midwest high school with high drop out and course failure rates, in part due to students either not doing homework well or not doing it at all (Watkins and Stevens, 2013). The school implemented a "No Excuses Homework Policy" that put in place specific steps to allow students to get the help they needed on homework and a strict accountability policy to make sure they followed through on the assignment. They saw both a 72% decrease in failure rates and improved student motivation and self-efficacy. It is important to note that students who were already doing their homework, or who were doing well in school without doing their homework, felt that the policy was unfair and did not see personal benefits like the other students did. From these reports, a school culture that values high quality homework and the efficacy of homework seems to be of paramount importance to the positive impact of homework on learning outcomes.

Formative Assessment and Feedback

Formative versus summative assessment.

Over the past few decades, teaching in higher education has gradually tried to incorporate formative assessment into classes that have traditionally used summative assessment. (Nicol & Macfarlane-Dick, 2006). The purpose of formative assessment is to improve student learning by providing ideas and helpful information to enhance their ability to complete a task (Brown et al., 1997). Summative assessment

focuses more heavily on ensuring students have sufficient knowledge in their program of study and is more a traditional approach to teaching and learning. While feedback is involved in both types of assessments, feedback is especially important in formative assessment. With formative assessment, students typically submit drafts of assignments and obtain feedback from the instructor before they submit the final version of assignments. This gives students the opportunity to reflect on any mistakes they originally made and actively correct for their previous misconceptions. While our study focuses heavily on the concept of formative assessment, traditional summative assessment techniques are beneficial as well. Specifically, summative assessments allow the instructors to gather feedback on their students' strengths and weaknesses, which can be used to improve future performance (Sadler, 1998). Thus, a combination of formative and summative assessment is most likely the most effective way to conduct classroom assessment.

One potential drawback of implementing formative assessment in the classroom is the growing student population and greater amounts of workload required from instructors (Covic & Jones, 2008). This idea is especially critical because learning is not as effective if either the teacher or student is overwhelmed. Maintaining a manageable amount of work for both the teacher and student is a common goal for all courses (Attwood, 2009). However, recent studies suggest that certain concepts can be taught without ever using summative assessment. Thus, instead of including formative assessment in addition to summative assessment, formative assessment can simply replace summative assessment in certain situations (Jenkins, 2010). For example, the required knowledge associated with summative assessment can be gained by encouraging students to independently seek and understand more reading materials, providing feedback on potential improvements in writing style and focusing on content during task-related assignments.

A brief summary of feedback history.

High-quality feedback interventions are critical to achieving successful formative assessment in the classroom. However, how to supply effective feedback to students has been a subject of debate for many years. In a somewhat controversial review of the literature performed by Ammons (1956), he concluded that feedback interventions improve student learning and motivation regardless of feedback strategy. This review is now one of the most cited reviews in feedback literature. However, Kluger & DeNisi (1996) pointed out that Ammons left out some key references that provided information that were inconsistent with his conclusions. Thus, one of the most influential reviews in the field ignored the high variability of feedback effectiveness. Kluger & DeNisi (1996) then sought to quantify the variability of feedback effectiveness with an extensive meta-analysis that analyzed 131 papers that led to 223,663 observations on feedback and performance. They found that the effectiveness of feedback was highly variable, and that approximately one-third of feedback interventions reduced performance. Based on these historical and seminal studies on feedback literature, a clear need exists in the present day to further investigate the types and timing of feedback that serve to improve performance rather than hinder it.

Categories and types of feedback.

One of the main obstacles in delivering effective feedback strategies in the classroom is determining what types of feedback consistently improve student performance and understanding. Although many models of feedback are available in the literature, one of the most influential models was developed by Hattie & Temperley (2007). They suggested feedback be separated into four categories (or "levels"): task feedback, processing feedback, self-regulation feedback, and self-feedback (Hattie & Temperley, 2007). Task feedback informs the student whether something is correct or incorrect, or if the student needs to provide additional information or problem-solving steps to reach the correct answer. Although this type of feedback is generally effective, it is a very basic form of feedback and thus does not lead to deeper processing or self-regulation skills. Processing feedback challenges students to think deeper about processes they are using to solve problems and try to relate processes to real-world scenarios. Self-regulation feedback helps students develop awareness of how to appropriately use their knowledge. This type of feedback can take the form of reminding students they already possess the skills necessary to accomplish a task, which encourages them to tackle problems with confidence. The last category of feedback is referred to as self-feedback, which is considered to be the least effective. Specifically, self-feedback simply informs students whether or not they did a "good job" on an assignment, and thus does not promote deeper thinking like processing feedback or self-regulation feedback. Although these four categories are all helpful in specific ways, keeping track of the types of feedback used in the classroom and emphasizing deeper levels of feedback such as processing and self-regulation is critical to maintain effective feedback practices.

The effects of positive versus negative feedback has

also been an active research area for many years. In a recent review investigating the effects of feedback on the brain and neural networks (Dion & Restrepo, 2016), positive feedback was shown to have consistent beneficial effects from childhood through adulthood, whereas negative feedback was more beneficial after adolescence due to specific areas in the brain needing to mature (Grammer et al., 2014). Thus, in higher education, both positive and negative feedback are useful in student learning. Furthermore, Dion & Restrepo (2016) noted the difference between direct versus indirect feedback. When feedback is ambiguous or indirect, it can potentially hinder learning (Ernst & Steinhauser, 2015) and activate neural functions associated with negative effects on learning (Sailer et al., 2007). Thus, feedback is most effective when it is "clear, explicit, and intelligible," and teachers are thus responsible to ensure the feedback they provide is understood properly by students (Dion & Restrepo, 2016).

Feedback models.

When effective feedback is successfully implemented in the classroom, students should be able to answer the following three questions according to a feedback model, also from Hattie & Temperley (2007). The questions are: "where am I going," "how am I going," and "where to next?" (Hattie & Temperley, 2007). The "where am I going" aspect of the model is focused on setting goals for students via specific cues and goal-oriented feedback. One objective in the classroom is to encourage students to commit to achieving a common goal. When this objective is accomplished, students are more likely to reach out for further feedback, which eventually results in self-regulation of their achievements (Locke & Latham, 1990). Thus, defining goals and supplying feedback to students related to those goals is a critical aspect of the "where am I going" part of the model. The second part of the model, "how am I going," involves a teacher, such as an instructor, peer, or the student themselves, supplying information related directly to performance on tasks and assessments. This part of the model is thought of as more traditional feedback, such as task or processing feedback. The last part of the model, "where to next," if used correctly with appropriate feedback, leads to enhanced challenges, developing strategies independently, and learning how to self-regulate tasks. Inappropriate use of this part of the model can lead to loading more work onto the students, which generally increases pressure and does not aid in learning. Importantly, the questions in this model need to be accomplished with clear objective statements and consistent, effective feedback.

When students are able to answer "where am I going," "how am I going," and "where to next," instructors have successfully implemented the model and students are capable of self-regulating their success. (Sadler, 1989).

Recent literature on formative assessment has expanded on the model from Hattie & Timperley (2007) and has made the goals of good feedback practice even clearer. Nicol & Macfarlane-Dick (2015) produced a seven-principle system that helps students become independent learners and improve self-regulation abilities through feedback. In summary, the seven principles are 1) clarify goals and expectations, 2) facilitate self-assessment and reflection, 3) deliver essential feedback information, 4) encourage dialogue between peers and teachers, 5) motivate and improve self-esteem, 6) provide opportunities to close the gap between current and desired performance, and 7) use student feedback to improve teaching. With these seven principals, students become aware of strategies they can use to regulate their own learning throughout college, and they can carry these strategies with them to remain life-long learners (Nicol & Macfarlane-Dick 2015).

Timing of feedback.

Timing is an aspect of feedback that is not clearly understood, and studies have produced conflicting results (Anderson et al., 1995; Butler et al., 2007). Specifically, research studies have analyzed whether feedback should be provided immediately following a task or if feedback should be delayed. Findings suggest that task feedback is most effective when it is given immediately (Clariana et al., 2000). This implies that students benefit most when they know whether or not they were successful at completing a task quickly after they attempted the task. This is especially true for the simpler tasks that do not require deep processing. For more complex problems, however, delayed feedback is preferred over immediate feedback (Clariana et al., 2000). By delaying feedback for complex tasks, students have the opportunity to process the problem and think independently for extended periods of time, which is more beneficial than intervening immediately. Timing of feedback is also dependent on the situation. For example, a meta-analysis determined immediate feedback is beneficial for classroom activities. For testing situations, however, delayed feedback is more effective (Kulik & Kulik, 1988). A more recent study compared how students performed on multiple-choice homework when given either immediate task feedback that indicated whether a question was correct or incorrect, or delayed feedback coupled with narrated

solutions. Students who received delayed feedback far outperformed students who received immediate feedback (Gladding et al., 2015). In summary, many of these studies are conflicting, although testing situations more often require task feedback, and classroom activities are associated with processing feedback. Nevertheless, appropriate timing of feedback is not entirely straightforward and is certainly context-dependent.

Misconceptions

In terms of teaching concepts, the need for the identification and engagement of student preconceptions and prior knowledge is widely recognized as an integral component of effective teaching, including in the National Research Council's study on how people learn (Bransford et al., 2000). According to this study, the interaction of prior knowledge with new material can result in three outcomes contingent upon the consistency of the two entities. First, if prior knowledge is consistent with new material, then the new material will be learned with ease. Otherwise, contradictions within this interaction, or misconceptions, will lead to confusion. The appearance of these misconceptions results in the two other possible outcomes. Resolution of these misconceptions will facilitate greater understanding of the material, while a failure to identify and challenge foundational misconceptions can lead to deeper misunderstanding and detriments to long-term learning as well as retention in engineering (Boud & Falchikov, 2006; Brown et al., 2009; McKeachie & Svinicki, 2013; Seymour & Hewitt, n.d.).

To address these misconceptions, there is a tremendous amount of research that has been conducted on the instruction of conceptual knowledge, as exemplified by the eclectic 8,400 articles documented in Duit's bibliography, "Science Teachers' and Students' Conceptions and Science Education" (Duit, n.d.). Amongst these articles, the seminal work of Hestenes, Wells, and Swackhamer in developing the Force Concept Inventory has greatly impacted engineering education (Hestenes et al., 1992). Their work created a test to assess and measure students' conceptual learning on force and Newton's laws of motion and has since inspired similar concept inventories in numerous subjects, including thermodynamics (Evans et al., 2003; Streveler et al., 2003). Based on the scores of these concept inventories, Hake noticed that interactive engagement of students as opposed to conventional lectures and homework assignments led to increased conceptual learning, as measured by the force concept inventory (Hake, 1998).

With this in mind, there has been a shift in strategies for teaching conceptual knowledge to more student-centered pedagogies that identify and correct misconceptions (Cakir, 2008; Longfield, 2009; Savion, 2009). Previous formal methods to identify misconceptions include the use of aforementioned concept inventories as method of assessment. However, the purpose and implementation of these concept inventories often determines the effect on student learning. For instance, if used at the end of a course, concept inventories can help instructors measure the effectiveness of their conceptual instruction to refine future iterations of the course, but these assessments would not directly affect the learning of the current students. If used at the beginning of the course, concept inventories can help measure students' grasp on prerequisite material and inform the foundational lectures of the course. Although these assessments are useful in highlighting more effective teaching methods, their impact on the student's ability to identify and engage their own misconceptions is more limited (Streveler et al., 2008).

In the shift to more student-centered pedagogies, frequent in-class assessments have gained prevalence (Mazur, 1999; 2013). In particular, immediate feedback assessment techniques, such as clickers and personal response systems, have gained widespread use for their amenability to large lecture-style classes (Sevian & Robinson, 2011). These techniques can be employed prior to and following each conceptual unit to help students identify their own prior misconceptions and gaps in knowledge post explicit instruction. The results can also help the instructor modify the course to help students learn the necessary material(Cakir, 2008). These immediate feedback techniques have already shown improvements in the correction of misconceptions and retention of conceptual knowledge (Epstein et al., 2002).

Similarly, "flipped" classroom approaches are becoming equally widespread. In these classrooms, students watch recorded lectures as homework and work on problem-based learning activities in class. In this way, instructors can see where misconceptions occur and immediately respond to correct these issues (Berrett, 2012). Both immediate feedback assessment techniques and flipped classroom approaches have made tremendous strides in empowering students as agents of their own learning and helping them identify and correct their own misconceptions. However, there is also merit in allotting students time to think through a concept and ensuring they reflect and act upon feedback to correct misconceptions.

Interventions in Thermodynamics

Regarding thermodynamics courses specifically there have been multiple studies of the paradox of students that receive high homework grades and low exam grades. Thermodynamics is a difficult, abstract subject that requires significant time input on the part of the student to master the material (Sokrat, 2013). National averages of passing rates for thermodynamics portion of the Fundamentals of Engineering professional licensing exam are between 40 and 65%, with minimal improvement in this area over the last ten years (Dukhan & Schumack, 2013). Some of the issues associated with poor performance on the exam stem from conceptual misunderstandings of the material. Many students can solve routine textbook problems yet provide incorrect answers to conceptual problems and are unable to sufficiently explain their answers (Dukhan, 2016). Some of these issues arise from historical misconceptions that lead to continued confusion in students today, while some stem from curricula that try to cover too much material in one semester (Sokrat, 2013).

Multiple studies have implemented online thermodynamics homework with rapid feedback or other more novel methods to try and combat these issues (Nasr & Ramadan, 2005; Taraban; 2005). Although these methods have met with some success the root issues still prevail. One educator who was implementing a homework system that focused on synthesis instead of memorization noticed that the students became frustrated when they were required to explain their answers (Nasr & Ramadan, 2005). After implementing homework that involves explanation and critical evaluation of problem solving and the resultant answer, another educator notes, "Most students respond favorably to this problem-solving approach. . . Of course, there still exists the typical student desire to have an example of every kind of problem that they will ever confront so that they can just change the numbers and crank out the answers" (Reardon, 2011). Even when students are given clear instructions to explain their answers they often avoid providing explanation and need to be repeatedly reminded to do so. Although some of these behaviors could stem from students being overwhelmed by the curriculum (Sokrat, 2011), it is likely that many of the issues arise rather from the climate of many undergraduate engineering programs and prevailing attitudes about homework that students develop in high school.

Three studies in particular focus on designing platforms to help students determine their misconceptions and synthesize course information in an engineering context. Armstrong (1995) implemented a novel "imperfect solutions" homework format in which students are provided answers to homework but some of them are secretly incorrect. The student needs to solve the problems and determine which solutions are incorrect and why. Although this particular study was primarily a case study, it shows strong student support for the method with only three of 67 saying it was not at least "good" on a scale from Excellent, Good, Average, and Poor scale, and with only 2 of 67 saying that they would not like it to be used in other courses. In general the students felt as if they were learning more while taking less time to complete the assignments.

Another homework method very similar to the one proposed here is Phase Arrayed Homework implemented by Godshall (2012). In this system, students prepare a problem before class, grade their own work in class as the answer is reviewed, and then have the opportunity to correct errors before the next class period to re-gain lost points. One time saving aspect of this homework system was student self-grading prior to submission of draft two, which could be implemented in this study if it is found that time is a major prohibitive factor. The Phase Arrayed Homework had multiple benefits, especially in student perception. Two of the main benefits were decreased stress associated with assignments and increased confidence in problem solving. However, improvement over control classes on exam scores was varied over class years, from no improvement to moderate (5-10%) improvement. One major result of this study was that since re-submission was optional, as the semester progressed and students became more overwhelmed by their coursework, the number of people who took advantage of the extra attempt decreased, mirroring loss in gains on the exam.

Another helpful homework system evaluated a mastery learning approach which aimed to make homework more of a formative assessment than a summative assessment (Moore & Ranalli, 2015). This technique allowed students to resubmit homework indefinitely after receiving a grade of 2 (mastered), 1 (not mastered), or 0 (not attempted). There was no grade penalty given for a not mastered problem unless the corrections were never made. This limited the amount of grading needed which helped to reduce the grading burden. It was noted that students almost always resubmitted, unless other more pressing assignments starting piling up, up to three attempts. Student homework grades increased and students felt like they learned more from this homework method, however, no correlations between homework and exams were made in the scope of this study.

Hypothesis and Scope of Study

We believe that the incorporation of conceptual drafts and intermediate feedback will motivate students to not only think conceptually but also to reflect upon their misconceptions. We also think that the removal of pressure in reaching the right quantitative answer on the conceptual draft will encourage students to engage and address their misconceptions instead of focusing on getting the correct answer. We also believe that this will help to decrease student work load/stress regarding their assignments and lead to enhanced learning.

Beyond this, our two-part homework system incorporates many beneficial features of formative and summative assessment, as well as effective feedback strategies that have been identified in past literature. We use a combination of formative assessment techniques (i.e. homework system) and summative assessment techniques (i.e. exams). The overarching goal is that students will gain strategic approaches that will lead to long-term benefits through formative assessment, and summative assessment will optimize short-term learning for students and provide feedback to the instructor to improve future performance of the students.

The primary questions that we will be investigating are: 1) Do students like the two-stage homework system?, 2) Does feedback improve student understanding?, and 3) Does the two-stage homework result in long term gains in understanding as evidenced by the exam?

In addition to these main questions, we will also be evaluating the time spent by the instructors and students for this course and compare it to other courses to determine if a combination of formative and summative assessment leads to excessive workloads. To try to better understand student response and behavior to this system, we will also evaluate self efficacy, student collaboration habits, and student study habits. Using these additional metrics, along with class and student observation, we will try to determine ways that we can improve upon the system to repeat the study for next year.

Methods

Participants

The participants in this study were 34 (four male and 30 female) students in a sophomore level Biomedical Engineering Thermodynamics course at Cornell University. The course was a mandatory course for all major students and was one of the first in-major courses that they took.

Homework Design

The two-stage homework system, and the goal behind it, was explained on the first day of class to the students along with a written explanation posted online. Students were given a copy of the rubric (Appendix A) and an example of full, partial, and no credit solutions to practice problems to use as a guide.

Student Performance Analysis

To assess student performance on Draft 1 and Draft 2 of each homework assignment and each exam, the TA (Earle) gave a number score between 0 and 2 that represented the student's understanding of the main concept of the problem. A score of 0 indicated that the student displayed no to minimal understanding of the concept being tested, a 1 meant that the student had a firm grasp of the concept, but was missing some key points, and a 2 represented complete conceptual understanding. With these numbers, we determined changes in understanding between Draft 1 and Draft 2 of the homework assignments, or relationships between understanding in homework and exams. Negative changes (i.e. "-2" and "-1") indicated a decline in conceptual understanding, or reversion to a previously held misconception. No change (i.e. "0") represented a student's understanding did not change between the two drafts. A positive change (i.e. "+1" and "+2") was indicative of an improvement in conceptual understanding.

Written Feedback Analysis

We analyzed the written feedback provided in homework assignments to determine if written feedback was effective in improving student performance and to determine which types of feedback were most effective. For the first three homework assignments of the semester, we transcribed all feedback given on Draft 1 for a diverse group of students (n=8-17 per assignment). We then categorized each instance of feedback into one of four categories: self, task, processing, or self-regulation. Next, we determined the change in student performance between Draft 1 and Draft 2 according to the scoring system previously mentioned. Relationships between feedback type and student performance were assessed.

Surveys

Likert scale surveys were given early in the semester (early March) and at the end of the semester (early May) to evaluate student opinion of the two-stage homework system, self efficacy, study habits, and collaboration habits (Appendix B and C respectively). The early semester survey was given after two homework assignments were completed but prior to the

first exam. The end of semester survey was given just prior to the final exam. It was combined with some general questions about the course as a whole. The self- efficacy questions were adapted from Bembenutty (2010) and Rowbautham and Schmidt (2013).

Focus Group

We held a focus group to gain personal insight from students on their opinions of the two-part homework system, feedback in the course, and work habits. Four students out of 34 participated in the focus group. All four were women. The discussion was guided by a focus group question guide (Appendix D) and led by two biomedical engineering graduate students who had no connections to the class. The session was recorded and transcribed. All responses remained anonymous to the instructors.

Results

Student Response to the Two-Stage Homework System

Overall, students had a positive experience with the two stage homework system, which agrees with previous studies on formative assessment based homework strategies (Godshall, 2012; Armstrong, 1995; Moore and Ranalli, 2015). In the early parts of the course, many students pushed back against the need to "justify or contextualize" their answers. This was the largest component of the conceptual draft grade and many students felt that it was "way less relevant than it seems" or felt that it wasn't necessary if they "already gave their assumptions." Seven out of the 34 students in the class specifically mentioned this key aspect of the homework assignment as something that they didn't like. This was the highest level of Bloom's taxonomy that was on the homework and required the students to familiarize themselves with the course content beyond just plugging known values into the equations from class. Like other studies, (Moore and Ranalli, 2015; Nasr and Ramadan, 2005) students needed to be reminded throughout the duration of the course to justify their answers and give a reason why their answers made sense in the context of the question and the class. However, more students found the system helpful than unhelpful early on, and the number of students who found it helpful increased as the semester progressed. This overall satisfaction and preference for this new stage of homework was most prominently shown in that 25 of 34 students recommended that this system would be used in their other classes and 12 out of those 25 strongly recommended its use in other classes.

Although the survey responses were informative, many of the student quotes from the surveys highlight the impact that the homework had on at least a subset of students. Some students really appreciated the ability to make a mistake without it penalizing their grade. One student said, "Makes me less afraid to make mistakes, so I don't mind answering questions wrong. Which is why I don't copy answers from online." And multiple students mentioned that reviewing their assignments specifically their mistakes helped them to learn, "the feedback points me in the right direction of thinking and helps a lot in my understanding of the homework," and "I think that the two-stage homework set up is such a great idea. After the second homework is submitted, I understand the concepts so much better than I did the first time around," and more specifically, "It forces me to make corrections and review my mistakes." Overall, these student quotes highlighted the positive impact that the students felt that the system had on their ability to learn the material.

One thing that was quite unanimous was that students were much less overwhelmed by the workload in this course than others. Based on the end of semester survey, 21 out of 34 students said that they were frequently overwhelmed by all of their coursework at a level at least 9/10. For this course in particular, no students said rated their themselves above an 8 (two students) and 17 out of 34 responded less than 5/10. These statements were further supported by the focus group. Students discussed how the two stage system helps them to organize their time and spread their work load out. Students also felt a confidence/grade boost from the system because the homework grades were high.

The focus group, which served as a more open-ended way to get feedback, illuminated some key ideas about how the students perceived the two-stage homework verses it's intention. It was originally stated that the first homework should have little to no math, however, almost none of the students actually treated the "conceptual" draft as conceptual. One student in the focus group said,

> "I think after the first homework, we didn't take it for conceptual learning. We more thought of this as a first draft of the homework. So maybe if the directions for the first draft were different? Like instead of doing everything, don't do any math, don't solve anything. Write down bullets for what concepts it is bridging. Like actual words instead of math and stuff."

What the student described as being more helpful, was the actual intention of the homework and what was described to the students both in class and in
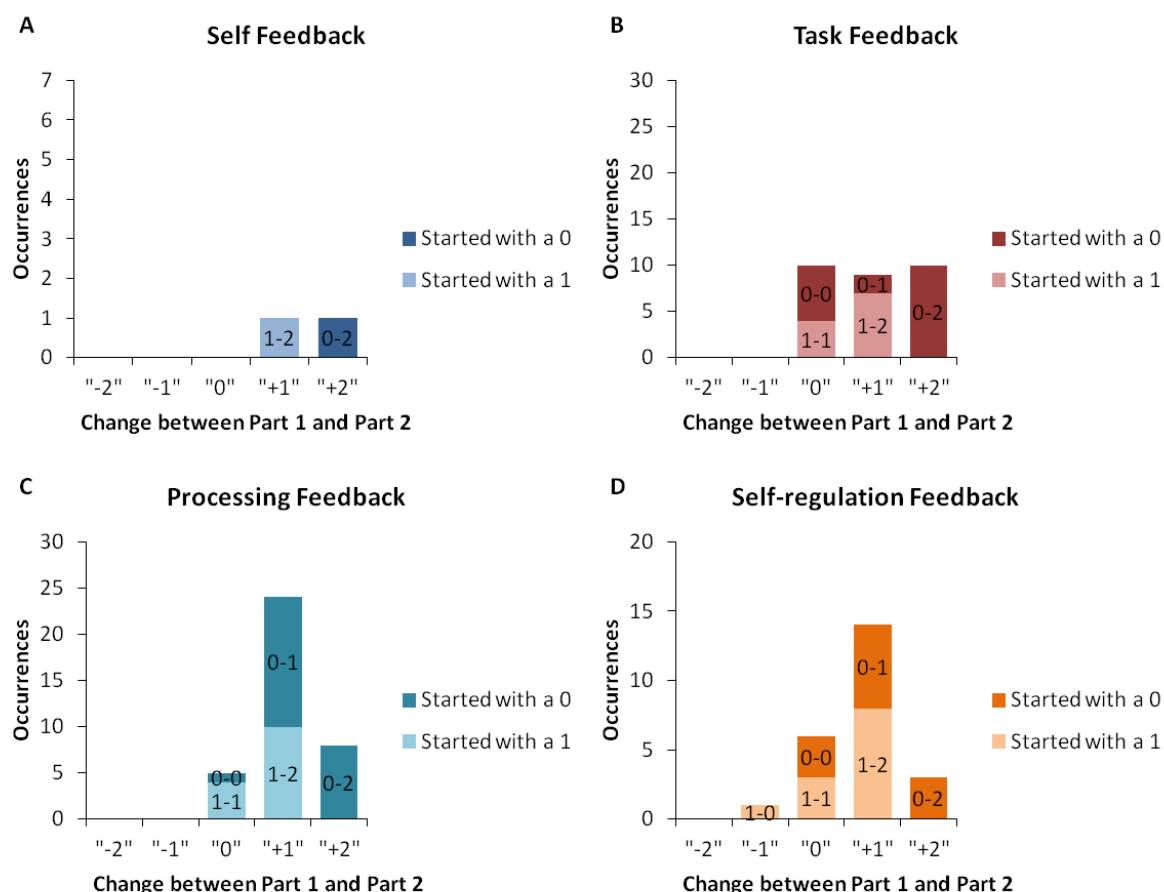
Figure 1: Effects of four feedback categories on student performance. A) Self feedback was very rarely the only feedback provided when the student lacked conceptual understanding in Draft 1. B) Task and D) self-regulation feedback improved student performance approximately 2/3 of the instances they were provided. C) Processing feedback improved student performance nearly 90% of the instances it was provided.

written documents. Another student expanded on this, saying,

> "So even in the student's mind, that wasn't clear. Ashley called it a conceptual draft the first class, and then we got our first homework, and we just thought of it as a practice round, we'll make stuff up if we don't know it, we'll put I don't know, we'll go to office hours Ashley, I promise. And then she'll check us off, and we'll be like, okay cool, and then we have a final draft. So I think it does increase understanding, but maybe not to the degree it wanted to."

What was interesting though, is that almost none of the students simply said I don't know on their homework. They at least discussed something related to the problems, or explained what they didn't know getting them to think about something. The focus group size was limited (only four students participated), so it may be that we are missing a significant amount of important feedback as to student perception.

Effect of Written Feedback on Student Performance

In the representative set of homework problems analyzed, self-feedback by itself was only given a total of 8 times. Task feedback was given 46 times, processing feedback was provided 51 times, and self-regulation was given 34 times. In our analysis, however, we focused on problems where students had room to improve (i.e. did not begin with a 2 in our scoring system). When only problems where students began with a 0 or 1, self-feedback was given twice, task 29 times, processing 37 times, and self-regulation 24 times.

To analyze the effectiveness of each type of feedback, we determined how student performance changed from Draft 1 to Draft 2 and grouped these changes based on feedback category. Self-feedback was so rarely the only form of feedback given (two occurrences) that we cannot provide any meaningful results regarding self-feedback (Figure 1A). Task feedback improved student performance 19 out of 29

occurrences, or 66% of the time (Figure 1B). Similarly, self-regulation feedback increased performance 17 out of 27 occurrences, or 63% (Figure 1D). Processing feedback improved student performance more than any other type of feedback, with increases in scores 32 out of 37 occurrences, or 86% of the time (Figure 1C).

Effect of Two-Stage Homework on Exam Performance

Using the previously described conceptual understanding grading system, differences in conceptual grades were found between homework drafts and each homework draft and the exam for multiple concepts. The differences across several concepts were averaged for each student for the first half of the year with homeworks 1-3 and Exam 1 (Figure 2A) as well as the second half of the year with homeworks 4-6 and Exam 2 (Supplemental Figure 1).

Within the first half of the year (Figure 2B), 33 of 34 students (97%) improved in conceptual understanding from homework draft 1 to draft 2, while the other student sustained their conceptual understanding. From homework draft 2 to exam 1, 32 out of 34 (94%) students sustained their increased conceptual understanding, and the other two students slightly reverted to their previous conceptual understanding. Nonetheless, overall, from homework draft 1 to exam 1, all students sustained (6 out of 34 or 18%) or improved (28 out of 24 or 82%) in their conceptual understanding.

We then evaluated whether this effect was present within the second half of the year (Figure 2C). From homework draft 1 to draft 2, 23 students improved (68%) and 11 students sustained (32%) their conceptual understanding. 26 of these students (76%) then sustained, 1 student improved (3%), and 7 students reverted (21%) their conceptual understanding from draft 2 to exam 2. This increase in reversion can be possibly due to an increasing stress load on students throughout the semester (see discussion). Still, from homework draft 1 to exam 2, 15 students improved (44%), 18 students sustained (53%), and 1 student (3%) reverted in their conceptual understanding. With this in mind, we believe that the two-stage homework system had a lasting impact on students' conceptual understanding.

Additional Factors in Conceptual Understanding

We also sought to evaluate other confounding factors outside of the two-stage homework system that could affect students' conceptual understanding. We used an end-of-semester Likert-scale based survey (see Appendix C: End of Semester Survey) with several questions graded on a scale from $1 - 10$ to study students' self-efficacy and study habits.

We first focused on self-efficacy, or belief in oneself, as it has been reported to be the main factor linking homework grade to exam grade. We primarily focused on student responses to the questions focused on their reaction to new material, their persistence, and their self-confidence. We then used these responses to create categories of students' self-efficacy (Table 1). We then found the average difference from homework draft 2 to the exam across several concepts for each student self-efficacy category.

As expected, students with a high self-efficacy had the highest increase in conceptual understanding. Interestingly, persistence, regardless of confidence level, also led to an increase in conceptual understanding (see categories O, F, and P).

However, confident students spanned a large spectrum (n=11) and did not rank as highly as expected (rank 5 of 6). We therefore further analyzed the study habits of these confident students (Table 2) by using the same end-of-semester Likert-scale based survey (see Appendix C: End of Semester Survey) based on questions related to distractions. We primarily focused on whether they had a preference for no distractions, how frequently they took breaks, and how successful they were in creating a distraction free environment. We found that confident students with good study habits (Category A in Table 2) showed the greatest improvement in conceptual understanding from home draft 2 to the exam.

Discussion

Overall student response and efficacy were quite encouraging regarding this system. However, there were some key limitations that bear discussion and require some changes to be made in the course before implementing again next year.

Student Perception and Limitations of the Two-Stage System

A major issue with the implementation of this system was that this was the first time teaching this course. There were additional hurdles in misunderstandings of students' prior knowledge, as well as some mistakes made during lecture due to unfamiliarity with the material that did not help the homework to reach its full potential. Aside from the naturally occurring benefit of next year being the second round of teaching, one thing that could specifically be done better in the future is to have fewer derivations during the lectures and more time spent talking about how to make an assumption, or evaluate the validity of that assumption. These steps will be implemented
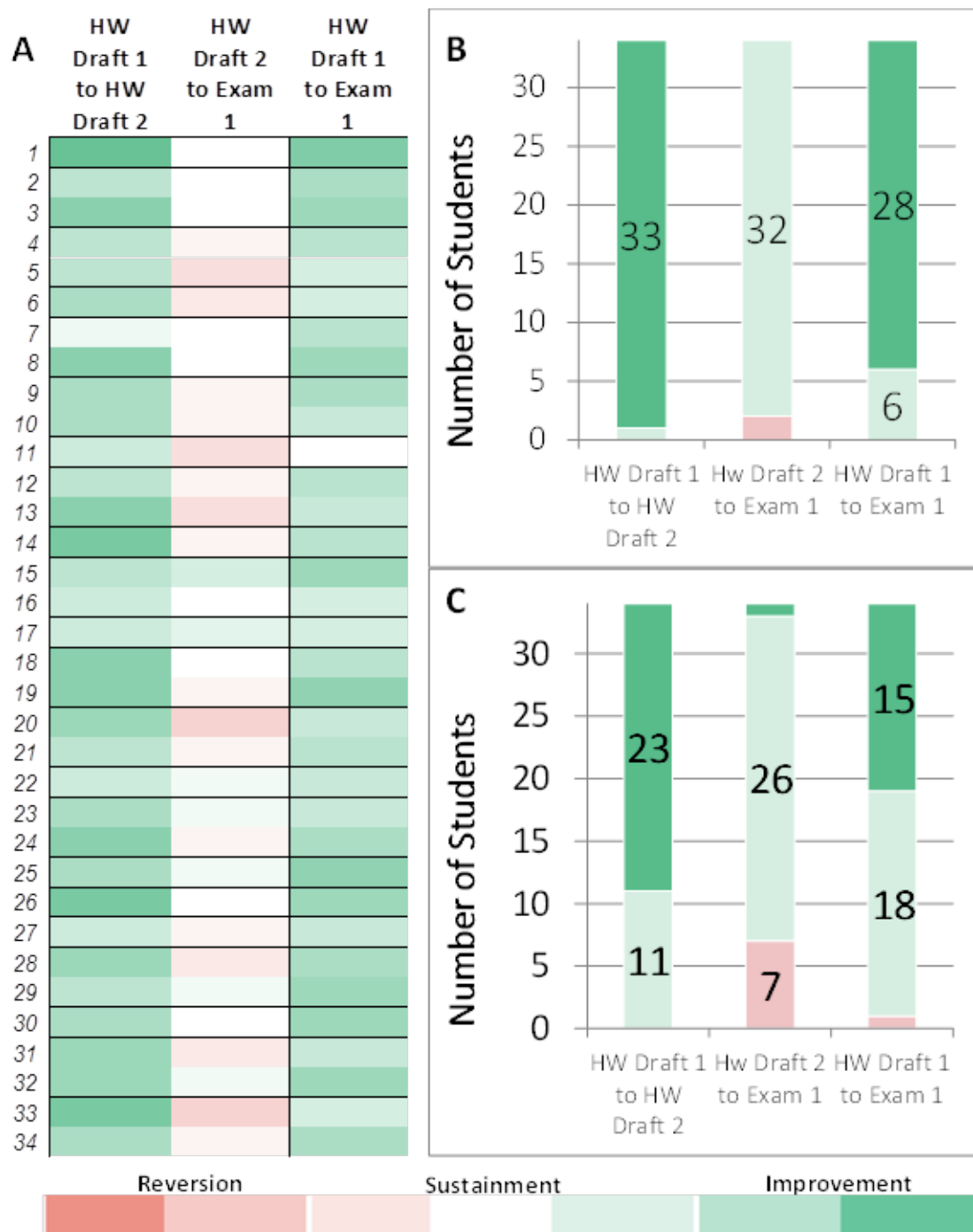
Figure 2. Effect of two-stage homework on conceptual understanding. A.) Average difference in conceptual understanding between specified assignments for multiple concepts for each student. B & C.) Total number of students who reverted, sustained, or improved conceptual within the first half (B) and second half of the year (C). Legend is included below figures.

Table 1: Categories of student's efficacy ranked in difference in conceptual understanding from homework draft 2 to the exam.

Categories	Definition	n	Ranked difference from HW Draft 2 to Exam
V: Very Confident	Average > 8	6	1
C: Confident	Average > 6.67	11	5
O: Overall Confident but Not Persistent	Believe > 7, Persistent < 5	2	6
F: Overall Confident but Bad First Impressions	Believe > 7, New Material < 5	5	4
P: Persistent but Not Confident	Believe < 5, Persistent > 7	2	2
N: Not Confident	Average < 6	6	3

Table 2: Categories of confident students' study habits associated with difference in conceptual understanding from homework draft to the exam.

Categories	Definition	HW Draft 2 to Exam
A: Prefers no distractions and succeeds	Creation > 7, Breaks < 3	0.083
B: Prefers no distraction and fails	Distractions > 7, Breaks > 4	-0.306
C: No Preference / Likes Distractions	Preference < 6	-0.333

next year to try and make the course self-consistent in its method. Beyond this, the whole major curriculum is also in its infancy stages. Different professors currently have different priorities which creates a slightly disjointed idea of what the outcomes of the program are. This manifests also in not having an entirely cohesive idea of what a student should have when they graduate. This has contributed to some courses being too difficult and causing significant stress to the students. In the end of the semester survey, 21/34 students reported being very overwhelmed (9 out of 10 scale) by their coursework. This no doubt negatively contributed to their performance as pressing, or more overwhelming deadlines in certain classes will take precedence over classes deemed as less overwhelming. One student even described it this way,

> "It's reached the point where we are asking our other professors if we can have extensions because we can't finish our work for Class A. It feels really guilty because we shouldn't prioritize one class over another. It feels wrong to ask another professor for an extension on things they don't control. It's not their fault. Every minute I'm not doing homework in one class; I'm doing work for Class A. So yeah this class doesn't really take up that much time."

These results are corroborated by Sokrat (2013) and Godshall (2012).

Beyond this, a larger issue that was alluded to above was that of cultural change. This class was a distinct cultural change from their other classes, at a time when they are already transitioning from lower to upper level course work. Sophomore year is when material typically becomes more conceptual and simply being able to plug and chug does not get you an A. The students are already in a cultural shock of their own as they start taking the classes that one day they will use for their jobs. This class, with the heavy conceptual focus of the homework and content, brought out a lot of that tension in the students as it forced this change into the light, and made them interact with it as part of their assignments. Like Watkins and Stevens (2013) found, the overall culture and expectation really matters. At a university like Cornell, students are generally competitive and grade focused. The homework system still seemed to have some impact, if not the intended one. As one student said,

> "I think it is helpful to have two drafts to any problem set. I think at this point with problem sets, people are just trying to get it done. You worry about understanding it the night before the prelim. That's the way

it works here. So if you have two weeks of it, it forces you try to understand it maybe. So I do think you get more than the average understanding."

Based on this feedback, I think this would be more effective if an entire department were to implement a system like this. That would provide the lasting conceptual grounding for students to develop the tool set of conceptual learning and then the culture shock would eventually become assimilation. A major, and less discussed part, of the reasons students' liked the system may be simply due to the fact that they felt like it boosted their grade. This may have influenced their self-reporting on learning because the demand for high grades is so prominent. Other researchers also found that the high homework grade was a factor in their student's enjoyment of the intervention (Godshall, 2012; Moore and Ranalli, 2005).

Another major point of discussion, although not significant in the feedback from the two surveys was student collaboration. Many students made extensive use of an online group chat, solutions manuals, and other supports. One student said in the focus group that, "The morning homework is due, it's frantic with people texting each other." Another student said, "Someone will say, "I went to office hours for ten minutes, and I got this," so they'll post that, and everyone else will be like "okay, great" and jumps on." This sharing of information, while it can be good for building cohesion in the class, could also negatively affect some students understanding of the material. This is very difficult, if not impossible to control, but it is important to keep in mind with regard to what the actual purpose of the homework is.

Limitations of Feedback Analysis

Our feedback analysis indicated that written feedback, in general, was effective in improving student performance. Furthermore, processing feedback was the most effective category of feedback compared to task and self-regulation, and self-feedback occurred too rarely to include in our analysis. Indeed, this is in agreement with previous literature (Hattie & Timperley, 2007) that deeper levels of feedback that requires students to process problems are more effective in improving performance. However, we must note that the written feedback analyzed was not the only form of feedback provided to students for these homework sets. First, attending office hours was most likely a major factor in predicting student performance on Draft 2. Second, the fact that the student had a second attempt at solving the same problem may have also played a role in the changes in scores between the two homework drafts. These two ideas correlate directly with our qualitative data from our focus group discussion with the students. Nonetheless, our analysis provides evidence that student performance improved with written feedback, and processing feedback was the most effective category of written feedback.

Limitations of Effectiveness Measures

We, as others, saw that homework is not a perfect predictor of exam performance, but similar to Bembenutty (2012), is modulated by other student specific factors such as self-efficacy, study habits, and collaboration with other students. One specific factor that we attempted to monitor was student's attitude toward seeking help, which was cited in their paper as being a prominent factor. Office hours were much busier than anticipated and student's extensive use of a class group chat made it more difficult to track their help seeking tendencies, this would be another factor that could help to understand the follow up results next year.

Beyond this, we have analyzed our data distinctly from other papers that have been published in evaluating homework and exam correlation. Instead of looking at exam grade versus homework grade, we concept scored both homework and exams and only evaluated the concepts that were directly linked back to specific homework problems. You could argue that this is removing the far-reaching effects that homework is supposed to have, but it also eliminates much of the noise in our data. It is possible, and noted for some students in our study, that they do very well on the concepts from homework, but show a drop off in understanding from problems that were focused on in class. We have not done any further correlation on conceptual understanding versus the overall course grade or even numerical exam grade. This may be an important factor to look into in the future.

Another important, and current limitation, is that the analysis of student performance is currently only based on the first exam. This was based purely on time constraints for this conference, but this analysis is on-going.

Implementation Limitations

We have yet to finish compiling our data on the amount of time that it takes for grading these versus more traditional types of homework assignments. Another factor is that it is difficult to be entirely consistent with giving feedback to students, and this increases as class size increases. For next year, we will explore using an online grading tool that should decrease the time and increase accu-

racy/consistency of commenting called Gradescope (`https://gradescope.com/`).

Another inherent limitation is the fact that only one person scored the understanding of the students on the 0-1-2 scale. In the future, we could have multiple people scoring the homeworks and exams to get a more objective idea of student understanding of concepts.

Lastly, and very importantly, this was the first time this course was presented. It was apparent that some mistakes were made in material presentation and in the overall organization that put road blocks in front of students in learning the material. Even more importantly, the homework was designed and implemented by the TA for the course, but the TA did not teach the majority of the course. This lead to differences in teaching styles and approaches that created some dissonance between the homework, lectures and exam layouts. In the future, we will need to work together more to present a cohesive picture to help the students fully grasp and jump on the conceptual learning band wagon.

Conclusions

Students seem to be positively affected by the two-stage homework. For the most part, students perceived that their learning was improved, and in an ideal world this would translate to enhanced interest in the subject matter. Beyond this, looking at the performance outcomes, students almost always showed increased understanding as a result of feedback between the conceptual and final drafts of the homework. This, for the most part, resulted in sustained understanding on the exams. This is particularly interesting in light of correcting misconceptions and long term learning. Although there were some limitations with the course, and lack of non-intervention control, we believe that the two-stage homework is worth pursuing to set the stage for a more concept focused engineering or STEM curriculum.

References

Ammons, RB. 1956. "Effects of Knowledge of Performance: A Survey and Tentative Theoretical Formulation." J Gen Psychol. 54(2): 279–99. doi:10.1080/00221309.1956.9920284.

Anderson, JR, AT Corbett, KR Koedinger, and R Pelletier. 1995. "Cognitive Tutors: Lessons Learned." J Learn Sci 4(2): 167–207. `doi:10.1207/s15327809jls0402_2`

Armstrong, B. 1995. "The Imperfect Solutions Homework Format." IEEE Transactions on Education.

Attwood, R. 2009. "Well, What Do You Know?" Times Higher Education.

Bas, G, C Senturk, and FM Cigerci. 2017. "Homework and Academic Achievement: A Meta-Analytic Review of Research." Issues in Educational Research.

Bembenutty, H, and MC White. 2013. "Academic Performance and Satisfaction with Homework Completion among College Students." Learn Individ Differ. 24: 83-88.

Berrett, D. 2012. "How 'flipping' the Classroom Can Improve the Traditional Lecture." The Education Digest 78(1): 36.

Boud, D, and N Falchikov. 2006. "Aligning Assessment with Long-Term Learning." Assess Eval High Educ. 31(4): 399–413.

Brown, G, J Bull, and M Pendleberry. 1997. "Assessing Student Learning in Higher Education."

Brown, MK, C Hershock, CJ Finelli, and C O'Neal. 2009. "Teaching for Retention in Science, Engineering, and Math Disciplines: A Guide for Faculty." Occasional Paper 25.

Butler, AC, JD Karpicke, and HL Roediger. 2007. "The Effect of Type and Timing of Feedback on Learning from Multiple-Choice Tests." J Exp Psychol Appl. 13(4): 273–81. doi:10.1037/1076-898X.13.4.273.

Cakir, M. 2008. "Constructivist Approaches to Learning in Science and Their Implications for Science Pedagogy: A Literature Review." Int J Env Sci Educ. 3(4): 193–206.

Clariana, RB, D Wagner, and LC Roher Murphy. 2000. "Applying a Connectionist Description of Feedback Timing." Educ Technol Res Dev 48(3): 5–22. doi:10.1007/BF02319855.

Clark, R, and J Andrews. 2014. "Relationships, Variety & Synergy: The Vital Ingredients for Scholarship in Engineering Education? A Case Study." Eur J Eng Educ.

Covic, T, and MK Jones. 2008. "Is the Essay Resubmission Option a Formative or a Summative Assessment and Does It Matter as Long as the Grades Improve?" Assess Eval High Educ. 33(1): 75–85. doi:10.1080/02602930601122928.

Dion, JS, and G Restrepo. 2016. "A Systematic Review of the Literature Linking Neural Correlates of Feedback Processing to Learning." Journal of

Psychology 224 (4): 247–56. doi:10.1027/2151-2604/a000260.

Dukhan, N, and M Schumack. 2013. "Understanding the Continued Poor Performance in Thermodynamics as a First Step toward an Instructional Strategy." ASEE Annual Conference.

Epstein, ML, AD Lazarus, TB Calvano, KA Matthews, et al. 2002. "Immediate Feedback Assessment Technique Promotes Learning and Corrects Inaccurate First Responses." Psychol Rec. 52(2): 187.

Ernst, B, and M Steinhauser. 2015. "Effects of Invalid Feedback on Learning and Feedback-Related Brain Activity in Decision-Making." Brain Cogn. 99: 78–86. doi:10.1016/j.bandc.2015.07.006.

Evans, DL, GL Gray, S Krause, J Martin, C Midkiff, BM Notaros, M Pavelich, et al. 2003. "Progress on Concept Inventory Assessment Tools." Front Educ. 1:T4G–1.

Fernandez, A. 2006. "2006-67: HOMEWORK AS AN OUTCOME ASSESSMENT: RELATIONSHIPS BETWEEN HOMEWORK AND TEST PERFORMANCE." Age.

Fleischmann, C, E Nakagawa, and T Kelley. 2016. "Redesigning the Student Learning Approach through Personality Types and Pedagogies, A Case Study in an Undergraduate Engineering Course." Universal Journal of Educational. http://eric.ed.gov/?id=EJ1089682.

Froyd, JE, PC Wankat, and KA Smith. 2012. "Five Major Shifts in 100 Years of Engineering Education." Proceedings of the IEEE 100 (SPL CONTENT): 1344–60. doi:10.1109/JPROC.2012.2190167.

Gladding, G, B Gutmann, N Schroeder, and T Stelzer. 2015. "Clinical Study of Student Learning Using Mastery Style versus Immediate Feedback Online Activities." Phys Rev Phys Educ Res. 11 (1): 1–8. doi:10.1103/PhysRevSTPER.11.010114.

Godshall, S. 2017. "Phased-Array Homework: Used to Shape and Steer Student Understanding." Usma.edu. http://www.usma.edu/cfe/Literature/Godshall_11.pdf.

Grammer, JK, M Carrasco, WJ Gehring, and FJ Morrison. 2014. "Age-Related Changes in Error Processing in Young Children: A School-Based Investigation." Dev Cogn Neurosci. 9: 93–105. doi:10.1016/j.dcn.2014.02.001.

Hake, RR. 1998. "Interactive-Engagement versus Traditional Methods: A Six-Thousand-Student Survey of Mechanics Test Data for Introductory Physics Courses." Am J Phys. 66(1). AAPT: 64–74.

Hattie, J, and H Timperley. 2007. "The Power of Feedback." Rev of Educ Res. 77(1): 16–17. doi:10.3102/003465430298487.

Hestenes, D, M Wells, and G Swackhamer. 1992. "Force Concept Inventory." The Physics Teacher 30(3). AAPT: 141–58.

Jenkins, JO. 2010. "A Multi-faceted Formative Assessment Approach: Better Recognising the Learning Needs of Students." Assess Eval High Educ. 35(5): 565–76. doi:10.1080/02602930903243059.

Jonassen, D, J Strobel, and CB Lee. 2006. "Everyday Problem Solving in Engineering." J Eng Educ. 9(2): 139-51. doi:10.1002/j.2168-9830.2006.tb00885.x.

Kluger, AN, and A DeNisi. 1996. "The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory." Psychol Bull. 119(2): 254–84. doi:10.1037/0033-2909.119.2.254.

Korpela, A, T Tarhasaari, and L Kettunen. 2016. "Structural Development of Substance in Engineering Education: Method of Cornerstones." International Conference. http://link.springer.com/chapter/10.1007/978-3-319-50337-0_54

Kulik, JA, and CLC Kulik. 1988. "Timing of Feedback and Verbal Learning." Rev Educ Res. 58(1): 79–97. doi:10.3102/00346543058001079.

Lafuente Martínez, M, IM Álvarez Valdivia, and A Remesal Ortiz. 2015. "Making Learning More Visible through E-Assessment: Implications for Feedback." J Comp High Educ. 27(1): 10–27. doi:10.1007/s12528-015-9091-8.

Litzinger, T, LR Lattuca, R Hadgraft, and W Newstetter. 2011. "Engineering Education and." J Eng Educ. 100(1): 123–50. doi:10.1002/j.2168-9830.2011.tb00006.x.

Locke, EA, and GP Latham. 1990. "Work Motivation and Satisfaction: Light at the End of the Tunnel." Psychol Sci. 1(4): 240–46. doi:10.1111/j.1467-9280.1990.tb00207.x.

Longfield, J. 2009. "Discrepant Teaching Events: Using an Inquiry Stance to Address Students' Misconceptions." International Journal of Teaching and Learning in Higher Education 21(2): 266.

Mazur, E. 1999. "Peer Instruction: A User's Manual." AAPT.

Mazur, E. 2013. "Peer Instruction."

McKeachie, W, and M Svinicki. 2013. McKeachie's Teaching Tips. Cengage Learning.

Moore, J, and J Ranalli. 2015. "A Mastery Learning Approach to Engineering Homework Assignments." 122nd ASEE Annual Conference, Seattle, WA. http://www.personal.psu.edu/jpm46/pubs/ASEE_2015a.pdf

Nasr, KJ, and B Ramadan. 2005. "Implementation of Problem–Based Learning into Engineering Thermodynamics." Age. https://peer.asee.org/implementation-of-problem-based-learning-into-engineering-thermodynamics.pdf.

Nicol, DJ, and D Macfarlane-Dick. 2006. "Formative Assessment and Self-regulated Learning: A Model and Seven Principles of Good Feedback Practice." Stud High Educ. 31(2): 199–218. doi:10.1080/03075070600572090.

Olds, BM, RA Streveler, and RL Miller. 2004. "Preliminary Results from the Development of a Concept Inventory in Thermal and Transport Science." Proceedings of the 2004 ASEE Annual Conference & Exposition, no. 703.

Reardon, FH. 2001. "Developing Problem-Solving Skills in Thermodynamics Courses." Age.

Rittle-Johnson, B, and RS Siegler. 2001. "Developing Conceptual Understanding and Procedural Skill in Mathematics: An Iterative Process." J Educ Psychol. 93(2): 346–62.

Rowbotham, M, and GS Schmitz. 2013. "Development and Validation of a Student Self-Efficacy Scale." J. Nurs. Care.

Sadler, DR. 1989. "Formative Assessment and the Design of Instructional Systems." Instr Sci. 18(2): 119–44. doi:10.1007/BF00117714.

Sadler, DR. 1998. "Formative Assessment: Revisiting the Territory." Assess in Educ Princ Pol Pract. 5(1): 77–84. doi:10.1080/0969595980050104.

Sailer, U, S Robinson, FPS Fischmeister, E Moser, I Kryspin-Exner, and H Bauer. 2007. "Imaging the Changing Role of Feedback during Learning in Decision-Making." NeuroImage 37(4): 1474–86. doi:10.1016/j.neuroimage.2007.07.012.

Savion, L. 2009. "Clinging to Discredited Beliefs: The Larger Cognitive Story." Journal of the Scholarship of Teaching and Learning 9(1): 81–92.

Sevian, H, and WE Robinson. 2011. "Clickers Promote Learning in All Kinds of Classes-Small and Large, Graduate and Undergraduate, Lecture and Lab." J Coll Sci Teach. 40(3): 14.

Seymour, E, and N Hewitt. 1997. "Talking about Leaving: Why Undergraduates Leave the Sciences, 1997." Boulder, CO: Westview.

Sokrat, H, S Tamani, M Moutaabbid, and M Radid. 2014. "Difficulties of Students from the Faculty of Science with Regard to Understanding the Concepts of Chemical Thermodynamics." Procedia-Social and.

Streveler, RA, TA Litzinger, RL Miller, and PS Steif. 2008. "Learning Conceptual Knowledge in the Engineering Sciences: Overview and Future Research Directions." J Eng Educ. 97(3): 279–94. doi:10.1002/j.2168-9830.2008.tb00979.x.

Streveler, RA, BM Olds, RL Miller, and MA Nelson. 2003. "Using a Delphi Study to Identify the Most Difficult Concepts for Students to Master in Thermal and Transport Science." In ASEE Proceedings.

Taraban, R, EE Anderson, and MW Hayes. 2005. "Developing On-line Homework for Introductory Thermodynamics." J Eng Educ. doi:10.1002/j.2168-9830.2005.tb00859.x/full.

Wankat, PC. 2001. "The Role of Homework." Age. https://peer.asee.org/9753.pdf.

Watkins, PJ, and DW Stevens. 2013. "The Goldilocks Dilemma: Homework Policy Creating a Culture Where Simply Good Is Just Not Good Enough." The Clearing House: A Journal of Educational Strategies, Issues, and Ideas. http://www.tandfonline.com/doi/abs/10.1080/00098655.2012.748642.

Figure SA: Effect of two-stage homework on conceptual understanding: Average difference in conceptual understanding between specified assignments for multiple concepts for each student within the second half of the year.

Appendix A: Grading Rubric for Conceptual Draft of Homework

Criteria	Mastering	Developing	Emerging
Problem Set Up	All known information is listed, a diagram is appropriately drawn and labeled and units are converted (when necessary). Unknown clearly specified.	All known information is listed, but either unit conversions are missing or diagram and labeling is missing or incorrect. Unknowns may be incorrectly determined.	Only some known information is listing with limited/no unit conversions or diagram with labeling
Assumptions and Problem Logic	All assumptions are listed and correct leading to identification of the appropriate equation.	Incorrect assumptions lead to incorrect equation.	Assumptions, and therefore problem logic, are missing. No way to determine conceptual grasp of problem.
Problem Solving Strategy and Skills	Student is able to link multiple concepts and multiple equations together to arrive at answer.	Student is able to partially develop a problem solving strategy, but at least one key concept is missing that prevents the student from completing the problem.	Multiple key concepts and/or relationships are missing such that the student is only able to begin the problem.
Organization and Neatness	Problem is written in a logical, step wise fashion that is easy to read and follow.	All problem components are present, but are disorganized, cramped, or otherwise difficult to read.	Problem is entirely disorganized, or illegible such that even if present, the other categories are impossible to score.
Critical Thinking	Student has provided a sound written and conceptual logic for feasibility or infeasibility of answer (or approach if problem was incomplete).	Student was able to determine whether an answer was feasible or not, but logic as to why is missing (or approach if problem was incomplete).	Student has not demonstrated that they evaluated the feasibility of their answer (or approach if problem was incomplete).

Appendix B: Early Semester Survey

Questions to be answered on a 1-5 scale.

• Pace of lectures. Do you find the pace of the lectures too slow (1), too fast (5), or about adequate (3)?

• Lecture organization. The lectures are well organized, and easy to follow.
Completely agree (5); completely disagree (1).

• Lecture style. Is the balance of PowerPoint slides to work on the blackboard adequate (3), or too board heavy (5), or to PowerPoint heavy (1)?

• Lecture outcome. Do you feel like you understand the "take home point" of the lectures, and achieve the learning outcomes?
Fully understand (5); completely confused (1).

• Friday Recitations. The Friday "recitation" style lectures are helpful.
Complete agree (5), not helpful at all (1).

• Homework Feedback. The homework feedback helps me learn and reinforce the material.
Completely agree (5), completely disagree (1).

• Homework Design. The two-stage homework helps me realize and correct my misconceptions.
Completely agree (5), completely disagree (1).

Open-ended questions (free text response):

• What is something you would like to keep the same in the course moving forward?

• What is something you would like to change in the course moving forward?

• What is something you would like to keep the same about the homework moving forward?

• What is something you would like to change about the homework moving forward?

• Do you have any other comments or suggestions for the course?

Appendix C: End of Semester Survey

How well do the following do the statements describe you, 1 being not at all and 10 being very well:

Homework Habits

1. I never work with other students on homework.

2. I disable notifications and avoid going on social media while I'm working.

3. When I don't know how to do a problem, I usually google the question to try and get help.

4. When I get stuck, I go on our class group chat to see what other students are doing.

5. I frequently answer peers questions on our group chat.

6. I like to do homework in a distraction free environment.

7. When I encounter new material I am confident that I can learn it.

8. I usually have facebook and other social media sites open while I'm working and take frequent breaks.

9. When I struggle to understand a concept, I find reading the textbook helpful.

10. I believe I can stay motivated, even when problems are difficult.

11. Do you have any specific homework habits that you find particularly helpful in learning from homework?

Homework Process

1. I feel like I make slow progress and then all of a sudden how to do the question "clicks"

2. I usually know how to do the problems when after I sit down and look through my notes.

3. When I find a new idea or application in the homework, I am usually able to connect it back to what I already know from this or other classes.

Response to Homework

1. I frequently feel overwhelmed by homework in this class

2. I frequently feel overwhelmed by the collective work that I have to do for all of my classes.

3. I spend much more time working on this class than I do my other classes.

4. I prefer the 2-stage homework system to normal homework.

5. I would recommend that other classes implement a similar homework system.

6. I feel like the 2-stage homework helps me learn.

7. I believe that I can learn even the most difficult material if I try hard enough.

8. Do you have any other particular comments on how the 2-stage homework impacts your learning?

Appendix D: Focus Group Questions

Background 1. Have you enjoyed the course?

2. How much do you feel you learned in the course?

Two-part homework system 1. What are your opinions of the two-part homework system?

2. Tell me about your positive experiences with the system.

3. Explain how your experiences differed from courses with a more traditional HW system (e.g. learning, course grade, time spent, etc.).

4. How could the instructors make the need to justify answers more clear?

Feedback 1. Were you more or less likely to pay attention to feedback on assignments?

2. Overall, how much do you like getting written feedback?

3. Would you have preferred a checkmark grade for the attempt of the first draft, and a numerical grade for the second draft?

Collaborations and resources 1. Did you use the solutions manual or other online sources for help on homework assignments?

2. How often did you collaborate on homework assignments?

3. How often did you use the online chat group?

Closing Remarks If you were teaching the course, what would you change about the homework system?

Shaping the Study Abroad Mind

Katryn Evinson

Department of Romance Studies, Cornell University
kw475@cornell.edu

Introduction

We might agree that, "What are the Humanities for?" and "What can you do with a degree in Humanities?" are questions that in recent decades have worked to undermine fields such as History, Literature and Anthropology insofar as these fields don't provide a clear professional path. At a time when Capitalism has attempted to render any little aspect of society as instrumental to profit, skills students learn by, for example, reading and analyzing seventeenth century memoirs of some forgotten poet, aren't directly transferrable to market needs. That is not to say that culture has experienced a lack of attention, on the contrary, different fields outside the Humanities see its clear value especially when it comes to business and finance (Nair-Venugopal, 2015). Humanistic approaches have been pushed aside arguably because of their potential to train critical and politicized students who won't sit well with market needs. In order to be a competitive professional in a global world, having the ability to work with different cultures, in a variety of international environments, showing a high degree of intercultural competence, only makes you a stronger candidate. Thus, culture has proven to be useful for Capital. Useful here is a key term, in that instrumentality has become central to how we think and relate to the world. In this context, study abroad has emerged as a great asset to students who are in search of improving their profile for the future job market. Having experience in a foreign country and speaking a foreign language has stood out as being an important asset which employers value in hiring processes. While it is true that many studies report that employers are unsure about how to rate study abroad when viewing their candidates' résumés (Herder et al., 2015), they do value international experiences and second languages, components in which study abroad programs are specialized. But study abroad is not only an increasingly valuable academic option for students' future professional careers, it is also an industry in and of itself. Study abroad providers have proliferated quickly in the past twenty years[1] and universities and colleges have rushed to develop their own programs and build their own study abroad divisions as they have seen a great source of potential income.

In this general context, language teaching at universities has not only become crucial for undergraduate education and preparing students for a competitive job market, but I would also argue that language teaching has been an indirect way to promote study abroad in the classroom. By making second language acquisition a requirement, study abroad programs find a majority of potential clients gathered together and easily targetable. In fact, at Cornell University where I have taught Spanish for four semesters, study abroad program ambassadors routinely visit Spanish language courses once or twice per semester to promote Cornell Abroad. That is not to say that study abroad is not an enriching opportunity beyond all professional applications, however, I am interested in reflecting on how financial goals belonging to study abroad as an industry can have an effect on how culture is being dealt with in second language acquisition courses at college level.

Culture in foreign language courses is often instrumentalized for language learning purposes, which means that culture constitutes a vehicle for language acquisition instead of a theme in and of itself. It is a common struggle to integrate culture into language curriculum especially when it is an "add-on" instead of an essential part of language learning. Such an approach to incorporating culture into language courses is premised on language levels standards demanding a certain competency in a target language whose only parameter is language acquisition, not cultural competency. Proof of this can be gathered in how exams and other assignments are designed, where cultural content is either inexistent or has very low weight in rubrics we are encouraged to use. Teachers are faced with having to squeeze in time for cultural knowledge into courses that are typically 45 hours per semester, on average. An amount of time that is already scarce to achieve language learning goals. At the same time, cultural knowledge is often barely considered in a given course's global learning objectives. It is no surprise that under these conditions, cultural analysis is sacrificed when time becomes an issue.

This being said, it seems unfair to require language

---

[1] http://www.chronicle.com.proxy.library.cornell.edu/article/Study-Abroad-Revenue-Drain-or/124850/

teachers, who are commonly linguists or philologists, to add "culture" to their syllabus when cultural analysis is not their field of expertise. Therefore, language teachers end up relying on their own cultural references and interests, if not only using those materials found in textbooks. At the same time, textbooks are frequently a difficult tool to work with regarding cultural insights. If it is already challenging to find well designed textbooks that address grammar in an adequate way, it is all the more difficult to agree with the way in which textbooks portray cultural materials. A variety of factors make it so culture becomes both a sort of burden for language teaching and extremely important, not only to Humanities, but also to develop better job candidates. The result is that culture is presented in "bits and bytes", easily digestible capsules that can be integrated in a classroom and effortlessly grasped by students. Such an approach to cultural analysis leaves students to think foreign cultures are uncomplicated. The idea that target culture can be understood through, for example, a short story by Gabriel García Márquez, becomes the chief logic in language learning and teaching. It gets worse when short stories are then combined with stereotypical family models and then completed with a twist: a virgin island just off Puerto Rico where you can go spend your vacations. Then, have language textbooks become mere vacation catalogues where you can choose from a flamenco experience, "El día de los muertos", and a stroll through an article on Buenos Aires' neighborhood Palermo? Are these digestible culture clips shaping the way in which we interact with foreign cultures? Are we preparing students to arrive to a city like Madrid to be closed to the complexities and contradictions that inform all cultures?

This particular kind of problem hasn't emerged today, it is part of a long tradition that has organized representations of the world and the other for centuries. If colonialism was a reality that justified its means of exploitation and oppression through picturing the other as non-human, as Edward Said's famous Orientalism examines (Said 1976, p 41) Nair-Venugopal (2015) argues that postcolonial theory is concerned with "identifying and challenging the legacies of colonialism in contemporary institutions and practices, such as tourism" (193) and therefore study abroad is not free from these legacies as well. So how are these forms of presenting culture nurturing ancient mindsets stemming from a colonialist background? Furthermore, if capitalism is considered to have thrived due to the colonial era (i.e. chattel slavery), how does a financial interest in study abroad feed into the cycle of representations of the other

as less human, and therefore less complex? In other words, if study abroad as an industry has the potential to generate significant profits based on presenting foreign cultures as enriching life-experience packages that can be easily consumed in a very convenient amount of time (i.e., one semester), what are the chances that those representations are either not questioning engrained beliefs inherited from our colonial mindset, or that they may even be consciously reinforcing them for the sake of business?

A Personal Trajectory

Before embarking in a PhD program I spent six years working in international education in Barcelona, Spain, for a study abroad third-party provider working with US and Canadian universities. I held positions both as an administrator and as an instructor for a Service Learning class during four semesters. Therefore, I had a chance to develop close relationships with students as their instructor, as well as working behind the scenes developing programs and linking courses with opportunities in the city, as an experiential learning coordinator and an academic advisor. Throughout my time there I learned about different challenges study abroad programs faced. The first and foremost concurrent challenge was student profile. No matter who you talked to, there was a shared perception and concern about students arriving with a very particular mindset that I can break down into four main traits: (1) a disinterest in integrating with local culture, which translated into not frequenting local bars, cafes and clubs but rather attending parties exclusively planned for North American students; (2) travelling profusely around Europe and Spain was a priority to see as many sites as possible, which could be read as cultural experiences being partially equated to doing tourism in order to visit their European heritage; (3) disregard towards educational and academic components of the program, which generally meant not showing interest in target language courses and/or culture classes on target culture; and (4) students arriving with an expectation that study abroad had to be their most exciting and positively intense experience they had ever had, adding a considerable amount of pressure that often meant their experience had to be perceived by others as successful and life-changing.

I must clarify that during my time working in this sector, we never surveyed students on these matters and we did not carry out any other type of study that would provide statistical information about these traits. However, in my time working with students in this capacity, not only did I witness these patterns being repeated semester after semester, but these

were perceptions shared among staff, among faculty, and across study abroad programs in Spain to the point that they determined our approach to improving our program. Therefore, our efforts devoted to bettering our programs revolved around increasing students' awareness of their expectations by designing mandatory workshops throughout the semester, encouraging students to travel less around the country and Europe by programming cultural activities in Barcelona and outskirts during weekends that would incentivize them to stay in the region, increasing experiential learning components of courses to make them more engaging, hoping to increase students' interest, and lastly, designing workshops for faculty and staff to provide tools to address these particular struggles. All in all, these constant attempts to improve the program were fundamentally addressing the four main challenges we faced once students arrived on site.

Now that I have experience being on the other side with respect to how we prepare students to go abroad, I have identified that (beyond societal elements which I don't want to understate, as they play a significant role in college students' disposition towards foreign cultures) language teaching and culture in home universities foster the previously described traits commonly seen in study abroad students very early on. One could argue that from a broader perspective, an approach to foreign cultures that renders them uncomplicated and unproblematic, to potentiate their interest as sites for consumer tourism, creates an increase of clients enrolling in study abroad programs.

Literature Review

Much research has been done on the ways in which culture is taught in language classes in order to uncover and scrutinize the ideological underpinnings that are subtly conveyed to learners under the pretext of language learning. In recent decades, there has been an increasing consensus in seeing second language acquisition as key in developing students' intercultural competence in a globalized world. A majority of the studies done on this matter agree that textbooks are one of the most crucial devices through which culture can be delivered and thus, a wide literature focuses on textbook discourse within language teaching and learning alone, establishing a new subfield of research. However, more sophisticated analysis underscores the fallacy of such an argument to understand how culture functions (Kramsch, 2013; Weninger, Csila & Tamas Kiss, 2013). Many argue that culture cannot be fixed in books since it is not a fixed component itself, but rather a process, dynamic and alive, that takes place in language courses

through the interaction of teachers, learners, and materials. Perhaps the notion of "classroom ecology" that Guerrettaz and Johnston (2013) suggest in their study on classroom materials and the ways in which students interact with them may be useful to think of a setting where this encounter with the other takes place first through the materials themselves, as alive, ecological. The definition they draw from describes an ecology as "relations of possibility between animals and their environment" (782), which they take from Ulric Neisser's (1987) explanation. They advance an idea of classroom as a site where participants and materials interact, as an ecology with its own dynamics and specificities.

Regarding the role of textbooks, in "(Re)Searching Culture in Foreign Language, Textbooks, or Politics of Hide and Seek," Germán Canale (2016) takes us through an analysis of fundamental assumptions that many textbook readers don't question when interacting with a textbook as it connects with politics of representation. To this regard, Canale falls on the side of those scholars who find textbook discourse analysis insufficient to provide a complex understanding on how learners are exposed to culture and produce their own meanings of it. Therefore, he stresses the textbook reader's role not as mere receptor of "static" information located in a textbook, but rather, he highlights the learner's active part in producing meaning that shapes the way in which foreign culture is perceived. In Canale's words "textbook readers are often unaware that textbooks don't contain mere facts, but rather that they are products of political, economic and cultural activities, however readers are not passive receptors" (227). In his study, Canale makes the choice of not limiting his claims to a specific language since he sees a drawback in this approach so prevalent in most of the literature on the topic. Instead, he decides to analyze textbooks designed for different languages in order to locate trends, patterns and constants tracing how language learning produces a particular understanding of foreign culture that, in turn, shapes communities' identities associated with the target language. Canale understands the textbook in the same waythat Kramsch (1987; 1998) does, as an artifact that "imposes social behavior and attitudes students are expected to agree with, and at the same time it creates a social reality for students to make sense of in a highly guided semiosis" (98). Canale defends that quantitative analysis does not provide a good method to delve into the tensions, nuances, and processes under which those norms are performed or not, and the level to which they achieve their ideological agendas. In order to complete his study, Canale reviews nine studies of cultural representation in text-

books, carefully selected according to specific criteria related to politics of inclusion or exclusion he identifies in teaching materials. His overarching claim is to say that even those exposures that try to familiarize students with crucial components of target language cultures frequently fall into the trap of reinforcing exclusion, and they do it in two general ways. The first understands that simplification of culture in class, although helpful to make culture accessible and attractive to students in ways that combine well with grammar learning, may in fact serve politics of exclusion by presenting the foreign culture as unproblematic. Simplification, then, brings us to the second aspect, in that it tends to rely on homogenization, thus looking to suppress those elements within a culture that don't fit into a formulation of an identity, free of difference. For instance, Tajeddin and Teimournezhad (2015) provide a thorough analysis of a strategy to do this is by presenting culture as neutral. They define "neutral" as not being associated with any particular region or cultural context within Iran for their study on Farsi language teaching and learning erasing those particularities, mostly regional, that define Persian cultures. They argue that to the extent that Persian culture is presented as unified and homogenous, it is actually unreal, inexistent, since no Iranian citizen experiences that "neutral" culture. Moreover, this strategy to present culture as homogenous and therefore, easily understandable, can have a very negative effect on students. Canale implicitly suggests that homogeneity may have an effect on learners' own identities for they will presuppose a norm of culture that should be applicable to cultures in general. Furthermore, such homogeneity is also in line with what is known as Perspectives, Practices, and Products (PPP) insofar as this approach sees culture as a cohesive element that is experienced in the same way by all the members of a particular community. Hence, even if learners are being confronted with other cultures, they are not being challenged to think about difference within those cultures, which then feeds into an understanding of their own culture as monolithic. For these reasons, Canale urges teachers to design classroom environments in which learners can develop a critical mindset towards the materials that will thus inform the type of interaction.

The objectification of culture cultivates and enhances beliefs in stereotypes, negatively affecting both the community's understanding of identities and their capacity to imagine other identities. Following Canale's critique of the tendency towards objectification, we find Weninger and Kiss' article "Culture in English as a Foreign Language (EFL) Textbooks: A Semiotic Approach" an essential contribution to the debate. In this paper, the authors not only challenge monolithic understanding of cultures but also highlight a normative trend that has consolidated in recent years: language learning must contribute to the development of critical and reflexive individuals that can grow to be highly competent in intercultural communication. The main critique they present is centered on revealing the challenge of exposing students to cultural realities when those are instrumentalized by language learning, specifically grammar and vocabulary teaching. If culture is only a pretext to grammar learning and we can agree that culture is a complex process of meaning-making that requires its own space for reflection, interiorization, and even questioning, then how will learners ever encounter a genuine cultural reflection when grammar is the focus and goal of the course? Weninger, Csila, and Tamas Kiss (2013) say "culture is used for grammar purposes and not to harness a critical and reflexive understanding to culture, the self and the other" (696). As an example of this unproblematic approach to culture, we see that in Musumeci and Aski's (2010) study of culture in Italian language instruction is thought of as a literal instrument to language learning but not as a complex dimension in and of itself.

On the contrary, Weninger, Csila, and Tamas Kiss (2013), believe that drawing from semiotic theory (eg., Charles Sanders Peirce) may help them to concentrate on the formal dimensions of the cultural process rather than using empirical classroom data, with the ultimate goal of analyzing culture as an extremely complex process (696). Their study arrives at two major conclusions. The first concerns the question of denotation, a technical term from semiotics that they refer to when describing the ways in which images and texts are connected to each other in the setting of language teaching and learning due to grammar constraints. Denotation is concerning because it hinders a freer process whereby learners and instructors would be able to explore other cultural connotations. They also conclude that materials are not the privileged object to analyze in language teaching and learning. Materials, such as textbooks, are not sole guardians of all embedded cultural assumptions in class because classroom activities typically catalyze meaning-making processes. The authors note that in-class activities are usually very controlled items, either by instructors or by instructions given in the book. This high level of control does not generate a space that invites learners to draw their own interpretations of the target language cultures. To create an environment that elicits cultural reflection and analysis from students, the authors believe that deno-

tation and control should be revisited, or altogether eliminated, as they run counter to reflexive thought necessary for cultural analysis.

But when did it become important to incorporate culture as intrinsic to language learning? When did just learning a language without a cultural context embedded in it became crucial? Claire Kramsch in her article "Culture in Foreign Language Teaching" (2013) traces (1) the history of the introduction of culture into foreign language teaching, (2) the shift and confusion about language teaching goals: is it supposed to raise awareness about language? or is it instrumental to future professional careers of learners in a globalized market? or else should it aim to teach effective communicative skills to travel?, and (3) the role of language teaching and learning in articulations of national and social identities. Kramsch's key argument is that language learning cannot be equated to learning an external object. That is to say, one does not learn language in the same way that one might study mathematics or a scientific phenomenon. By making this presupposition central to her discussion, she purposefully draws from a post-structuralist perspective to emphasize how in language learning learners undergo personal transformations when exposed to a different language. Therefore, Kramsch considers that language is not a rational structure that can be grasped in its totality, but rather something we live in and are shaped by. Language, consequently, is what allows for culture to be produced and although she believes it is arbitrary—in line with the Saussurians—she also advocates for a historically marked motivational aspect of language that must be examined. In this sense, the encounter with language as an "other" requires learners to delve into the historical traces that can be found in language. This necessary encounter with the other through language is what causes Kramsch to comment, "language learners learn who they are through the encounter with the other and they cannot understand the other if they don't understand the historical and subjective experiences that have made them who they are. But they cannot understand these experiences if they don't view them in the eyes of the other" (61). This implies that it is through learners' own culture that they establish analogies with the other culture that allow them to understand and penetrate the world that the other represents. This is why Kramsch believes more in working with students to help them develop intercultural competence than in teaching them culture as we commonly refer to it (60). Kramsch's framework emphasizes building on this unfolding of an inside and an outside, common to the post-structuralist framework, in order to emphasize the importance of

this split that, she argues, is constitutive to foreign language learning. She expresses this idea via the concept of "transgredience," which she borrows from Bakhtin that refers speakers' ability to see themselves from the outside generating what she calls "a third place." This in turn, leads her to underscore the role of the learner in producing meaning when entering in contact with another language that brings an intrinsic culture, a world, which needs to be processed by attributing meaning to it.

Based on these studies I intend to focus on three major aspects that I think are central to the discussion of culture as it is addressed in language instruction in addition to how culture in language teaching and learning contributes to shaping presuppositions that students rely on when encountering foreign cultures in this setting. The first study closely looks at how culture is a process, an alive element that cannot be located in the books as a confined and immobile aspect in and of itself. Such premise provides context to understand culture in this research paper as a meaning-making process where students have a predominantly active role. This argument aids us in building our case to suggest that students are affected by the forms in which culture is presented to them while, at the same time, it affords them a schema to reproduce the same logic in their own work.

Weninger, Csila, and Tamas Kiss provide a basis to consider how language instruction has repeatedly instrumentalized culture, reinforcing an idea that Canale implicitly points to regarding the objectification of culture. From their semiotic approach we gain perspective on the complexity of culture and possible ways to generate spaces in class that elicit less connoted and controlled approaches to its learning. If culture is a meaning-making process, then we need for language instructors to stop utterly controlling the setting in which that meaning is produced. To this point, this study offers an affirmative critique that defends incorporating culture as a process into the classroom, helping this project think about alternatives to how culture is presented in SPAN2090.

The historical analysis that Kramsch brings to the ground is an excellent background to think about the historical relationship between language and culture in language instruction, its limits, and possibilities. Arguably, her advocacy for moving towards an intercultural competence model may also find its own limits, but it can be a great way to think about how monolithically presented culture calls to be challenged and transformed.

Methods

For this study I followed a qualitative approach to analyze the extent to which ideological beliefs, conscious or unconscious, are inferred from the ways Spanish language and culture are taught in Intermediate Spanish Language (SPAN 2090) which I taught in Spring semester 2017. This study examines how textbooks and other material sources present Hispanic culture in a way that directly informs students' perception of said culture as uncomplicated. Particularly, I would like this study to explore how these materials not only affect the ways students perceive Hispanic cultures, but also how students think of foreign cultures in general. Moreover, the textbook's portrayal of Hispanic cultures as uncomplicated, in fact, falls in line with the increasing interest of universities in the US to enhance their finances by promoting study abroad options among undergraduate students, often presenting them as attractive tourist destinations. The tourist-y framework of textbooks and study abroad packages serve to elicit students' interest as long-stay visitors in popular destinations such as Spain (ranked number three in top student study abroad destinations according to a study done in 2012 by CBS News[2] shaping students' subjectivity not as future residents of a country, but as consumers. I focus on how culture is presented as small digestible clips, precisely to frame culture as uncomplicated and consumable. Interestingly, as a language instructor I have witnessed the effect of these aspects on how students see Hispanic cultures, a phenomenon which I also intend to analyze in this paper. For instance, in an in-class group oral presentation students were assigned in my class, it was noticeable that many students made significant reference to tourist attractions.

In order to analyze how students' relationship to Hispanic cultures is indirectly conditioned by the study abroad industry and directly conditioned by the ways that textbooks, and language instruction by extension, present culture, I consider three material sources. The first material sources are teaching materials used in SPAN2090, including the textbook, assignments referring to cultural components, and a library research guide specifically designed for class projects[3]. The second material source is the work produced by my eleven SPAN2090 students, specifically the presentations on one aspect of Hispanic cultures of their interest that they researched and presented in class. The third source of data is a survey of over one hundred students in eleven sections

of SPAN2090, an intermediate to high-level Spanish course. In the survey I ask students to respond anonymously to questions addressing a) their initial motivations to take the course, b) their likeliness to enroll in a study abroad program and their reasons to do so, and c) the way they see Hispanic cultures as they have been exposed to them in this course. These data are analyzed both quantitatively and qualitatively, as I am interested in looking at both students' preferences and the level of student attachment to this mindset.

To contextualize the role that teaching materials have in this class I offer a literature review of the influence of textbook discourse on instruction of Spanish and other languages. Because this project is interested in analyzing how culture is strategically framed in order for students to perceive it as a dimension they can easily grasp, focusing solely on Spanish language is not strictly necessary.

This study can help us think of the ways in which language instruction deals with understandings of culture as a component that can be totalizable, that is as a whole unit that can be consumed and comprehended.

Analysis of Source Materials

Spanish Language Textbooks or Tourist Guides?

Glancing through the pages of Enfoques, an Intermediate level Spanish language textbook, one soon realizes that language and culture textbooks are challenging items. Among language professors, there is a shared understanding that textbooks are not fully functional materials. Sometimes grammar is only roughly explained, other times cultural content is embarrassing, and more often than not they are overly expensive considering their quality. Nonetheless, language teachers prefer to use them not only to avoid copyright issues, but also to provide students with a coherent book that they can rely on, instead of a variety of written materials that may not all follow the same conventions. Enfoques is no exception, every instructor I spoke with at Cornell that was using this textbook had serious issues with its structure, content, and grammar explanations, however our coordinator was very good at filling the gaps and selecting which parts of the book should be utilized. For this study I am interested in looking at some examples of how cultural content is presented in the book in order to provide a material appreciation of what students were exposed to in class.

As the following images show, the portrayal of Hispanic cultures fall into hackneyed stereotypes that have long overflown Spanish and Latin American identities. In the first image we see two pages devoted

[2] http://www.cbsnews.com/news/20-most-popular-study-abroad-destinations/

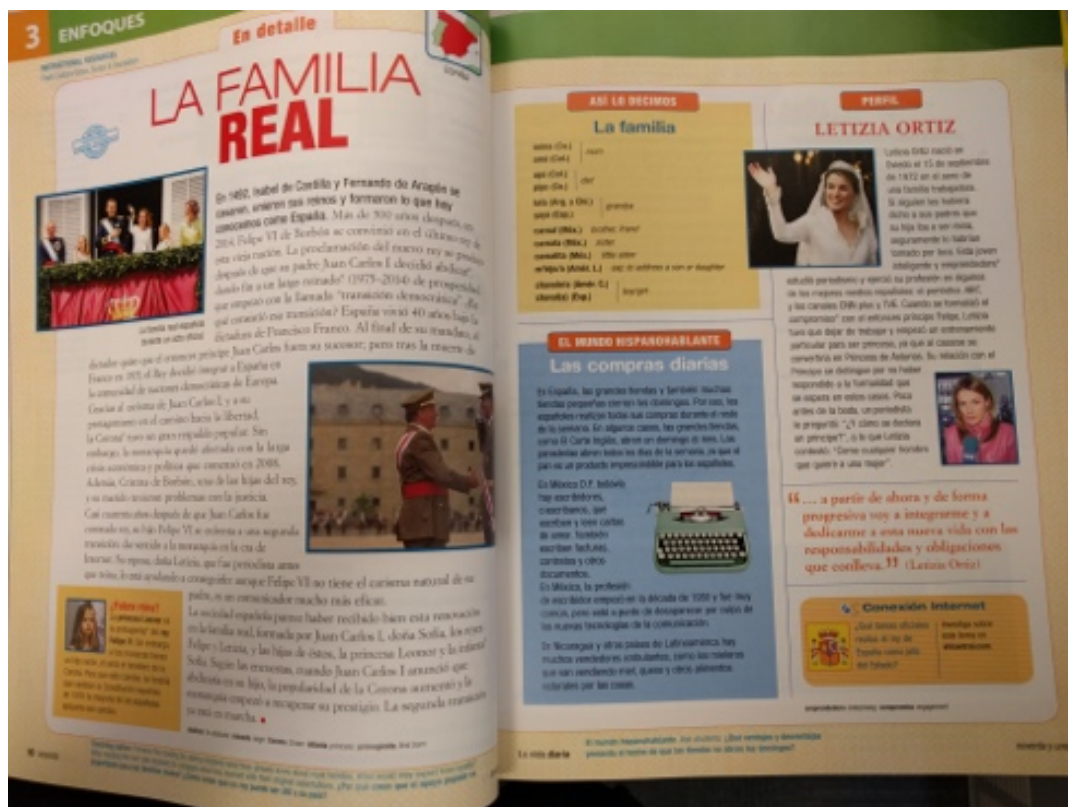[3] http://guides.library.cornell.edu/SPAN2090

Figure 1: Example of textbook material discussing the Spanish royal family.

to "La Familia Real", the Spanish royal family that not only represents a very controversial issue in Spain, but also models a traditional understanding of nuclear families (Figure 1). Spain lived through two Republican regimes, both very progressive, that abolished monarchy. The Second Republic (1931-39) was overthrown by a military coup in 1936, followed by a Civil War that lasted three years and which left terrible misery and poverty. When in 1975 Franco, the fascist Spanish dictator, died in bed, there was a process to transition into democracy that was extremely problematic. Delving into details about Spanish Transition exceed the purpose of this study, but in order to understand the problematic nature of this course material, it should be understood that the monarchy was reinstated and Spain became a monarchic state, even though the coup had taken over a Republic and half of Spain's population was against the decision. In these pages, Spain's Royal Family is pictured as essential to Spain's transition into democracy when in fact King Juan Carlos played a very minimal role in the process. But it becomes much more scandalous when these pages address Spain's financial crisis in 2008, which resulted in the national unemployment rate increasing from 8.3% to nearly 27% between 2006 and 2013[4], to say that the Royal Family suffered from

economic struggles when in 2015 they were allocated 7.7 million euros from national GDP. Therefore, the portrayal of the Spanish Royal Family as a role-model of a family and who also suffered from the Spanish financial crisis is not only a simplification of what monarchy means in Spain, but also a misrepresentation of events.

The section on "Telenovelas" reinforces a long standing stereotypical image of Latin American societies. Although the text acknowledges widely extended criticisms that "Telenovelas" (soap-operas) have received for being extremely sexist and emphasizing oppressive class structures, the text concludes with a statement from a viewer who says that "telenovelas" are innocuous, pure entertainment (Figure 2). It would be one thing if the textbook addressed sociological controversies at stake in Latin American "telenovelas" and compared them to gender roles and class structures in US soap-operas to analyze whether differences are as striking as the text silently leads us to think. In an exercise of this sort, instructors could work with students in pinpointing hidden ideological assumptions in class materials that could help develop a critical perspective in college students. However, as it stands, two pages to flag "telenovelas" as essential to understanding Latin American societies constitutes a very superficial approach to real social issues. Especially when it pictures them as being representative

---

[4]http://www.elmundo.es/mundodinero/2007/01/26/economia/1169798824.html
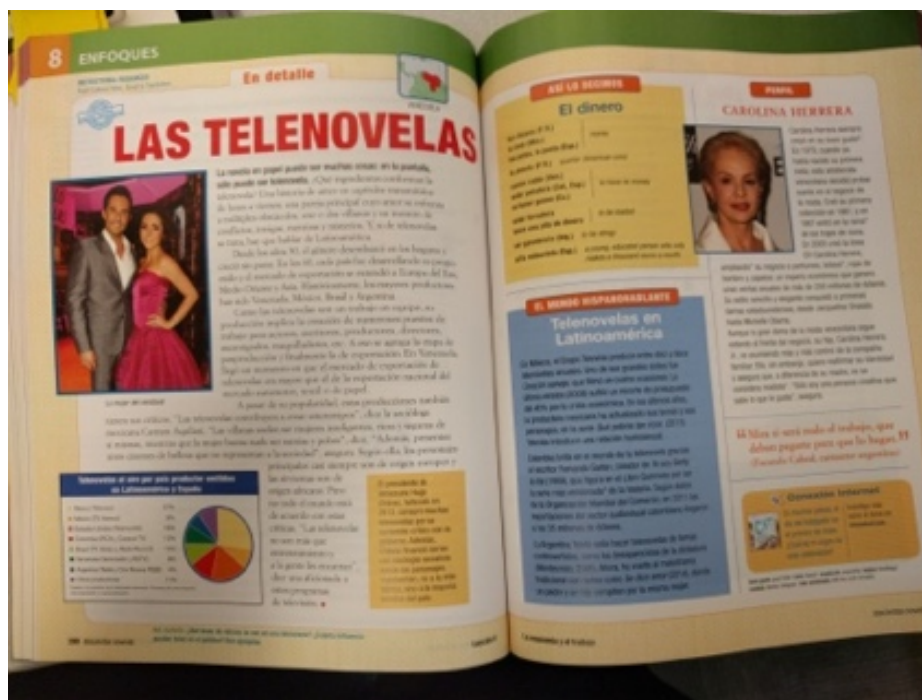
Figure 2: Example of textbook material discussing Spanish language soap operas known as "telenovelas".
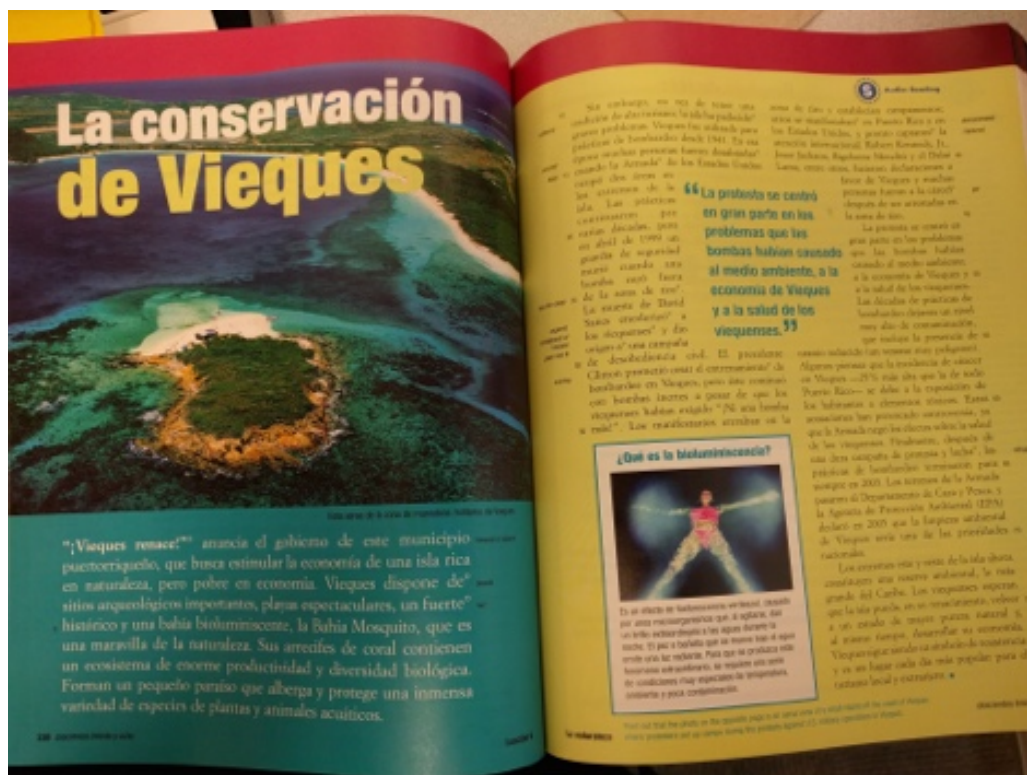

Figure 3: Example of textbook material discussing tourism and conservation of the island of Vieques, Puerto Rico.

of a conglomerate of countries such as Mexico, Argentina, Brazil, and Peru, which have very different social compositions, political histories, among many other factors.

Vieques in Puerto Rico is a great example of how textbooks borderline vacation brochures (Figure 3). The opening paragraph starts with "Vieques reborn" and continues with some sentences on the island's biodiversity, its financial struggles counterpointing its "spectacular beaches," framing Vieques as a great vacation destination: it's close, inexpensive, and beautiful. The following text does address that the US Government used Vieques for bomb testing causing high levels of pollution in the area. After many protests and international attention, the US ceased all military activity and Puerto Ricans have worked to restore the island's profuse nature. The excerpt ends on a hopeful note for Vieques' impoverished economy by highlighting its rising interest as a tourist destination.

These three examples extracted from Enfoques provide a succinct but clear sample of ways this Spanish language textbook addresses cultural content. It should be said that other book sections contain small excerpts from literary texts , as an example of what is often referred to as "authentic" materials, to underscore that these materials haven't been specifically designed for language learning, rather, they are materials that belong to a given society's cultural production. Nonetheless, Enfoques is plagued with "bits and pieces" of so-called cultural analysis similar to those examined above. We can therefore have an idea of the level of depth and complexity students are exposed to when attending a class that meets three times per week for fifty minutes per session.

Student Work: Reproducing the Textbook Frame

In this section I look at students' classroom work in order to evaluate the degree to which they base their own work on models taken from class materials, which as we have seen, tend to reinforce stereotypes and Spanish-speaking countries touristic interest.

Three groups of three to four students were assigned to carry out twenty minutes in-class presentations after interviewing a native Spanish speaker in the community. In order to ensure the interview was well prepared, students had to research a topic relevant to the interviewee's country of origin (i.e. social issue, cultural issue, financial issue, medical issue, environmental issue). As instructors we provided students with a two-page, detailed assignment explaining steps they should follow in order for their research to be thorough. Cornell University's library

developed a website [5] with materials especially designed for this course where they could find extensive bibliographies, websites, and other sources that could aid them in their research. Students were given examples of what constituted a good presentation and what type of clichés and errors should be avoided. Even with this information in hand, students in my class prepared presentations that repeatedly fell into worrying simplification or pictured a particular country by highlighting touristy attractions.

Here I present a brief summary of students' work in order to analyze the types of problems I saw. Group 1 presented on Mexican Food. Their presentation focused on the different varieties of foods traditionally consumed by Mexicans and provided a very superficial history from pre-Colombian times until today aiming to show how certain foods have evolved. Group 2 presented on poverty in Honduras where they provided data on distribution of poverty according to specific demographics and compared it to US rates. Group 3 presented on Puerto Rico and divided their presentation in three parts: 1) overview of geographical information, 2) overview of historical relationship with the US, and 3) beaches and natural parks that should be visited if one should go to Puerto Rico.

Without wanting to analyze each presentation in depth, I observed a few patterns recurring through all presentations. Themes that the students chose were not guided by a research question as prompted. Therefore, their themes were unproblematic and superficial. Students also barely carried out research on their topics. This was confirmed by a lack of both source citation and contextualization of their topic. Most information students presented was descriptive and did not contain either creative or critical approaches to their materials despite creativity and critical analysis being prompted by assignment guidelines. All presentations showed a lack of coherence in structure where Students would jump from one major issue to another without connecting the parts. Finally, interviews were not prepared based on previous research as instructed, rather, mostly random questions were directed to the interviewee and the interview content was not an integral part of students' presentations.

To be fair, in previous courses I have taught I made sure to meet with all groups prior to their presentations in order to advise them on all these matters. I ask students to prepare their materials at least one week in advance and after I give them feedback they can make all necessary changes before presentations are due. Taking into account that I was doing this

_____

[5]http://guides.library.cornell.edu/SPAN2090

study, I decided to not intervene in students' work related to this assignment in order to see what kind of results students showed without instructor advising.

As seen in the very brief descriptions I gave of students' work, for the most part, students presented for twenty minutes on very generic matters, with hardly any research backing their presentations, even though many sources were made available to them including a website designed specifically for this assignment. As far as I could evaluate their cultural understanding of their chosen topics, students were oversimplifying and generalizing particular Hispanic realities as well as emphasizing aspects that would make those realities attractive for future travel. Therefore, they were to some extent, mirroring the Enfoques textbook, which as we have seen tends to follow this pattern.

Data Analysis: Students' Perceptions

Students were surveyed (Appendix 1) anonymously across eleven sections with IRB approval [6]. A total of seventy-three students completed the survey, representing roughly 65% of all students taking SPAN 2090. Questions that were addressed ranged from asking them about their interest in study abroad before and after taking this course, their preferred destinations, and how competent they felt culturally speaking in regards to Hispanic societies.

Here I present some relevant results that can help comprehend how students perceive their competency regarding cultural content after taking SPAN 2090. For question #10: "Has your perspective on Hispanic communities within the US changed after taking Spanish courses at Cornell? If yes, how? If not, why not?" 31 out of 73 students (42%) thought they hadn't learned anything new about Hispanic culture in the US after taking this course. Their answers touched on a variety of matters but there was a pattern that repeated where these students felt they were aware of most cultural topics discussed in class either because of previous classes, or through their own knowledge. So although cultural analysis is a fundamental part of this course, almost half of the class felt as though they hadn't learned anything new. Of course, we should distinguish between what they think they learned and what they actually learned. But since we are measuring their perceptions in order to understand their mindset when it comes to Hispanic cultures, I focused on their own perception.

For question #4 "Has taking this class made you consider doing a study abroad program in a Spanish speaking country?" 34 out of 73 students (47%)

---
[6]Approval Date: May 26, 2017 / Protocol ID#: 1705007182

responded that they did consider going abroad after taking this course. This means that this course proved to be a good source to generate potential future study abroad students.

For question #7 "Do you feel like the cultural component included in your language classes at Cornell provided a complex approach to cultural materials (films, literature, music, politics, etc.)?", interestingly 57 out of 71 students (80%) responded affirmatively expressing they thought cultural materials presented in class provided a complex approach to culture. This data is striking given that as we have seen, cultural content in this course tended towards simplification, stereotyping, and touristic interest as a means of presenting culture as a component that could be easily grasped and interiorized by students, even though we know that cultural realities are, more often than not, complex, contradictory, and difficult to understand. Therefore, we could argue that after taking this course, we see students produce overly simplified cultural analysis, yet we witness witness that they believe to have been exposed to complex approaches about Hispanic cultures.

Discussion

We have seen how financial interests have attracted investment in study abroad programs in the US for at least twenty years. Following a business mindset that is in line with corporatization of education, cultural experiences abroad have been framed as consumer packages that sell experiences to college level students looking for enriching life-experiences while bettering their résumés. This cycle can be explained not only from a market perspective, but also according to a postcolonial perspective. Said critique would see in study abroad businesses a prolongation of colonial legacies. Students are indirectly addressed as first world citizens (implicitly above all other non-Northern European cultures) that should travel the world and learn about the other, who are always consciously or unconsciously perceived as less human and less complex. In order to fit into "culture clips", foreign cultures need to be simplified.

The humanities loss of control over cultural realms has not necessarily meant a devaluation of culture. Rather, its interest has been transferred to other disciplines which, far from being interested in questioning colonial legacies whose reinforcement of imbalances and inequalities is still prevalent, are conversely more focused on how to become proficient in foreign cultures in order to make profitable business with them. Language & Culture has thus become a discipline at the service of such endeavors, which is one of the ways in which disciplines in the Hu-

manities have survived. Struggling with its own challenges, language instruction has experienced a demand to incorporate culture when instruction hours necessary to achieve language competency are already scarce. Language instructors who are trained philologists and linguists have been gradually asked to become cultural analysts to train students in intercultural competence—though I'm sure administrators' lexicon avoid these terms. Relying on instructors' own knowledge and unsatisfying textbooks, culture in language classes has tended towards being packaged as digestible clips at the service of language learning, not as themes in and of themselves. This approach to culture in some "bits and bytes" falls into the same logic as those colonial legacies that base their understanding of the other and foreign cultures as "less" for they both are forced to simplify and uncomplicated in order to move on, there's no time for a deep discussion on culture. Language learning is often thought of as a window to the world, a form of expanding our knowledge, since it is believed that learning a language is learning a different system of thought. Even if unwillingly, language instruction has in fact become a mechanism to nurture oppressive logics that we carry since colonial times, if examined from a postcolonial perspective.

We have seen that students tend to reproduce the same models found in textbooks regarding cultural understanding. Even if they are provided with additional sources and rigorous guidelines, they base their understandings of foreign cultures on the types of materials used in class day-to-day, whose format in "bits and bytes" reduces reality in order to remove any contradictions. But what becomes more striking is that in their own perception of how foreign culture is being taught, they seem to believe that in-class approaches to culture are deep and complex when we have seen they are quite the opposite.

This can lead us to think that such a mindset determines students' predisposition at their study abroad destinations, making them establish a very superficial relationship with the target culture, which in turn feeds into the colonial logic discussed earlier. More and more, study abroad programs have transitioned from working with local universities to offering their own programs disconnected from host universities, and many times classes are not even instructed in the target language. This trend has proven more profitable since study abroad programs deliver all services, but it also allows the programs to retain more control over students' experiences. It is important to also mention that the phenomenon of "helicopter parents" has played a significant role in this change given that parents, often the ones financing students' experi-

ences abroad, are very concerned with their children's safety and well-being, and are anxious about students not being able to navigate their host country. Such an obsession with their children's situation has left little space, if any, for students to confront challenges and frustration, typical experiences and feelings to have when living in a different country where the language is not your own. Therefore, study abroad packages have become more and more controlled and providers have grown their student services extensively.

But if part of becoming immersed in another culture has to do with giving up control insofar as you find yourself constantly making mistakes, failing to understand your context, and feeling disoriented and lost, how can utterly controlled study abroad packages, as they exist today, actually provide a platform for cultural competency? That is to say that the study abroad industry is, in fact, framing their product as offering a controlled experience of foreign culture, claiming it provides authentic experience leading to competency in the target culture (which could be transferrable to other cultures). But we know that these types of programs build off the premise of simplifying culture to make it seem easy and unchallenging for students. We also know that authentic cultural experiences are full of complexity, challenges, and misreadings —not to say that they are less adventurous, exciting, and fun because of these challenges. Also, in order to have access to these experiences students need to be less protected in order to have the liberty of building meaning without an extremely controlled environment.

Culture is a living thing that students, in this case, interact with and on which construct meaning. How can we build a curriculum in language instruction that acknowledges that and plans classes accordingly? It seems to me that only if we start rethinking what culture is and how to trust students that they are able to take on deep critique, will we be close to preparing them for complex abroad experiences. This is not to say that all students are determined by this exposure to culture. Of course many other factors such as previous experiences, family background, class, etc., inform our mindset and some students have a better preparation to understand the complexities of foreign cultures. But the question for me should focus on deeply transforming the forms in which culture is incorporated into language instruction. There's a lot of potential to encourage critical, creative, and emancipated students.

References

Blanco, J. & Colbert. (2016). Enfoques, 4th edition. Vista Higher Learning. Canale, G. (2016).

"(Re)Searching Culture in Foreign Language, Textbooks, or Politics of Hide and Seek." Language, Culture and Curriculum 26: 2, 225-243.

de Gordon, M; McDonough, C; Palmerio-Roberts, R. (2016). "Hispanic Faces: An Exploratory Study of How University-Level Spanish Language Instruction Impacts Perceptions of Hispanics" Hispanic Journal of Behavioral Sciences. Volume 38, Issue 4, 559-571.

Guerrettaz, A. M. & Johnston, B. (2013). "Materials in the Classroom Ecology." The Modern Language Journal, Vol. 97, No. 3, 779-796.

Harder, A., Andenoro, A., Roberts, T. G., Stedman, N., Newberry, M., Parker, S. J., & Rodriguez, M. T. (2015). Does study abroad increase employability? NACTA Journal, 59(1), 41.

Kramsch, C. (2013) "Culture in Foreign Language Teaching." Iranian Journal of Language Teaching Research 1 (1), 57-78.

Musumeci, D. & Aski, J. M. (2010). "The Integration of Culture in Italian First Year Textbooks Author(s)." Italica, Vol. 87, No. 1, 21-36.

Nair-Venugopal, S. (2015). "Issues of Language and Competence in Intercultural Business Contexts". Language and intercultural communication (1470-8477), 15 (1), 29.

Tajeddin, Z. & Teimournezhad, S. (2015). "Exploring the hidden agenda in the representation of culture in international and localised ELT textbooks," The Language Learning Journal, 43:2, 180-193.

Said, E. W. (2003). Orientalism. (25th Anniversary ed.). New York: Vintage Books.

Weninger, C. & Kiss, T. (2013). "Culture in English as a Foreign Language (EFL) Textbooks: A Semiotic Approach." TESOL Quarterly, Vol. 47, No. 4, 694-716.

Appendix 1

Student Survey

This survey is being conducted as part of a research study related to the relationship between culture as it is dealt with in Spanish 2090 and the study abroad industry. Completion should take approximately 5 minutes. The results may be presented in summarized form, and viewed only by the researcher, Katryn Williams Evinson. The results are anonymous and may be shared in publications and/or academic presentations. All responses are kept strictly confidential, with only numeric coding on the surveys.

Your participation is completely voluntary. You may choose to refuse to participate, discontinue at any time, or skip any questions. Participating in this survey will not result in any risks or benefits (e.g. no extra credit is awarded for completing this survey) and no compensation is offered (e.g. no extra credit will be awarded for completing the survey). Participating in this survey has no impact of any kind on your grade.

Any questions concerning this project can be directed to the researcher, Katryn Williams Evinson (kw475@cornell.edu). If you have any questions or concerns regarding your rights as a subject in this study, you may contact the Institutional Review Board (IRB) for Human Participants at 607-255-6182 or access their website at http: www.irb.cornell.edu.

Thank you for your time.
Katryn W. Evinson

1. Why did you decide to take Spanish as a second language? What were your motivations?

2. Have you considered enrolling in a study abroad program travelling to a Spanish speaking country?
Yes / No / I don't know If yes,
a) What country or countries have you considered? b) What most attracted you from this particular location? c) Do you think you will/would acquire a competent understanding of the culture after being there for the academic year/semester/summer program? In what way? d) Do you think there is some component (cultural, personal, academic, etc.) that will be challenging for you? Which one(s)?

3. Has taking this class made you consider doing a study abroad program in a Spanish speaking country?

4. Has taking this class made you want to travel to Spanish speaking countries?

5. Please rank from 1 to 5 how familiar you feel to Hispanic cultures after taking Spanish language courses at Cornell (1 being the least and 5 being the highest): a. 1 b. 2 c. 3 d. 4 e. 5

6. Has your perspective on Hispanic communities within the US changed after taking Spanish courses at Cornell? If yes, how? If not, why not?

7. Do you feel like the cultural component included in your language classes at Cornell provided a complex approach to cultural materials (films, literature, music, politics, etc.)?

8. Please explain what was the most meaningful cultural component considered in the class that you found was the most complex to comprehend (films, literature, etc.):

9. When you decided on a topic for your in-class presentation with your peers (group 20 min presentations on culture) what guided you to choose your geographical region?

10. What was the most interesting information you learned during the preparation of your presentation?

11. What was the information you learned from other presentations that you most appreciated?

Trust, Power, and Transformation in the Prison Classroom

Frances Fairbairn

Department of Philosophy, Cornell University
fhf22@cornell.edu

Abstract

This article asks two main questions: i) what is the role of power and trust in transformative learning in the prison classroom, and ii) in particular, what is the role of power and trust in the decision of whether to transform one's meaning scheme in the face of new information or whether to simply reject the new information. I interviewed 19 educators who teach in correctional facilities about their experiences. These interviews were transcribed and the data was coded. In light of the themes that emerged, I constructed a five-stage model of transformative learning which should be thought of as complementary to (not exclusionary of) Mezirow's original ten-phase account. The data collected indicates that trust and power are particularly important at stages 3 and 4 of my suggested five-stage model, where the student is at a crossroads as to whether to revise their existing meaning schemes or merely to reject the new information.

Keywords: transformative learning; prison education; transformative experiences; Jack Mezirow; power; trust; student resistance; student push-back

Introduction

Power is inherent within the teacher-student relationship: the teacher has the power to dictate what it takes to succeed in a class, and often has the power to influence the student's future fairly significantly. We have all had those students who are applying to law school and need A's across the board to get in. In the prison classroom, these power dynamics become more distinct: misbehaving in a class can mean being kicked out of the program and even put into solitary confinement. In addition, students are already disempowered by virtue of their situation as incarcerated individuals, and more often than not, are disempowered through membership in various marginalized social groups (for example class, socioeconomic status, race). And yet, prison education has proven to be vastly beneficial. Recidivism rates are lower for individuals leaving prison having completed some form of education (Cecil et al. 2000; Coley & Barton 2006; Davis et al. 2014), behavior within prison is improved for individuals in education programs (Brewster 2014; Davis et al. 2014), and students themselves report improved quality of life when enrolled in education programs (Davis et al. 2014; Dhami et al. 2007). After considering these issues, I wanted to find out more about the role that power dynamics play in the prison classroom and how they impact the classroom dynamic, the aims of the teacher, and the experience of the students. I was also interested in whether and how students undergo transformative experiences in the prison classroom, experiences which result in a shifted world-view for the students. These interests led me to the following research questions: what way does prison education transform incarcerated individuals? How do the dynamics of power and trust influence this transformation? This paper presents a qualitative research project that I conducted with an aim to answering those questions.

Purpose

I am a graduate student at a prestigious university in the United States. My institution is involved with a program that brings college-level classes into local correctional institutions on a volunteer basis. In teaching various classes with that program, I became interested in the way my role as an educator changed, and the way it changed when I taught at different correctional institutions. In particular, I was interested in the role that power played in the prison classroom, and in the means of fostering trust in an environment that is largely seen to be lacking in trust.

In the course of my project, it became clear that transformative education would be of central interest. Many interviews clustered the topics of student push-back and resistance with students' trust and empowerment/disempowerment. A recurring theme both from the interviews and from my personal experience was that trust and power are particularly important elements in dealing with student resistance and in encouraging student growth. When we consider the role of prison education as a way of facilitating change in the outlooks, paradigms and prospects of incarcerated individuals it becomes clear that there is some connection between this kind of transformation and

the balance of power and trust in the prison classroom. The more I researched for this project, the more these themes presented as central to my research and necessary to be developed.

Literature Review

Prison education is held to have many benefits, among them lowering recidivism, improving behaviour, and quality of life. One particular benefit that I'd like to highlight is that education can transform not just the prospect, but the outlooks of incarcerated individuals. In general, education can have a transformative effect on individuals by changing the way they see themselves and the world they live in. The paradigm-shifting effect of education is explored in the works of Paulo Freire and Jack Mezirow, both of whom have enjoyed an extensive following literature. Jack Mezirow's work focuses on adult education, going so far as so suggest that transformative education can happen only with adult students. Freire's is more general, and both have had their theories extended and applied by other authors interested in the topic. It seems to me that the transformative effect of education is especially important when considering prison education, particularly given the forward-looking goal of reducing recidivism – which is (to an extent) transformative in itself: we are asking individuals to change their previous ways of life. It is curious, then, that there is relatively little research conducted on transformative learning in the prison setting. This paper seeks to fill that gap by considering the role of and possibilities for transformative learning in prison classrooms.

Jack Mezirow

Jack Mezirow introduced the concept of transformative learning in the 1970's (e.g., Mezirow 1978). His view was refined and developed across the subsequent 20-30 years until his death in 2014. In general, Mezirow claimed that transformative learning occurs when, as a result of exposure to new ideas, theories, behaviour etc. students experience a profound shift in their paradigms of thinking, their presuppositions (1978, 1990, 1991, 1995). So, through transformative learning, students critically reflect on their ways of assessing and interpreting the world, they question these methods and assumptions, and they reject/revise some of their assumptions in light of new evidence/information.

According to Mezirow's view, when we receive any type of information from the outside world, we interpret and understand it via meaning structures. Meaning structures can be understood as general frames of reference, and are made up of various smaller assumptions and perspectives such as individual beliefs, values, emotional responses, and pieces of knowledge (Mezirow 1991, pp. 5-6) as well as general world views, theories, presuppositions and goal orientations (Mezirow 1990, p. 44). These smaller components together constitute a frame of reference in the sense that, when I receive a new piece of information, say that I got an A on my recent math test, I assess and interpret that piece of information against a background of beliefs, assumptions, presuppositions, and general world views. These background components that make up my meaning structure or frame of reference can mean the difference between my interpreting that information as evidence of my mathematical ability and interpreting is as evidence of the ease of the test. If my frame of reference tells me that I am hopelessly bad at math and could never ace a math test, then I am likely to assess my achievement on that test with suspicion.

Transformative learning occurs when, either via a single profound experience or through a series of experiences, the new information I receive results in a paradigm change to my current frame of reference. Having a frame of reference which tells me I am hopelessly bad at math may initially make me interpret the fact that I got an A on my math homework as evidence that the test was too easy. But after taking many tests and doing well, and after taking upper level courses and still doing well, I may be forced to adjust my frame of reference in order to accommodate this data. In this case it is no longer rational or feasible to interpret that data as evidence of the test being too easy. The only reasonable way to interpret the data is as evidence of my being good at math. That information, however, is not consistent with what my existing frame of reference tells me, namely that I am no good at math and will never be any good at math. In this case, I have undergone a transformative educational experience. What I have learned has resulted in a change in reference frame – a change in my meaning scheme.

Transformative learning is distinct from other types of learning in that the new information is not simply assimilated into an existing reference frame; rather, the reference frame itself is altered in order to be able to accommodate the information. In ordinary cases it is the former that takes place. In this way, a transformative learning experience, where an existing reference frame is rejected in order to accommodate new information, is a profound and relatively unusual thing.

Transformative learning can occur in any number of different situations given the right conditions and any type of believe whatever can cause transformation in the right context. There is, however, a

general structure to the way transformative learning plays out that Mezirow presents (1995 p. 50):

Phase 1: The student experiences a disorienting dilemma.

Phase 2: The student performs self-examination in light of the disorienting dilemma which often brings on feelings of guilt and/or shame.

Phase 3: The student critically assesses their epistemic and sociocultural assumptions in order to try to resolve the dilemma.

Phase 4: The student recognizes that others have been through a similar experience.

Phase 5: The student explores possible changes to their actions, assumptions, and beliefs.

Phase 6: The student makes a plan for moving forward.

Phase 7: The student gathers the skills to put that plan into motion.

Phase 8: The student tries out their new assumptions, beliefs etc.

Phase 9: The student, via testing, gains self-confidence in their new meaning-scheme.

Phase 10: The student integrates the new assumptions, beliefs etc. into their meaning scheme.

Mezirow did go on to add to and alter these phases in later work (Kitchenham, 2008) but the core idea remained the same. Phases 5-6 focus on the mechanics of altering one's frame of reference. I will be focusing here on the question of what conditions best facilitate getting to Phase 5 – that is how students get to the point where they wish to alter their reference frames rather than rejecting of reinterpreting the new information. So I will be focused on Phases 1-4. As mentioned earlier, this question is particularly relevant in the case of prison education, given that a big part of the aim of prison education is rehabilitation, which assumes some level of transformation. There is much room for work on such a topic.

Paolo Freire

Paolo Freire, along with philosopher Thomas Kuhn (Kuhn, 1962) and philosopher and linguist Jurgen Habermas (Habermas, 1971), formed a theoretical basis for Mezirow's theory of Transformative Education (Kitchenham, 2008, pp. 105-108). Freire argued that the standard conception of successful teaching and learning was misguided due to its focus on knowledge transference (1970). On the standard conception, the teacher is seen as the bearer of knowledge and the students as empty vessels waiting to be filled with that knowledge. In this scenario, the student is a passive receiver in the classroom and the instructor holds all the power. Freire argued that successful learning essentially involves active, critical reflection on the part of the student. We can see this emphasis on reflection mirrored in Mezirow (1990, pp. 1-19) as well as the essentially active role of the student. In Freire, the transformation that takes place is thought to be a broader, societal transformation, whereas in Mezirow the transformation is individual (Taylor, 2008). However, the reflection of values is undeniable given the primacy of critical reflection and self-reflection in Mezirow's framework as well as the role of transformation in the learning process.

Thomas Kuhn

Thomas Kuhn was a philosopher of science whose book, The Structure of Scientific Revolutions (1962), was a primary influence for Mezirow's theory of transformative learning (Kitchenham, 2008, pp. 105-107). In this book Kuhn suggested that science does not evolve continually through the steady gathering of new information but rather that scientific knowledge cycles through periods of settled paradigm agreements and periods of revolution and upheaval. A scientific paradigm is set whenpractitioners of science generally agree on a paradigm that unifies and helps to explain currently accepted theories and predicts future theories and results. Every so often a paradigm shift happens, in which some new data is gathered which is so radically incommensurable with the existing paradigm that the paradigm must be thrown out and a new one built. This idea of paradigm shift in the face of incommensurability was the model for Mezirow's idea of transformation, where one student adopts a new frame of reference in response to new, incommensurable information. A frame of reference, for Mezirow, is an individualized version of Kuhn's paradigm. It is a way of organising and understanding the world, but also a means of making predictions about future information.

Keen and Woods

In 'Creating Activating Events for Transformative Learning in a Prison Classroom' (2016), Cheryl Keen and Robert Woods investigate the ways that disorientating dilemmas and activating events may facilitate

transformative learning in prison classrooms. They interviewed 13 educators who worked in correctional facilities about their experience and found that those educators did perceive transformation in their prison class. In particular, they found that educators reported that performing the roles of role model and of counsellor, as well as establishing humour and respect in the classroom, helped to facilitate transformative learning (Keen and Woods 2016, p. 28). There was an emphasis from the interviewees on mutual respect between student and teacher needed for transformation to occur. This draws close connections with this project and my questions about the effect of trust and power on transformation. I hope that this project will serve as an extension of the Keen and Woods paper but with some important additional components: i) I interviewed educators working at a college level as opposed to at a K-12 level, ii) the educators I interviewed were all working on a volunteer basis, which I believe to be a significant factor based on my findings.

Methods

I performed a qualitative study by interviewing instructors and TAs for college-level courses in correctional facilities over a semester. The interviews were recorded and transcribed and the resulting transcriptions were subsequently coded and analyzed for recurring themes. Grounded theory was used to construct an analysis that draws some significant conclusions for the data gathered.

Source of Data

I interviewed 19 educators who taught in prison. Of these, six were Teaching Assistants and 13 were Instructors. All interviewees were teaching college-level courses, and were teaching on a volunteer basis. The courses were taught across three different facilities and covered a wide range of topics including English, neuroscience, and law. Interviewees each taught one or more classes, some taught several.

Interview Method

The interviews lasted for approximately one hour and were generally un-structured: while I did have a skeleton of eight to ten prompts or probe-questions, I did not read directly from those questions nor did I ask questions in an identically-structured manner from one interview to the next. Rather, I let the interview feel more like a natural conversation, and I followed different avenues depending on what the interviewer said, what they seemed interested in and how talkative they were. A selection of interview prompts and probes are included in Appendix A. I did this purposefully in order to create a relaxed and open environment. Talking about prison education can be very difficult and emotional. Opening up can be hard, especially when there is a level of secrecy around what is said and done behind prison walls that generally makes the educator wary of sharing experiences. I wanted interviewees to feel able to talk openly about their experiences, and I wanted the environment to feel supportive and not cold or clinical. I believe that this methodology allowed me to gather deeper and truer data from my participants. For this reason also, I felt free to share experiences of my own when interviewing. In doing so, I intended to create a two-way flow of trust: the interviewee was not expected to share openly while I stayed silent, rather we were both sharing stories and experiences with each other. I believe that this strategy was essential in gathering the rich data that I did. However, it should be noted that such methodology will have affected the course of the interviews in various ways. Interestingly, this method was reinforced by my investigation into the role of trust in the classroom – one repeating theme was that the students wanted to hear more from the instructor about what their personal views were and about what brought them to teach in a prison. Whether the instructor was willing to open up to the students appears to have had a large effect on whether the students trusted the instructor. For more on this, consider in particular the codes 'trust' and 'caring'.

Analysis of Data

The 19 interviews were coded using Alan Bryman's four-stage technique (Bryman, 2008). First I skim read through a selection of interview transcriptions in order to identify major themes. After that, I read through the interviews carefully, marking the text where those themes occurred, making notes, and identifying labels for codes. During the course of this, I identified various sub-codes inductively. So, for example, the general theme 'Trust' was exemplified via the following sub-codes: 'trust breaking', 'trust and caring', 'trust and testing', 'trust and insecurity'. Subsequently, 'caring' became its own code. In the third stage, I read all the interviews and marked them systematically with the themes and sub-codes. In the process of this I identified several exemplifying questions within the sub-codes, for example the code 'Trust and caring' was exemplified by the question 'why are you here?' A full depiction of themes, subcodes and thematic questions is given in Appendix B. Finally, I used grounded theory to identify general theoretical ideas that came out of this data.

Limitations

In conducting my research, I encountered one significant obstacle that forced me to change the course of my project somewhat. It was not possible for me to conduct interviews with the students themselves due to the requirements of the program, the correctional facility, and the Institutional Review Board. This is a general problem for researchers interested in prison education and it raises its own questions about trust and power in the prison classroom since the opinions and reflections of the students become invisible on this model. In this case, I chose to explain to my students, via a written handout, that I was conducting this research, what my questions were, what I would be doing, and why I wouldn't be talking with them personally about it.

Results & Discussion

Using the interview data I was able to construct a five-stage schema that allows us to better understand the process of arriving at a transformative learning experience. This schema builds on Mezirow's original ten-phase description because it reveals more about the mechanics involved in coming to transformation as opposed to rejection/rationalization. The five-stage approach developed here also allows us to focus in on an important question – what makes the difference between a transformative experience and a rationalization or rejection of new information?

Five Steps Toward Transformation

As noted earlier, for this project I was interested mainly in Mezirow's first four phases, that is, the phases that lead up to the decision to transform. In particular, I was interested in what influences the decision of a student that chooses to alter theory reference frame, or alternatively, chooses to reject or reinterpret the new data. This is an important aspect of Mezirow's theory that, I believe, doesn't get as much discussion as it should.

According to Mezirow's theory, during Phase One a student is confronted with a dilemma: some new piece of information (or perhaps a packet of various pieces of information) is presented to them that does not easily consolidate with their existing reference frame. During Phase Two, the student critically reflects on their reference frame in an attempt to understand the dilemma. During Phase Three, the student tries to resolve the dilemma via critical assessment. During Phase Four, the student recognises 'that one's discontent and process of transformation are shared and that other have negotiated a similar change' (Taylor, p. 8). Phases Five to Ten describe the process of identifying, considering and auditioning potential new frames for reference.

Notice that at Phase Four, there is already the assumption that a paradigm shift will take place for the student. Yet there is a lot that must take place between Phases Three and Four in order for that to take place. Phases One to Three simply present us with a student who has encountered a dilemma – where a new piece of information has been received that cannot be assimilated into their existing reference frame - and is trying to resolve it. It is wrong to assume that such a situation, by itself, will lead to a change in paradigm. After all, a much more likely response to such a dilemma would be to reject or reinterpret the new information. On an epistemological level, we tend to assess new evidence or information against a background of assumptions, beliefs, and other factors as a way of testing that evidence. We generally give the background assumptions more weight than we do the new evidence because we have built up those assumptions from years of experiences, teachings and inferences. So when we introduce a new piece of information, we compare it against out background reference frames in order to test that new information. If the new piece if information doesn't easily assimilate into our background beliefs, then we take that as evidence that the new information is faulty.

For example, imagine that someone tells you that rats are naturally very clean and gentle creatures –cleaner, even, than cats. This new piece of information does not fit with your existing frame of reference which indicates strongly that rats carry diseases and are very dangerous. So, in this situation you have a problem: the new piece of information that you just received doesn't fit with your existing reference frame. You have several options: you might reject the new information as false or misguided, you might reinterpret the information taking it to have a grain of truth to it (perhaps rats are clean in comparison to some other dirty animal. . .) or you may change your reference frame to accommodate the new piece of information. Usually, we are inclined to reject the new information before changing our reference frame. Individual reference frames are deeply entrenched and strongly held. They are made up of both individual beliefs (such as 'rats are dirty, diseased animals') and habits of mind (such as automatically moving from the idea of rats to feelings of disgust and fear), the latter of which may be extremely difficult to overturn.

My question, then, is 'what influences the decision either way?' When faced with a disorienting dilemma of this kind, under what circumstances will a student opt for a change of reference frame as opposed to rejection or reinterpretation of the new information?

In light of these considerations, and backed by the data gathered from my interviews, I developed a schema for the process involved in the decision to transform. This can be understood as an adaptation of Mezirow's first four phases (a full table including exemplifying quotes for the various stages of my model is included in Appendix C).

Stage 1: Initial Position - Student comes to the classroom with initial frame/perspective.

Stage 2: New Data - Student encounters new and incommensurable data.

Stage 3: Assimilation - Student attempts to fit new data into existing frames.

Stage 4: Incommensurability - Student realizes that assimilation is impossible within current frame.

Stage 5: Altered Frames - Student either rejects new data or adopts new frame/paradigm.

Stage One represents the student's initial position; their initial frame of reference or meaning scheme. At Stage Two, the student is confronted with some new information or data. The student then moves to Stage Three, where they attempt to assimilate that information into their existing reference frame/meaning scheme. In most cases of learning the student will assimilate the new data without issue. In cases of transformative learning, the student will move to Stage Four, where they find that the new data is incommensurable with their existing reference frame[7]. Once the student realizes this, they move to Stage Five where they choose whether to reject the new data, reinterpret it, or revise their existing reference frame in order to accommodate the new data.

This schema helps us to identify the point of decision in transformative learning. Stage Three involves an attempt by the student to assimilate the new data into their existing reference frame in the usual way. If this is achievable, then we have a case of 'garden-variety' learning (as opposed to transformative learning). So the student must reach Stage Four in order to transformative learning to take place. However, reaching Stage Four is still not enough by itself. As demonstrated here, transformative learning also requires that the student make a certain choice upon reaching Stage Four, namely that they will adopt a new reference frame. That the student will do this

[7]This stage reflects the Kuhnian influence on Mezirow (Kuhn, 1962).

is not a foregone conclusion, they may just as likely choose to reject the new data and preserve their existing reference frame.

This schema arose in response to analysis of the interview data. Each of these stages was exemplified in the data (see Appendix C for examples). This way of schematizing the first part of the transformative learning process also helps us to home in on an important question: once the student has reached Stage Five of the process, what influences whether they then go on to have a transformative learning experience, as opposed to either rejecting the new data or reinterpreting it?

The Roles of Power and Trust in the Five-Stage Process

After identifying my five stages discussed above, I was able to focus on my original research question: what is the role of power and trust in the process of transformative learning? Significantly, the data gathered from the interviews indicated that a balance of power and trust is particularly important for the students when they reach Stages Three and Four of the five stage process.

Stage Three is the 'assimilation' stage, where the student attempts to assimilate the new data into their existing reference frame. Stage Four is the 'incommensurability' stage. This is the point where the student discovers that the new information cannot be reconciled with their existing reference frame. It is well recognised that the process of transformation can be very distressing and alienating for students (Mezirow 1978; 1990). This is particularly well described in Mezirow's 'disorienting dilemmas' stage: it is disorientating to discover that a new and compelling piece of information is incommensurable with your existing frame of reference. A situation like this demands critical self-reflection, which can be difficult and disorienting in itself. This is why in Mezirow's ten-phase characterization there is such an emphasis on connecting with the experiences of others (Phase Four). A student who realizes that new information is not compatible with their existing reference frame can experience strong feelings of guilt and shame when they turn to critically addressing their existing reference frame. The experience of shame and guilt can be aided by the knowledge that others have been through something similar before. But, as I have been emphasising, the student will never reach this point if they choose to reject or reinterpret the new information rather than choosing to critically revise their reference frame. One important theme that came out from the interviews was that students in the prison setting are particularly likely to

be confronted with new information that was incommensurable with their current set of beliefs, and thus is disorienting. The educators interviewed were particularly aware of and concerned with this fact. As a result, they tended to experience a lot of student push-back and resistance, particularly in the form of distrust. The students were often concerned about why the educators were there: what their aims were, who they were there for, and what they intended to do with the power that their position afforded them. See thematic questions linked with coding for examples (Appendix B). Whether a student chose to revise their reference frame, as opposed to rejecting or reinterpreting the new data, depended on how the assimilation process went and on their experience at the incommensurability stage. If an adequate level of trust was not established by the time the student reached the assimilation stage, they were much more likely to reject the new data at Stage Four.

Students often come to the prison classroom with difficult past experiences with trust and power. This can often be reflected in their initial position (Stage One). Data gathered from the interviews indicated that students' initial positions were reflected in certain insecurities. Thematic questions arose from the interviews as followed:

Why are you here?

A general question on the part of the student asked why the educator had chosen to take the time to come to the class and teach them. Interviews indicated that students were concerned about whether educators where there just to pad their resumes or as part of some sort of 'saviour complex', or whether the educators really cared about them and their views. This thematic question came mainly out of codes for 'trust' ('Should I trust this instructor, or are they using me?') and 'caring' ('Does the instructor care about me as an individual?'), but also has connections with 'insecurity' ('Is the instructor using me to get something else?'). It might also represent Stage One of the five-stage process outlined above: The students' initial position is one that causes them to be wary of educators' motives for coming to teach them, especially where the work of the educators is unpaid or voluntary. This might be due to students holding a particular reference frame which tells them that individuals volunteering to teach them must have certain motives.

"...at the very beginning he did ask me about my qualifications for giving the class. (Q: Really?) Yeh, yeh, at the end of the first class he came up and asked me, 'so why did you want to teach this class, 'cause

I like to get a feel for my instructor.' So maybe what I said made him not trust me. (Q: What did you say?)I said that, [...] that all of our work has to do with our own career, like helping ourselves, like writing our dissertation, and even TA-ing in classes, that's what we get paid for and it's also to get teaching experience, but that you wanted to do something to help someone else. And I thought that really inspired, that was one of the reasons why I applied for the program. I also needed the money; that was something I didn't tell him that maybe I should have, 'cause I wasn't being completely honest, I wasn't telling the whole truth, but...so anyway that's what I said."

"(Q: How did they regard you, do you think when you first came into the prison classroom and met them all?) I don't know, they were very friendly, we shook hands on first meeting, umm...they asked me to tell my life story (laughs) (Q: Really?), Yeah, at least my background to teaching this class. (Q: Do you remember the wording of the question?) I don't. I think I had asked them to all introduce themselves, and then they asked me about my background, and so I told them my college and grad school background, where I'd studied and time that I'd taken off from study to work, and when I'd finished telling that story, I was sort of in my head, but I was suddenly jolted out of my own thinking to hear them all going 'Hmmm, hmmm', nodding approval and, like 'thanks' for opening up to them."

"Actually, after I asked them all why they wanted to take neuroscience, one of the students asked me why I was teaching there, which I was a little taken aback by. I guess looking back on it, it's not surprising why they want to know that, I mean I ask them why they were there, so why shouldn't they ask me, right? And I told them the same reasons why, basically, that I told you; that I wanted to bring neuroscience to people who wouldn't maybe have the opportunity to, and also I mentioned but I didn't focus too heavily on it that I think that there should be education in prison, and I felt like I was just doing my part to make that happen. And they seemed to accept my answer."

What are your aims?

It was often stated that students were concerned with whether the educators were interested in hearing and considering their point of view or whether they just wanted to project knowledge and information. This theme is particularly interesting because of its close ties with Freire – 'Do you see me as an empty cup to be filled with information, or do you see me as an individual with ideas and thoughts and feeling of their own?'. This question came out of codes for 'trust' (linking back to the 'Why are you here?' question), 'caring' ('Does this instructor care to know what I think or are just interested in what they know?'), and 'power' ('Are you on a 'power trip' and only interested in hearing yourself talk, feeling the power of an instructor, or showing how smart you are?'). This thematic question may also represent Stage One of the five-stage process by representing the initial position and expectations of the students based on their reference frames.

> *"There has been, I think, one instance in which I did an activity where I invited the students to ask me a question about myself and the question needed to be . . . . Typically I teach a scientific critical thinking class, and to do that I say, 'so let's get to know each other; you don't know me, what questions can you ask me the answer to which will inform you whether or not I have the authority to be the instructor? And then if you can give a good reason for why that would be informative, then I'll answer the question.' And only at that time has there ever been students who've sort of challenged, have felt comfortable to challenge, and there, their concern was 'do you care about us?' not, like, 'do you have the credentials?' or whatever . . . . 'do you care about us?'"*

> *"I do think that one teaching philosophy that I do have with all students that I think adapts very well to a prison context, in particular an adult learning context is that I'm very careful to remind people that my expertise is in psychology and I'm here to sort of present you with resources. You're the one who does the learning. In order to facilitate that, everything we do in class, I'm going to be very transparent about why we're doing it. So when I assign an assignment, I'll tell you exactly the goals or the outcomes and how the individual aspects of*

> *the assignment help you to achieve that, but it's up to you to go and achieve that. And so I think there is very strong structure in the class. I'm very, very open about exactly why we're doing anything. And so I think that might be really good in a prison setting because the students feel that I'm not, like, condescending or whatever, they know why they're doing things."*

> *"I would often thank them at the end of class, and they would thank me, you know, they'd say 'No thank you, we think it's awesome that you guys come from Cornell to teach us.' And Jay Henderson said 'It's unbelievable.' You know, and you could see the expression on his face; 'I can't believe that you guys would actually come here to teach us.' I never got the sense from any of them that they felt at all entitled to an education. Everything about prison and even teaching in prison, the prison makes it clear to them that it's a privilege and that they're not entitled to it at all. And so that's definitely drilled into them and I think that they know the perception that a lot of people on the outside have of them, and so I think that they were just generally curious why someone would drive an hour once a week to go teach a bunch of prisoners, when I could teach maybe people who were more deserving of it, you know. Or just spend my time doing something that people who aren't in prison do. Why would I? I think that they deserve it, if they want it."*

Is this a 'real' class?

It was often reported that students were very concerned with ensuring that the materials of their classes were not changed at all from the content of a 'regular' college class. They were reported to be worried that some classes might be 'dumbed down' for them. This thematic question came mostly out of codes for 'insecurity' ('Don't give me special treatment'). It may form a good representation of Stages Two and Three of the five-stage process. This question was usually expressed in response to new data and might indicate the attempts of the student to assimilate that new data into their existing reference frames.

> *"I think that a lot of the guys are insecure about the legitimacy of academic credentials, I get a lot, they write about it in their*

papers, when they come up with examples, and it also seeps into a lot of the discourse in the class, they're worried that this isn't a real college class, this is just somebody coming in and doing their little thing, but it's not like real education, and we're not being challenged like real college students are challenged, and I don't think that has to be the way I'm doing the classes, I think that's a worry that they carry into the program, so they're really insecure about that, and it affects their behavior, so I have people saying, if I'm lenient with them about things 'no don't be lenient, I want to earn it... I think you made a mistake grading my paper but don't change the grade, let me rewrite the paper' and that's actually kind of inspiring. It's also related to the feeling that they don't want to be coddled, because they're insecure that this isn't a real class. And another thing that happened was in the mid-term exam, I had... this is also example of push back... I gave them 20 study questions and I said there are 10 questions on the exam paper taken from these questions, and Cornell students would be thrilled by that, 'Awesome! If I answer these questions I can get an A or B exam!' But some of these guys were mad about that; they were, like, 'don't do that, that's making it not a real class, that's making it not hard like it's supposed to be,' and one guy was, like, 'a real college class would never do that,' Like 'I've a lesson for you about how things are run at [name of home institution]!' But sort of it's a trust issue because they don't trust the class to be like a real college class and they're insecure about that. So that's not so much a personal trust, but it is an issue of trust, so that's why I bring it up."

"....prison is sort of a depressing environment, where if something seems too good to be true, it probably is, in general, um, and it's probably just a place where people have a lot of discouraging things to say about one another, so I wouldn't be surprised if on the one hand, the students themselves are used to not getting their hopes up about things, and on the other hand other prisoners reinforce that by insulting good things people do have. I don't have evidence of either of those things being the case but it wouldn't surprise me

based on the general things that I believe about prisons. So, here's an example that might be illustrative: so, a student... we were talking about rationalization, this was part of the critical thinking material, what's rationalization? And on the exam, one of the questions was 'define rationalization and give an example,' and a student gave a brilliant and also really heart breaking example of somebody who is intimidated by their prison class and, as a result of that, doesn't take the class, and then rationalizes that by saying to himself 'it's not a real class.'"

"Maybe they're intimidated by the class, and they tell themselves it's not a real class. Maybe some students who tell themselves it's not a real class do it because they're intimidated, as a way of justifying not working as hard. I could also see other prisoners who are intimidated by those who get into the class trying to deal with that by minimizing achievement, minimizing their achievement by saying 'it's not a real class anyway.'"

Testing power dynamics.

Interviewees often spoke of cases where students seemed resistant to them or their methods and where students seemed to want to test the power dynamics in class. This thematic question mostly came out of coding for 'power' and 'resistance'. As a thematic question it might represent Stages Four and Five of the five-stage process outlined above. The students might be resisting in response to incommensurability of the data either as a way of critically reflecting or as a way of rationalizing that new data.

"But the prison students, I think there was a testing, and I didn't invite them to ask that, but I already felt a lot of .... Like, people make jokes or side comments....I had them do an activity where....I think they were trained to test also the boundaries there but in a less expressive way.....I had them do an activity where they write a letter to themselves, about something, someone, or some circumstance that they don't understand, that they've always been curious about. Then we seal it, at the end of the class we open those letters again, and the final is to write a response to themselves using psychology, like, all of what

you've learnt. And so that was the first class, and some students were, like, 'I don't think there's anything I don't understand,' or, they might be.... my TA asked them if there was anything they preferred to be called, their names, during the introduction activity (which I didn't think was necessarily a good idea) but it did invite some (...) she started shouting random things, like, what are you calling me, 'the King', this kind of stuff, so I think a lot of those dynamics were implicitly testing authority."

"Yes, there are a few resistant students in the classroom, a few more at the beginning, we kept losing a few students because one went to the box, solitary, so we lost him, even though they investigated him, he started a fight or something, and they said the fight was not his problem, not his fault, but he still went to solitary, causing him to continue with the program even though he's out, so he struggled a bit with the material, he just doesn't see a point in it, that's reflective of a lot of the students in our psychology class at the beginning, they just felt there was a lot of content and not much to navigate with, they didn't find the structure, what the purpose of it was, and they weren't making those connections like they were in Jennifer's (pseudonym) class with the readings and, you know, they were putting themselves in the character's shoes and they weren't finding much of a relationship, the importance of the work, just doing the readings, and so I did have one student, actually just last class...they all got back their first exam and I guess he didn't do as well as he had hoped, and so he told James (pseudonym), 'don't expect too much from me because I'm not going to reach those expectations, I'm just here to pass the class and get my credits and leave, I don't care about any of the information I'm learning', or something like that. That's just so hard to hear, because, you know as an instructor you want all of your students to pick up the information and find it interesting and engaging. But, yeah he just pushed back against James (pseudonym), and he thinks that it might be something to do with self-handicapping, just like resisting authority maybe."

"I mean he's one of the students who ba-

sically thinks...he's a pretty conservative guy, he doesn't want to read any stuff that has a political bias, which gets tricky. He's prickly. So I felt like the power dynamics....we had one moment where I felt really, really stressed by the power dynamics in that class last semester. Because he's kind of a History buff, so we'd have all these very specific questions about dates and events...stuff....and I felt like he was sort of crossing a line and actually, like, testing me, rather than just asking because he was curious. He was, sort of, like 'I don't think you actually know anything,' you know, 'History? You just know this wishy-washy stuff.'"

Pushback and resistance to educators may indicate that the students are experiencing the discomfort of the incommensurability stage or that they are at the final stage of the process of transformation - 'altered frames/rejection of new data'. Data from the interviews consistently indicated that establishing a level of trust with the students early on encouraged them to not reject new data even in the face of uncomfortable critical self-reflection. Interviews suggested that a student who is comfortable with their teachers' aims, methods, and intentions is more likely to try to assimilate difficult or contradictory data at a later stage. This may indicate that students are very protective of their reference frames in the prison classroom, and very suspicious of individuals offering them contradictory information. A good question might be whether such protectiveness is more pronounced in certain settings (i.e. in the prison classroom).

Future Research

This project constitutes an early investigation into an area that is ripe for investigation. There are ample opportunities for further research both on the basis of the data gathered here and utilizing additional data. I would like to take a little time before concluding to suggest some particularly fruitful avenues.

A broader range of data.

This project involved in-depth interviews with 19 prison educators across three facilities and numerous disciplines, with all classes taught at the college level. There are many more facilities that run college programs, and many more instructors who could give valuable insight into the experience of teaching college in a prison. It also bears mentioning that the fact that educators in my study were working on a volunteer basis may have influenced the data gathered. It may make a difference to the security of the

students as well as the power dynamics and behaviour of the instructors that they were working on a volunteer basis (some of these aspects are explored above). It would be valuable to gather evidence from educators who work on a contracted, paid basis, though this may be difficult given the underfunded nature of college programs in prison. In addition, it would be valuable to gather data from educators working on a pre-college (K-12) level. Keen and Woods work towards this goal, but more data is needed to draw wider conclusions.

The student's experience of transformation.

I was unable to conduct interviews or gather data from the students themselves as part of this study. This is problematic both analytically and ethically. It is problematic analytically because it narrows the scope of the data I have gathered: in investigating the transformative learning experiences of students I have not heard from the students themselves. It is ethically problematic because it further silences an already silenced group. I had the opportunity, several times, to ask students about their feelings towards prison teaching research. Their reactions were always guarded and sceptical. The mean response was that the students were frustrated that there were so many academics talking about the experience of being in a prison without actually talking to the people who are in prison. Therefore, this needs to be remedied.

The ethics of transformation.

Much of the literature on transformative education focuses on the conditions that are required for transformation, what the experience of transformation is like, and other questions having to do with how we can facilitate transformative learning. But given the profound effects of transformative learning, and the fact that it is an essentially uncomfortable experience, one question we must consider is whether and when aiming for transformative learning is ever ethical. There is often an assumption in the literature that transformations of this kind are always positive, but that is not necessarily the case. Transformative experiences may be extremely negative. An individual in an emotionally abusive relationship may come to have a transformative learning experience through repeatedly being told that they are worthless. Arguably, women and members of other minority groups systematically go through transformative learning experiences that teach them they are essentially not rational, strong, emotionally stable...etc. Imposter syndrome is a well-documented experience, but it is also, arguably, a learned experience. Should we think of the experience of imposter syndrome as

having come about through transformative learning? As I encourage further investigation into the role of transformative education in prison classrooms, I feel that I must also encourage further investigation into the ethics of transformative education.

Conclusion

This project focused on the role of trust and power in a particular stage of the process of transformative learning – specifically, the stage at which the student either decided to alter their reference frame or to reject/reinterpret the incommensurable new information. I asked the following two questions: i) In what way does prison education transform incarcerated individuals? and ii) How do the dynamics of power and trust influence this transformation? Data from 19 interviews with prison educators suggests that prison education may facilitate transformative learning in some situations where the correct environment is created. Data also suggests that power and trust may be particularly important at Stages Three and Four of the process of transformative learning and that building a foundation of trust and power-sharing/empowerment early on, students may be more likely to change their reference frames when encountered with incommensurable data rather than rejecting the data itself.

Acknowledgments

References

Brewster, L. (2014). The Impact of Prison Arts Programs on Inmate Attitudes and Behavior: A Quantitative Evaluation. Justice Policy Journal, 11 (2), 1-28.

Bryman, A. (2008). Social Research Methods. 3rd Ed. New York: Oxford University Press.

Cecil, D., Drapkin, D., MacKenzie, D. Hickman, L. (2000). The effectiveness of adult basic education and life-skills programs in reducing recidivism: A review and assessment of the research. Journal of Correctional Education, 51, 207–226.

Coley, R.J., Barton, P.E. (2006). Locked up and locked out: An education perspective on the U.S. prison population. Princeton, NJ: Educational Testing Service, Policy Information Center.

Davis, L.M., Steele, J.L., Bozick, R., Williams, M.V., Turner, S., Miles, J. N., Steinberg, P.S.

(2014). How effective is correctional education, and where do we go from here? The results of a comprehensive evaluation. Santa Monica, CA: Rand Corporation.

Dhami, M.K., Ayton, P., Loewenstein, G. (2007). Adaptation to Imprisonment: Indigenous or Imported? Criminal Justice and behavior, 34 (8), 1085-1100. Freire, P. (1970). Pedagogy of the Oppressed. New York: Seabury Press.

Habermas, J. (1971). Why More Philosophy? Social Research: An International Quarterly Critical Perspectives on the Social Sciences. Mack, A. (Ed), 38 (4), 633-654.

Keen, C.H., Woods, R. (2016). Creating Activating Events for Transformative Learning in Prison. Journal of Transformative Education, 14 (1), 15-33.

Kitchenham, A. (2008). The Evolution of John Mezirow's Transformative Learning Theory. Journal of Transformative Education, 6 (2), 104-123.

Kuhn, T., (1962). The Structure of Scientific Revolutions. 3rd ed. Chicago: University of Chicago Press

Mezirow, J. (1978). Perspective transformation. Adult Education Quarterly, 28, 100–110.

Mezirow, J. (1990). Fostering Critical Reflection in Adulthood. San Francisco: Jossey Bass.

Mezirow, J.(1991). Transformative Dimensions of Adult Learning. San Francisco: Jossey-Bass.

Mezirow, J. (1995). Transformation Theory of Adult Learning. In Welton, M. R., ed. In Defense of the Lifeworld. New York: SUNY Press (1995), 37-70.

Taylor, E.W. (1998). The Theory and Practice of Transformative Learning: A Critical Review.

Appendix A

The following is a list of interview questions and probes that I used as jumping-off points in my interviews with prison educators. As emphasised earlier, I did not ask every question of every individual, nor did I follow a strict, predetermined structure for the interviews. Rather, I tried my best to make the interviews feel like informal conversations, in order to encourage the interviewees to feel comfortable and to open up.

1. What is the most defining feature of your experience teaching in a prison?

2. Was there anything that surprised you about your students in prison?

3. What differences have you noticed in teacher-student dynamics comparing prison teaching with non-prison teaching? Probe: what steps have you taken in light of these considerations? How do they affect the class?

4. How does your relationship with your incarcerated students differ from your relationship with your non-incarcerated students? How does it differ from students' relationships and interactions with each other?

5. Have you experienced student resistance to certain tasks or assignments? Probe: Can you tell me a bit about those cases? What was the task, in what ways were the students resistant, why were they resistant? How did you deal with the resistance?

6. Have you experimented with any techniques for improving poor relationships with incarcerated students? How well do you think they worked?

7. How do you foster trust in your prison classroom? Probe: Was this easy or difficult? How does it compare to fostering trust in the non-prison classroom?

8. Did you notice any behavioural changes in your students over the course of the semester? Probe: Can you remember any particular cases/circumstances? What do you think prompted their change?

9. Have you noticed any 'breakthroughs' with students? Probe: Can you tell me a bit about that breakthrough? What prompted their change?

Appendix B

Themes	Subcodes	Exemplifying Quotes	Exemplifying Quotes	Thematic Questions
Trust	Trust	"He came up and asked me 'so why did you want to teach this class, cause I like to get a feel for my instructors'"	"I think, well, trust right? I think he felt very, um, he felt like I cared about his engagement in the class and with the material, and maybe it made him trust me because.... OK, I'm not sure I totally understood the sources of his initial distrust, whether, you know the object of his distrust, whether it was me or [the discipline I teach]? Because if it was me then that would make sense, 'cause then the things that he said later would make sense, because what he seemed to be responding to was me showing that I care about him being engaged in the class and, you know, trying to cater to his interests. But, on the other hand, the things that he said at the beginning that were negative, I'm not sure that they, they didn't seem to be so much directed at me as toward [the discipline I teach]."	Why are you here? Is the teacher here for me or for themselves?
		"I just handled class systematically with comments and the residual level if I'm not sure that I really understand your comment, but if you've learned my trust now or I'm going to go with... what the see media other responding to was me showing that it are about him being engaged in the class and, you know, # trying to cater to his interests."	"I think at some point, because there was something about the material that I gave them that I then freaked out about, I wasn't sure if it was OK to give that to them.... anyway so at some point I think I just said to them 'I feel like I've learned some things since I started coming here, but I'm still, there's still a lot of things you know better what the rules are and you know what's OK or not, and you need to tell me if I'm doing something that's stupid, or if I give you something that you shouldn't have, or ...' I mean...just sort of being....there's so much I can learn from you, teaching here, and you have a better sense of where the limits are. And I trust you, also, to tell me if something's off."	
	Caring	"They do not respect teachers who they think are there for résumé building or to write a paper [...] and so I think that that dynamic is really important with respect and trust, because they don't trust people they don't respect, and they're not going to respect you, or trust you, if they think that your motives aren't there."	"I'm sure they do trust me in terms of being somebody who isn't going to give misinformation about psychology. I think that's probably true. Maybe the way that I show that is, if I don't know something I just say I'm not sure but I'm going to look it up and then the next week I do tell them what I've looked up; I'm very careful to do that, because I think it's important that students trust that I'm going to do what I say I'm going to do."	
		"There has been, I think, one instance in which I did an activity where I invited the students to ask me a question about myself and the question needed to be .... Typically I teach a scientific critical thinking class, and to do that I say,' so let's get to know each other; you don't know me, what questions can you ask me the answer to which will inform you whether or not I have the authority to be the instructor? And then if you can give a good reason for why that would be informative, then I'll answer the question.' And only at that time has there ever been students who've sort of challenged, have felt comfortable to challenge, and there, their concern was 'do you care about us?' not, like, 'do you have the credentials?' or whatever ....'do you care about us?'"	I discussed this actually with some of my friends. ...I know some people who are in a down-and-out situation don't trust the supposed selflessness of people who are trying to help them, they think 'you're just helping me because it feels good to you, you don't really care about me,' or 'you're just helping me because you have some political agenda, but really, I don't care about your political agenda, I just want to get out of here, and I want you to help me get out of here,' something like that, you know.	
	Power/ Empowerment	"rather than [...] dogmatically going ahead with whatever they think is going to work; there are science guys who've tried to do that by 'filling the empty cups' and end up with really unhappy students. [...] They don't want to sit and listen for three hours, they want to be active participants, and if an instructor doesn't take advantage of that, student learning I think is jeopardised."	"But, you know, we have an authority as the person in front of the classroom and as the instructor. It's a different type of authority, though. It's an authority over the subject-matter, and over how we're going to approach the material, and generally what we're going to talk about, but I gave them a lot of opportunities to think freely for themselves and to share those ideas. And so, in that sense, even though I was maybe an authority figure in a sense, I was also a liberator in a cognitive sense. You know, a chance for them to really think about something really deeply, and be able to talk about it with other people, including me, that had nothing to do with prison at all, really, a lot of the times."	What are your aims? Do you want to hear what I have to say or just transfer facts/feed your ego?

97

Stage	Stage Description	Example Quote	Application to Prison	The Role of Trust and Power
1: Initial Position	Student comes to the classroom with initial frame/perspective	"I mean, they're in prison, right, everyone around them can exert power on them…other prisoners, guards, the state, the federal government, like everyone has control over them to some extent. They have such little autonomy, and they've extremely wary of the notion that there are people involved in [teaching] because they can get something out of it, for themselves."	Incarcerated students reside in a space that can be disempowering and not conducive to trust.	
2: New Data	Student encounters new and incommensurable data	"I think that their expectation of lots of structure, of lots of things decided for them conflicted with my wanting to empower them and give them some […] power over the direction of their course or their assignments and so I think we've been working to figure that out because I might not have given them the structure that they've used to."	The classroom experience introduces new data, either because they have less educational background or because the classroom environment is so different from that of the prison generally.	
3: Assimilation	Student attempts to fit new data into existing frames	"a lot of the guys are insecure about the legitimacy of academic credentials, […] they write about it in their papers, when they come up with examples, and it also seeps into a lot of the discourse in the class, they're worried that this isn't a real college class, this is just someone coming in and doing their own little thing, but it's not like a real education"	In the prison classroom, assimilation often takes the form of the questioning of intentions and the testing of power roles.	Data from the interviews indicates that trust and power-sharing is important for getting students past the assimilation stage. The student must trust the teacher enough not to throw out the disorientation new information. The student must also be empowered enough to feel comfortable questioning the new data and investigating its role in his or her meaning scheme.
4: Incommensurability	Student realizes that assimilation is impossible within current frames	"I just handed essays back with comments and he said 'well I'm not sure that I really understand your comments but you've earned my trust now so I'm going to go with it.' what he seemed to be responding to was me showing that I care about him being engaged in the class and, you know, trying to cater to his interests."	Students reach this stage if they find that the new data is incommensurable with their existing meaning schemes.	Whether a student reaches the incommensurability stage depends on what goes on at the 'assimilation' stage. If a student successfully assimilates the new data into their existing frames, then incommensurability is not encountered.
5: Altered Frames	Student either rejects new data or adopts new frame/paradigm	Altered Frame "The [incarcerated] students […] have a great deal of appreciation for having classes offered to them, because it takes them so far outside their normal environment. […] at most universities […] I don't think, at least in general, that they see that their professors are people who are there because they care about them or want to see them succeed, they've just there to perform some task that they need to do so that they can get their grades [and] go to law school or whatever." Rationalization "on the exam one of the questions was 'define rationalization and give an example,' and a student gave a brilliant and also really heartbreaking example of somebody who is intimidated by their prison class and, as a result of that, doesn't take the class, and then rationalizes that by saying to himself 'it's not a real class'"	Students either reject new data (rationalization) or adopt new paradigms (altered frames) in order to assimilate the data.	

Why Are We Losing Students? Enrollment and Motivation in Chinese Language Learners

Annetta Fotopoulos

Department of Asian Studies, Cornell University
ajfotopoulos@gmail.com

Introduction

Declining Language Enrollment

Data compiled in the last several years has shown an alarming overall drop in foreign language enrollment at US universities. The latest Modern Languages Association (MLA) report "Enrollments in Languages Other Than English in United States Institutions of Higher Education" that was released to the public on February 11, 2015 illustrated that between 2009 and 2013, U.S. university student enrollment in foreign languages decreased by 6.7% (Goldberg et al. 2015). Although Chinese language enrollment nationwide saw a slight average increase during this period (2%), this was significantly less than previous growth rates. There was a 50.4% growth rate between 2002 and 2009 and 16.5% growth rate between 2006 to 2009, indicating that Chinese language enrollment may be beginning to plateau (ibid., 23).[8]

Furthermore, at Cornell University specifically, data collected by the Language Resource Center has shown an average decline of 8.4% in overall language enrollment between 2008 to 2013 (Feldman 2014). Chinese language enrollment specifically peaked in 2009 and then steadily dropped between 2010 and 2015, an average of 15.6% per year (Feldman 2016). Two Cornell employees who have been following issues of language enrollment closely in the last decade indicated to me that they believed the decline in enrollment at Cornell was consistent with nation-wide trends at universities that had already been offering Chinese. Yet the MLA report showed a slight overall increase in enrollment because many institutions that had not previously offered Chinese have since initiated Chinese language programs. This hypothesis is supported by the MLA report itself, which notes that the number of institutions reporting enrollments in Chinese has more than doubled between 1990 and 2013 (Goldberg et al. 2015).

The MLA report also notes that the ratio in enrollment between introductory and advanced courses in Chinese (defined as first and second year vs. third and fourth year respectively) has been a steady 4:1 over the past three survey periods (2006, 2009 and 2013) (ibid., 8). This may indicate that even if overall enrollment rates have declined for universities with existing programs, retention rates have been relatively consistent, although this is complicated by the fact that some students with pre-existing background in the language may test into upper levels. At Cornell specifically, figures tracking students' completion of successive levels of Chinese each year showed an increase in the ratio between students who had completed beginning versus advanced levels from 4:1 (rounded to the nearest whole number) between 2007-2011 to 5:1 from 2008-2015.[9] This indicates that, contrary to national standards, the retention rate of students progressing to upper levels of Chinese language is declining at Cornell.

The question of why student enrollment in Chinese language has been declining is a significant source of anxiety among Chinese language instructors at Cornell, who have recently been forced to cancel summer intensive language programs due to insufficient enrollment. Having enrolled in numerous Chinese language courses myself including four years of Mandarin as an undergraduate from 2004-2008, two years of Mandarin and two years of Classical Chinese in graduate school from 2008-2010, and four different study abroad programs between 2007-2012), I am also extremely curious about why interest in learning the language at the university level seems to be declining. Thus, when I taught Chinese as a teaching assistant for Beginning Mandarin I at Cornell in the fall of 2015, I decided to conduct a preliminary study on the matter by using survey methods to ask my students about their motivations for enrolling in Chinese, their expectations for the course, and later, their motivations for continuing or discontinuing with the next semester. While my motivation for conducting research was to attempt to account for recent drops in Chinese language enrollment at Cornell, my spe-

---

[8]For information on how this data is collected and what it includes and excludes, see the "Notes on Methodology" section: Goldberg et al. 2015, 11-13.

[9]This is based on data I obtained from the Language Resource Center at Cornell tracking the number of students who had completed successive levels of Chinese language each year. Students who did not start from the beginning level but tested into upper levels are excluded from this data.

cific research questions were: What kind of students enroll in beginning level Chinese?, How does the way Chinese courses are designed and taught affect enrollment?, What are some of the reasons that students begin taking Chinese?, What are some of the reasons that they continue?, What are some of the reasons that they don't continue?, How can we optimize curriculum to attract and retain students?

Motivation and Second Language Learners

Motivation, which I define here as a reason or reasons for pursuing a certain course of action, has long been recognized as a crucial factor in effective learning. In 1972, Gardner and Lambert's groundbreaking study argued that motivation was a key factor in second language acquisition as well (1972), a hypothesis that was supported by subsequent research (Cooper and Fishman 1977; Dörnyei 1994; Pintrich and Maehr 1995). Since then, myriad studies have been conducted on this topic. Several different tests have been developed for measuring motivation in second language learners (Gardner, Tremblay and Masgoret 1997; Horwitz 1987; Kuhl and Beckmann 1994; Oxford 1990; Pintrich et al. 1991). Earlier studies tended to conceptualize motivation as the result of social stimulants and psychological processes affecting the individual. Particularly influential was the notion of integrative motivation, learning in order to better understand, function in, or integrate into the target language group (Gardner, Day and MacIntyre 1992). This has often been compared to instrumental motivation, learning in order to achieve a specific pragmatic goal.

Since the 1990s, the concept of integrative motivation has come under criticism for its inability to capture the situation of many language learners for whom a clearly defined target language group does not exist, or whose motivation to learn a foreign language has little to do with the prospect of integration into another society. Thus in many more recent studies, the focus has been less on motivation as a quantifiable attribute of individuals and more on motivations as multiple, shifting, and the product of dynamic social, cultural, and psychological factors. For example, Dörnyei has proposed a dynamic systems approach to studying language learning, in which "the complex interaction of language, agent, and environment" would be operationalized according to specific theoretical and measurement terms (2009). Noels et al. proposed that language learning motivation be examined from the perspective of self-determination theory, which views the individual as the locus of decision-making about language learning and examines both internal motivation and various types of external motivations, arguing that the most effective motivators are those that support the learner's experience of autonomy, competence and relatedness (2000).

Comanaru and Noels, in a study of Chinese language learners specifically, used survey methods to measure multiple different motivational variables, which they termed intrinsic motivation (feelings of enjoyment), integrated regulation (equivalent to integrated motivation), identified regulation (related to self-identity), introjected regulation (feelings of obligation), external regulation (similar to instrumental motivation), and amotivation (2009). My own study draws heavily on Comanaru and Noels' model, analyzing student motivation to enroll in and continue Chinese language study according to multiple motivational variables.

Methods

Using Qualtrics online survey software, I sent a series of four email surveys to Cornell students enrolled in CHIN 1101 Beginning Mandarin I and CHIN 1102 Beginning Mandarin II at the beginning and end of each semester for the 2015-2016 academic year: two entry surveys in August 2015 and February 2016 and two exit surveys in December 2015 and May 2016. There were 59 students enrolled by the second week of the first semester course and 56 students enrolled by the second week of the second semester course (although of the latter 56, only 47 or 75% were actually students continuing from first semester; the rest tested into the second semester course based on preexisting skill in the language).[10] Response rates to the surveys were as follows: 61% (36 of 59 students) completed the August 2015 survey; 30% (18 of 59 students) completed the December 2015 survey; 32% (18 of 56) completed the February 2016 survey; and 34% (19 of 56) completed the May 2016 survey.[11] Question formats varied, including multiple choice, Likert scale, numbered ranking, and open-ended questions. Each survey had between 11 to 13 questions and asked about the following items: demographics, including gender, class year, major and native language; students' intentions about future Chinese language enrollment and language application; students' motivations for beginning and continuing Chinese class (entry surveys); and students' relative satisfaction with the course (exit surveys).

---

[10]Students who dropped the course in the first week were excluded from the survey since fluctuating enrollment in the first week is a common occurrence for all university courses.

[11]Higher response rates for the first survey were likely due to the fact that I was a teaching assistant at the time and personally made an announcement at the beginning of class encouraging students to take the survey.

Table 1: Demographics of students that completed surveys.

Gender	Aug. 2015	Dec. 2015	Feb. 2016	May 2016	Mean
Female	21/36, 58%	13/19, 68%	14/21, 67%	12/20, 60%	**63.25%**
Male	15/36, 42%	6/19, 32%	7/21, 33%	8/20, 40%	**36.75%**
Class Year	Aug. 2015	Dec. 2015	Feb. 2016	May 2016	Mean
Freshman	8/36, 22%	6/19, 32%	10/21, 48%	10/20, 50%	**38%**
Sophomore	16/36, 44%	9/19, 47%	7/21, 33%	9/20, 45%	**42.25%**
Junior	8/36, 22%	4/19, 21%	4/21, 19%	1/20, 5%	**16.75%**
Senior	1/36, 3%	0/19, 0%	0/21, 0%	0/20, 0%	**0.75%**
Graduate Student	3/36, 8%	0/19, 0%	0/21, 0%	0/20, 0%	**2%**
Major	Aug. 2015	Dec. 2015	Feb. 2016	May 2016	Mean
Undecided	3/36, 8%	1/19, 5%	3/21, 14%	1/20, 5%	**8%**
Decided	33/36, 92%	18/19, 95%	18/21, 86%	19/20, 95%	**92%**
Native Language	Aug. 2015	Dec. 2015	Feb. 2016	May 2016	Mean
English	28/36, 78%	19/19, 100%	19/21, 90%	16/20, 80%	**87%**
Other	8/36, 22%	0/19, 0%	2/21, 10%	4/20, 20%	**13%**

Several circumstances ensured that students enrolled in beginning Chinese classes were highly motivated before even beginning the course. While a certain amount of language study at Cornell is compulsory for all undergraduate students and some graduate students in order to fulfill curricular requirements, the choice of language is up to the individual student. Thus, even those students taking Chinese in order to fulfill curricular requirements had actively chosen to study Mandarin Chinese, ranked as one of the three most difficult languages to learn for English-speakers, as opposed to one of more than thirty other languages offered.[12]

The Cornell Chinese program has a reputation for being rigorous. It requires a permission code for students to be able to enroll, is worth six credits, meets seven times a week (five Chinese immersion sections and two English-language lectures on grammar and culture), has a strict attendance policy, and delivers a heavy homework and examination load (daily assignments, weekly quizzes, and three sets of oral and written exams). It thus requires an enormous intellectual and time commitment, which is made clear from the very first day of class. Therefore, in general, those who choose to take Chinese language courses at Cornell do so because they are highly motivated to do so. Furthermore, because students see each other every day, sections classes are small (five to twelve students), and sections are conducted exclusively in Chinese, students become very familiar with each other quickly and often develop strong social bonds, which is another potential factor in motiva-

tion.

Results

Demographics

In this study, four different surveys were administered at four different points of time and participation was completely voluntary each time. As a result, the group of respondents varied from survey to survey and each group's responses are not necessarily representative of the overall group of students enrolled in beginning Chinese courses. Because respondents were slightly different for each survey, I collected separate demographic data each time on gender, class year, major, and native language and then took the mean percentage to get an idea of the overall demographics. Relatively low standard deviation suggests that demographics were fairly consistent across the four groups of survey participants and there was likely considerable overlap.

All four surveys indicated more female respondents than male by an average ratio of approximately 2:1 (Table 1). Veteran Chinese language teachers at Cornell that I spoke to suggested that gender ratio varied from year to year and so, although there were more female students this year, this was probably not reflective of a greater trend in Chinese language enrollment.

Reported class year of respondents indicated, on average, a majority of freshman and sophomores (collectively about four fifths of the class) some juniors (about one sixth of the class), and a very small number of seniors and graduate students (Table 1). This makes sense for a beginning language class, as most students tend to start taking language in their first two years and advance to higher levels in their later years. The fact that some seniors and graduate students responded to the first survey but none

---

[12]Cornell undergraduate language requirements require either study of a second language to the intermediate level or 11 credits of coursework. Since Beginning Mandarin was six credits per semester, two semesters of Chinese language study were enough to fulfill these requirements.

Table 2: Students' previous experience with Chinese language.

Q: What is your previous experience with Chinese language and/or culture? (select all that apply):	Aug. 2015
I am of Chinese heritage	11/36, **31%**
My close friend or relative is Chinese or speaks Chinese	11/36, **31%**
I have no previous experience with Chinese language or culture	11/36, **31%**
I previously studied Chinese language or culture	8/36, **22%**
I have traveled to China or a place where Chinese is spoken	7/36, **19%**
Other[7]	4/36, **11%**
I like to watch Chinese movies, listen to C-pop or am interested in other aspects of Chinese pop culture	2/36, **6%**
I lived in China or a place where Chinese is spoken	1/36, **3%**

responded to subsequent surveys may indicate that these students subsequently dropped the course or did not enroll in second semester. In my personal experiences as a Chinese language student and teacher, I have noticed that seniors and graduate students are more likely to drop language courses or discontinue, both because course credits are generally not mandatory for these students and because they often have conflicting obligations.

The vast majority, an average of 92%, of respondents had declared majors and were asked to identify these themselves. The most commonly reported (filled in more than once on the same survey) majors were: Computer Science, ILR, Biological Sciences, Math, Engineering, Economics, and Hotel Administration. Notably, Asian Studies and CAPS (China and Asia-Pacific Studies), the only two majors that would specifically require students to take an Asian language (usually Chinese, Korean, or Japanese) were both reported only twice in total, once each on two separate surveys. This seems to indicate that most students taking beginning Chinese are in fields of study not directly related to Chinese language or culture and are probably not primarily motivated, at least initially, by the direct applicability of Chinese to their future careers, although many students who advance through many years of Asian Language ultimately end up taking on an additional minor in Asian Studies or CAPS.

Most students reported that they were native speakers of English (an average of 87%). Of the 13% who were non-native speakers of English, the majority were native speakers of Korean and there were also a few who had grown up speaking a dialect of Chinese such as Cantonese or Teochew. Other native languages reported were Portuguese, Thai, and Spanish. According to the 2015-2016 report from the Cornell International Students and Scholars Office 21.2% of school-wide enrollments and 10.1% of undergraduate enrollments were international students in 2015 (Cor-

nell 2016, 10). Therefore, an average of 13% non-native English speaker respondents does not appear to be significant in the greater context of university demographics. There have been studies suggesting that the relative difficulty or ease of second language acquisition is strongly dependent on the similarity of the student's base language to the target language (e.g., Jorden and Walton 1987). The Asian language program at Cornell specifically tailors pedagogy to U.S. native English speakers by explaining linguistic and cultural nuance in terms of their similarity and difference to U.S. norms.[13] In the context of this study, I did not collect enough data to draw a convincing correlation between motivation to enroll and persist in Chinese language courses and the learner's native language, but this would be a fruitful topic for further studies.

In order to further answer the question who takes Chinese? in the first survey, I also asked students about their previous exposure to Chinese language and culture (Table 2). About one-third of respondents (31%) indicated that they were of Chinese heritage (generally defined as having one or more parents who speaks some dialect of Chinese), while another third (31%) indicated that they had no previous experience with Chinese language or culture. This means that the remaining third were non-heritage learners but had some previous exposure to Chinese language or culture (although this cannot be broken down numerically since the other choices are not mutually exclusive).

There have been numerous studies suggesting that the motivations of heritage learners of Chinese are somewhat different than those of non-heritage learners (Comanaru and Noels 2009; Tsai 2004; Xu and Moloney 2014). In this particular study (only looking at the August 2015 survey), the majority of both

---

[13]I was trained in this methodology during my time at Cornell and observed it in action during lecture classes for both Chinese and Japanese courses.

Table 3: Number of semesters students intended to enroll in Chinese language courses.

Duration	Aug. 2015	Dec. 2015	May 2016	Mean
1 semester	5/36, 14%	-	-	**14%**
1 year	10/36, 28%	8/19, 42%	5/20, 25%	**31.67%**
1.5 years	-	-	1/20, 5%	**5%**
2 years	11/36, 31%	2/19, 11%	0/20, 0%	**14%**
3 years	7/36, 19%	3/19, 16%	3/20, 15%	**16.67%**
4+ years	3/36, 8%	3/19, 16%	7/20, 35%	**19.67%**
Discontinue for now	-	2/19, 11%	2/20, 10%	**10.5%**
Other	-	1/19, 5%	2/20, 10%	**7.5%**

students of Chinese heritage and students with no exposure to Chinese language or culture at all indicated that they had decided to study Chinese because they thought it would be fun (intrinsic motivation) and because they thought it would be useful to their future career (instrumental motivation). However, whereas an equally great motivation for those with no exposure was a future intention to travel to China (a kind of integrative motivation), by far the greatest motivation for heritage students was the desire to communicate with Chinese speakers they knew (another kind of integrative motivation).

Intentions

In three of the four surveys (August 2015, December 2015, and May 2016), I asked students how many Chinese courses they intended to take during their academic career (Table 3). The most frequent answer was one year, most likely because that is the amount needed to fulfill undergraduate language requirements. The second most frequent answer was four or more years, although the number of students who selected this answer rose considerably in the May survey when the second semester class was ending.

In the two exit surveys (December 2015 and May 2016), I asked whether students intended to continue onto the next semester. While 90% (19/21) of respondents indicated in December 2015 that they were planning to continue onto second semester Chinese, 74% (42/57) of students who finished first semester actually continued on to second semester Chinese. In May 2016, 55% (11/20) of students indicated that they planned to continue onto the next semester.

Motivations

Although it is clear that the majority of students enrolled in beginning Chinese language at Cornell were initially highly motivated to take the course, my pilot study aimed to discover both what their specific motivations were and whether these motivations changed as a result of actually taking the course. In order to do this, in each survey, I included a multiple-choice question about why students were taking Chinese and divided them up into categories (Tables 4 and 5).

In the initial survey I asked why students had decided to take Chinese (Table 4). The most popular answer was utility in future career (instrumental motivation) followed closely by the desire to communicate with Chinese speakers (integrative motivation) and anticipation of enjoyment (intrinsic motivation). The least popular answers (chosen by only one respondent each) were that it was related to the student's major (instrumental motivation) and that the student felt obligated by someone else to take the course (introjected motivation).

In the subsequent three surveys, administered in December 2015, February 2016 and May 2016, I asked about students' reasons for continuing Chinese class (for December and May because they had expressed the intention of continuing; for February because they were enrolled in the new semester). By far the most popular answer was enjoyment (intrinsic motivation) with other answers coded for intrinsic motivation also chosen by more than half of participants.

Also very popular, selected by over 70% on average, were the two integrative options: a desire to talk to Chinese speakers and a desire to visit Chinese-speaking places. The least selected types of motivation were Chinese heritage (identity-based), although this was still selected by a quarter of the students, and a direct relation to the student's major (instrumental), which supports demographic data indicating that most students were not Asian Studies or CAPS majors. The motivations that (based on high standard deviation) appeared to change the most over time were fondness for the teachers (students seemed to like their teachers more first semester than second semester), job utility (initially a major motivator only for a third of students; by the end of second semester a motivator for nearly two thirds) and curricular requirements (after completing two semesters of Chinese class the undergraduate language requirement

Table 4: Students' initial motivations for enrolling in Chinese language courses.

Q: Why have you decided to study Chinese (Select all that apply):	August 15
Instrumental motivation	
I think Chinese will be useful to my future career	67%
I need it to fulfill program requirements	31%
I plan to major in a Chine-related field	3%
Integrative motivation	
I want to know how to communicate with someone I know who speaks Chinese	58%
I plan to travel to China and would like to know some Chinese	44%
Intrinsic motivation	
I was curious or thought it would be fun	53%
Identity-based motivation	
My previous study or experiences have gotten me interested in Chinese language	36%
Introjected motivation	
I am being pressured by my family or someone else to take Chinese	3%
Other[8]	8%

was fulfilled so students intending to continue onto the next year were no longer motivated by this).

In the two exit surveys (December 2015 and May 2016), I also asked students who had indicated that they were not planning to continue onto the next semester of Chinese their reasons for discontinuing Chinese language classes. However, because response rates were very low (two students for the first semester and seven for the second semester), the data is insufficient to help inform us about how motivations changed from first semester to second semester. Based on the limited data, time commitment seems to be the biggest factor in discontinuing. Somewhat more telling are the options that were not selected by any students: graduation, grades, and lack of enjoyment. The data seems to indicate that most students quit for pragmatic reasons rather than because the course failed to deliver utility or enjoyment.

As a supplement to multiple choice and Likert scale-type questions, I also included three open-ended questions at the end of the last survey asking students to describe their initial motivations for taking Chinese, how these had changed, and what role Chinese would have in their future. Responses to the initial question were quite varied, but can be largely subsumed under the three main categories of instrumental (for language requirements or job success), identity-based (on account of heritage or a previous connection to the language) and intrinsic (based on interest, curiosity or love of the language). Answers to the second question about how motivations had changed overwhelmingly focused on integrative motivations (desire to speak to friends and function in Chinese-speaking places) and intrinsic motivations (intensified appreciation and enjoyment of the lan-

guage). Answers to the final question about the role of Chinese in each student's future were centered around three main themes: living in or visiting China, working in China or at a China-related job and communicating with friends and relatives.

Satisfaction

In order to gauge students' relative satisfaction with beginning Chinese courses and the extent to which they felt that what they were learning was useful, I asked a series of Likert scale questions about how useful they thought Chinese would be to them, how much they enjoyed taking Chinese class (for the initial Aug. 2015 survey I asked instead how enthusiastic they were to take Chinese class), and how effective each semester-long Chinese class had been in developing their Chinese skills (Table 7). For all of these items, the average rating was consistently above four for all four surveys, suggesting that students generally felt that their Chinese skills were useful and that the Chinese courses were both enjoyable and effective.

Discussion

The demographic data I collected suggests that students who enrolled in beginning Chinese language courses for the 2015-2016 academic year were majority female, undergraduate freshman and sophomores, non-Asian Studies majors, and native speakers of English. The characteristic that stands out in terms of Chinese language enrollment is the fact that most students were from majors outside of Asian Studies. According to data collected by Cornell's Language Learning Center, from 2009-2015, an average of 66% percent of students enrolled in Chinese classes at all levels were in Asian Studies. This suggests that many

Table 5: Students' motivations for continuing to take Chinese language courses after fulfilling curriculum requirements for languages.

Q: What are your reasons for continuing Chinese class? (Select all that apply):	Dec. 2015	Feb. 2016	May 2016	Mean
Intrinsic motivation				
I enjoy it	15/17, 88%	14/18, 78%	9/11, 82%	**82.67%**
I really like the language	13/17, 76%	13/18, 72%	8/11, 73%	**73.67%**
I find it highly rewarding[10]	12/17, 71%	11/18, 61%	-	**66%**
I really like my teachers	13/17, 76%	8/18, 44%	5/11, 45%	**55%**
Integrative motivation				
I want to be able to use it to talk to Chinese-speakers	-	-	8/11, 73%	**73%**
I want to be able to function in a place where Chinese is spoken	-	-	9/11, 73%	**73%**
Instrumental motivation				
It will probably help me get a job	6/17, 35%	5/18, 28%	7/11, 64%	**42.33%**
I need it to fulfill requirements	8/17, 47%	9/18, 50%	2/11, 18%	**38.33%**
It's directly related to my major	0/17, 0%	2/18, 11%	2/11, 18%	**9.67%**
Introjected motivation				
My family or friends want me to continue	5/17, 29%	8/18, 44%	4/11, 36%	**36.33%**
I feel like I should continue	-	-	3/11, 27%	**27%**
Identity-based motivation				
It's part of my heritage so I think I should continue	6/17, 35%	4/18, 22%	2/11, 18%[11]	**25%**
Other	1/17, 6%	2/18, 11%	1/11, 9%	**8.67%**
I think it will be useful to me in the future	14/17, 82%	14/18, 78%	-	**80%**

Table 6: Students' motivations for not continuing to take Chinese language courses after fulfilling curriculum requirements for languages.

Q: What are your reasons for dis-continuing Chinese class? (Select all that apply):	Dec. 2015	May 2016	Mean
It's too much of a time commitment	2/2, 100%	4/7, 57%	**78.5%**
Schedule conflict	2/2, 100%	1/7, 14%	**57%**
It's not really related to my major	1/2, 50%	4/7, 57%	**53.5%**
I don't like the way the program is run	1/2, 50%	1/7, 14%	**32%**
I have already fulfilled the requirements I need	0/2, 0%	4/7, 57%	**28.5%**
It's too difficult	1/2, 50%	0/2, 0%	**25%**
I don't think it's useful to me	1/2, 50%	0/7, 0%	**25%**
I am graduating or moving	0/2, 0%	0/2, 0%	**0%**
I didn't get a high enough grade	0/2, 0%	0/2, 0%	**0%**
I don't enjoy it	0/2, 0%	0/2, 0%	**0%**
Other	0/2, 0%	0/2, 0%	**0%**

Table 7: Perceived course effectiveness, enjoyment, and language utility, as indicated in student surveys.

Q1: On a scale of 1-5, how useful do you think knowing Chinese will be to you in the future? (1=useless, 5=indispensable)	Aug. 2015 36 responses	Dec. 2015 19 responses	Feb. 2016 18 responses	May 2016 20 responses
Mean rating on Likert scale:	4.14	4.32	4.06	4.20
Q2a: On a scale of 1-5, how enthusiastic are you to learn Chinese? (1=don't really want to learn it, 5=can't wait to learn it)	Aug. 2015 36 responses	-	-	-
Mean rating on Likert scale:	4.44	-	-	-
Q2b: On a scale of 1-5, how much did you enjoy taking Beginning Chinese? (1=you hated it, 5=you loved it)	-	Dec. 2015 19 responses	Feb. 2015 17 responses	May 2016 20 responses
Mean rating on Likert scale:	-	4.32	4.06	4.64
Q3: On a scale of 1-5, how effective in developing your language skills did you think Beginning Chinese was? (1=completely ineffective, 5=extraordinarily effective)	-	Dec. 2015 19 responses	Feb. 2015 17 responses	May 2016 20 responses
Mean rating on Likert scale:	-	4.53	4.18	4.55

students who enroll in beginning Chinese class do not initially begin with a major or minor in Asian Studies, but that a significant portion of those who continue into higher levels pick up a major or minor in the field, enough that the majority of students enrolled in Chinese language courses are in Asian Studies. What this means is that many students, rather than enrolling in Chinese in order to fulfill the instrumental goal of requirements for an Asian Studies major, are doing the reverse and likely adding or switching to an Asian Studies major because of their previous enrollment in Chinese classes. In fact, this describes my own past scenario as an undergraduate at a U.S. university taking Chinese language courses perfectly; initially I enrolled out of personal interest and later ended up adding a second major in Asian Studies.

The fact that approximately a third of respondents were heritage learners, a third had no experience with Chinese culture and language at all, and a third (by deduction) were not heritage learners but had some kind of previous exposure to Chinese language or culture suggests that beginning Mandarin Chinese at Cornell is broadly appealing to students with diverse backgrounds and experiences, not just to those with a preexisting connection to the language.

The survey results about intentions to enroll in further Chinese classes were also telling, not because they can actually accurately predict future enrollment, but because they seem to indicate that after completing a full year of Chinese courses in May 2016, students were more likely to aspire to reach advanced levels (four or more years) of Chinese language study, with 55% indicating they were planning on continuing past the first year and and 35% aspiring to four or more years as opposed to in August 2015 when the first semester was just starting and 58% expressed the intention to continue past the first year, but only 8% to continue through four or more years. This, along with an actual second semester retention rate of 74%, would suggest that the first year language courses are relatively successful at motivating students to continue to pursue Chinese language studies. This hypothesis is corroborated by data on course satisfaction, in which respondents in all four surveys overwhelmingly rated the effectiveness and enjoyment of Chinese classes highly and felt that knowing Chinese was useful.

The hypothesis that beginning Chinese courses were largely successful in cultivating student motivation to continue to enroll in Chinese courses is also largely supported by the data on motivation, which showed that initially students enrolled primarily on account of a combination of instrumental, intrinsic and integrative motivations, but later continued because of overwhelmingly high intrinsic and integrative motivation. Instrumental motivation in the relevance of Chinese to a future career still remained a significant factor, but seems to have become secondary to intrinsic and integrative motivation. Although data on students who did not intend to continue Chinese language study was sparser, not a single one of these students indicated that they did not enjoy the course and several noted that they were planning on picking up with Chinese again in the future.

Based on preliminary findings from this broad pilot study, my hypothesis is that recent drops in Chinese language enrollment at Cornell are either due to drops in initial recruitment of students or lowered retention at upper levels beyond the second year, as first year courses appear to be relatively successful at cultivating student motivation to continue. The latter hypothesis is supported by figures showing a relative decline in enrollments in advanced as opposed to beginning Chinese courses at Cornell (Cf. Footnote 2). At the present, my recommendation is that the Chinese program at Cornell increase emphasis on attracting and recruiting new students by initially emphasizing the utility of Chinese for communication and career, the vast number of Chinese speakers both in the U.S. and worldwide, study abroad and other visiting opportunities to Chinese-speaking places, and the immense satisfaction derived from learning a language often regarded as "difficult."

Methodologically, this study, which was heavily focused on quantitative data, could benefit from more open-ended questions, focus groups, or personal interviews to provide a more detailed picture of students' complex motivations for enrolling in, continuing and discontinuing language study. Further studies on this topic would benefit from surveying a broader range of students beyond just first year, a more stable group of participants whose motivational changes could be specifically tracked over time, more careful coding of response options to different types of motivation, and a larger sample of students who discontinue language study.

Acknowledgments

References

Comanaru, R. and Noels, K.A. 2009. "Self-Determination, Motivation, and the Learning of Chinese as a Heritage Language." The Canadian Modern Language Review 66 (1): 131-158.

Cooper, R.L. and Fishman, J.A. 1977. "A Study of Language Attitudes." in J.A. Fishman, R.L. Cooper, and A.W. Conrad, eds. The Spread of English. Rowley, MA: Newbury House. 239-276.

Cornell University International Students and Scholars Office. 2016. 2015-2016 Annual Statistics. Accessed 5 June 2016. http://isso.cornell.edu/sites/isso.cornell.edu/files/isso/documen 16.pdf

Dörnyei, Z. 1994 "Motivation and Motivating in the Foreign Language Classroom." The Modern Language Journal 78 (3): 273-284.

Dörnyei, Z. 2009. "The L2 Motivational Self System." in Dörnyei, Zoltan and Ema Ushioda, eds. Motivation, Language Identity and the L2 Self. Clevedon, GBR: Multilingual Matters, 2009. 9-42.

Feldman, D. 2014. Foreign Language Enrollments, 2003-2013. Unpublished raw data.

Feldman, D. 2016. Cornell Chinese Enrollment. Unpublished raw data.

Gardner, R.C., Day, J.B., and MacIntyre, P.D. 1992. "Integrative Motivation, Induced Anxiety, and Language Learning in a Controlled Environment." Studies in Second Language Acquisition 14 (2): 197-214.

Gardener, R.C. and Lambert, W.E. 1972. "Attitudes and Motivation in Second Language Learning." Rowley, MA: Newbury House Publishers.

Gardner R.C., Tremblay, P.F., and Masgoret, A.M. 1997. "Towards a Full Model of Second Language Learning: An Empirical Investigation." The Modern Language Journal 81: 344–362.

Goldberg, D., Looney, D., and Lusin, L. 2015. Enrollments in Languages Other Than English in United States Institutions of Higher Education. Modern Language Association of America.

Horwitz, Elaine. 1983. Beliefs About Language Learning Inventory. Austin, TX: University of Texas. Aptitude test.

Jorden, E. and Walton, A.R. "Truly Foreign Langauges: Instructional Challenges." Annals of the American Academy of Political and Social Science: Foreign Language Instruction: A National Agenda 490: 110-124.

Kuhl, J., and Beckmann, J. 1994. "Volition and Personality: Action Versus State Orientation." Seattle: Hogrefe and Huber Publishers.

Noels, K.A., Pelletier, L.G., Clément, R., and Vallerand, R.J. 2000. "Why Are You Learning a Second Language?" Motivational Orientations and Self-Determination Theory." Language Learning 50 (1): 57–85.

Oxford, R. 1990. Language Learning strategies: What Every Teacher Should Know. New York: Newbury House.

Pintrich, P.R. and Maehr, M.L. 1995. Culture, Motivation and Achievement. Greenwich, CN: Jai Press.

Pintrich, P.R., et al. 1991. A Manual for the Use of the Motivated Strategies for Learning Questionnaire. Ann Arbor, MI: National Center for Research to Improve Postsecondary Teaching and Learning.

Tsai, S.C. 2004. A Comparative Study of American-born Chinese and Non American-born Chinese Students' Language Learning Motivation, Language Acquisition, and Ethnic Identity Development in Chinese Language Schools. Doctoral Dissertation. Old Dominion University. May 2004.

Xu, H.L. and Moloney, R. 2014. "Identifying Chinese Heritage Learners' Motivations, Learning Needs, and Learning Goals: A Case Study of a Cohort of Heritage Learners in an Australian University." Cercles 4 (2): 365-393.

Gender Differences in Students' Sex-Specific Social Comparisons and Goal Adoption for STEM and Non-STEM Majors

Emily J. Hangen

Department of Clinical and Social Sciences in Psychology, University of Rochester
emily.hangen@rochester.edu

Abstract

Over 600 university students from both STEM and Non-STEM disciplines were surveyed about their achievement goal adoption, their normative performance expectations (i.e. how much they expect to perform better than their male and female classmates), and the degree that their goals focus on performing better (or avoiding performing worse) than their male classmates and their female classmates. Findings reveal that despite opposite gender stereotypes, students generally compare to classmates of their same-sex regardless of being in a STEM or Non-STEM discipline. However, across all students, greater comparisons to females was associated with performance-approach goal adoption and greater comparisons to males was associated with performance-avoidance goal adoption. Furthermore, within STEM majors performance-approach comparisons to male classmates were associated with a higher GPA, but performance-approach comparisons to female classmates were conversely associated with a lower GPA even after controlling for actual goal adoption.

Introduction

Do college students compare their performance in school to their classmates? And do gender performance stereotypes in science, technology, engineering and mathematics (STEM) influence whether students compare more or less to classmates of a certain gender? And of practical importance, how does comparing to same-sex, or other-sex, classmates influence whether students adopt adaptive or maladaptive goals and influence their school performance? These are a few of the questions examined in the present research. The purpose of the present research was to combine research on negative female STEM stereotypes, achievement goals, and social comparison theory to better understand to whom college students compare themselves to, how comparisons are affected by stereotypes, and how sex-specific comparisons are association with students' goal adoption and academic achievement.

Theoretical Framework and Previous Research

Performance goals.

Performance goals are a subset of achievement goals that focus on one's relative competence, or how well one is doing as compared to others (For example, "Did I get a better exam score than my classmates?"; Elliot & McGregor, 2001; Elliot, McGregor, & Gable, 1999). There are two types of performance goals: performance-approach and performance-avoidance. Although both of these goals measure competence by using normative standards, they differ as to whether a person is focused on achieving a positive outcome (performance-approach goal) or focused on simply avoiding a negative outcome (performance-avoidance goal; Elliot & Harackiewicz, 1996).

The distinction between performance-approach and performance-avoidance goals is critical because these goals lead to opposite outcomes. Performance-approach goals are generally adaptive, being associated with deep study strategies, persistence, effort, self-efficacy, intrinsic motivation, and better grades, Performance-avoidance goals are consistently shown to be maladaptive, being associated with greater evaluation anxiety, self handicapping, avoidance of help-seeking and lead to poorer performance (see meta-analyses by Hulleman, Schrager, Bodmann, & Harackiewicz, 2010; Lochbaum & Gottardy, 2015).

Presumably, the difficulty level of reaching one's performance goal will vary widely depending on whom one is comparing to (e.g. trying to do better than the class know-it-all versus trying to do better than the class clown), Yet research has not focused on the target comparisons of people's performance goals and whether some people may switch between adaptive performance-approach goals and maladaptive performance-avoidance goals by modifying what group they select as their target comparison. However, choosing whom to compare to has received a great deal of attention in social comparison theory.

Social comparison theory and performance goals.

According to social comparison theory, people compare themselves to others as a means to assess their own opinions, beliefs, and capabilities (Festinger, 1954). Although social comparison the-

ory has focused heavily on the motive for seeking comparison (self-evaluation or self-enhancement) and on the choice of comparison (upward or downward comparison), less attention has been given to the consequences of these comparisons (see Suls & Wheeler 2014 for review). Because the outcomes of performance goals have been heavily studied, examining the relations between social comparison with performance-approach and performance-avoidance goals could greatly inform how social comparisons impact student's performance.

Thus, the current study examines the relation between the target of comparison and the type of performance goal (approach or avoidance) adopted by having students report who they are comparing themselves to for each performance goal separately. By measuring who students compare to for their performance-approach and performance-avoidance goals separately also allows examination of whether the proxy model of social comparison holds true between both goals. The proxy model of social comparison developed from social comparison theory and suggests that people seek to determine their capabilities on a task by looking to a person (a proxy) who shares performance-relevant characteristics with them and who has already performed the task in question (Wheeler, Martin, & Suls, 1997). According to this model, students in school are likely to compare to classmates of the same sex. However, this may not hold true for females in male-dominated fields.

Social comparison theory and the gender gap in STEM.

Females are under-represented and face negative stereotypes in STEM fields (Alper, 1993; Halpern et al., 2007; National Science Foundation, 2009). Because these fields are male dominated, it is hypothesized that female students who assimilate into STEM fields may more often compare themselves with their male, STEM-stereotypical classmates, relative to the frequency that female students compare themselves to their male classmates in Non-STEM disciplines.

Research Questions

The current research addresses questions regarding whether students compare more to classmates of their same sex, how this may differ between performance-approach and performance-avoidance goals, and between STEM and Non-STEM female students. Specifically, it is hypothesized that 1) in general, students compare to classmates of their same sex (as expected from social comparison theory), 2) students compare to different classmates between their approach and avoidance goals, and 3) females in

STEM fields compare more to their male classmates than do females in Non-STEM disciplines. Furthermore, exploratory analyses will examine whether students' sex-specific comparisons (comparing to females, or comparing to males) predict adoption of performance-approach or performance-avoidance goals and how these comparisons may be associated with students' GPA.

Methods

Procedures

635 undergraduates (Mage = 20.0 years, SDage= 1.42 years; 68% female; 48% white, 31% Asian/Asian-American, 8% Hispanic, 7% Black, 6% other) were recruited for an online study. Both STEM and Non-STEM majors were well represented with 375 students majoring in STEM disciplines and 260 students majoring in Non-STEM disciplines [14]. In a single online session, students reported their major (or intended major), performance goal adoption, goal-specific sex comparisons, their sex-specific performance expectancies, their perceived gender ratio of their major courses, and various demographics.

Measures

Performance goals.

Elliot and Murayama's (2008) 12-item Achievement Goal Questionnaire-Revised (AGQ-R) was used to assess students' performance goals for their major courses. Participants responded on a 1 (not at all true of me) to 7 (very true of me) scale, and the respective items were averaged to form the performance-approach (α= .90) and performance-avoidance (α= .92) goal indexes.

Goal-specific and general sex comparisons.

For each performance goal separately, students were asked to report to what extent they were comparing to their male classmates and to what extent they were comparing to their female classmates. They used a 1 (not at all) to 5 (extremely) scale in response to, "To what extent were you thinking of other [male/female] students as the 'other students'?" [15]. For non-goal specific (i.e. general) sex comparisons, responses across same sex comparison were averaged to create overall index of comparison to females (α= .86) and comparison to males (α= .90).

---

[14]Majors were coded as a STEM or Non-STEM major on the basis of whether a bachelor in science or a bachelor in arts degree would be awarded.

[15]Students were also asked to what extent they were comparing to male and female students in the same major, in the same discipline (STEM or Non-STEM) excluding their major, and to students from another discipline (STEM or Non-STEM)
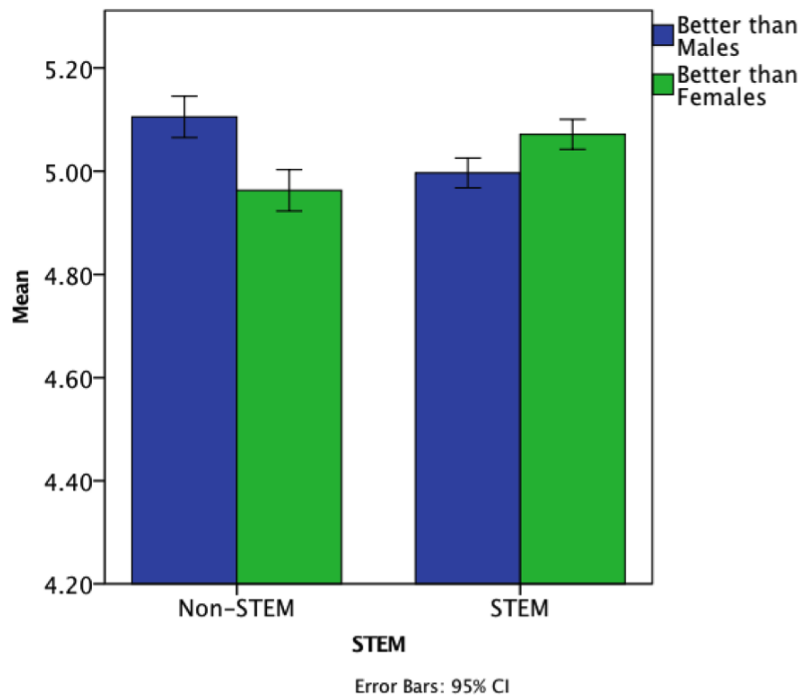
Figure 1. Mean ratings of expecting to perform betterthan male and better than female classmates between STEM and Non-STEM majors.

Sex-specific performance expectancies.

Students were given a single item measuring how well they expected to perform in their major courses as compared to their male classmates, and a single item for how well they expected to perform as compared to their female classmates, "How do you think you will do in you [major] classes compared to other [male/female] students?". Students responded on a 1 (very poorly) to 7 (very well) scale[16].

Perceived gender ratio, GPA, and demographics.

Students reported their perceived gender ratio of their major classes with a 1 (Almost entirely male students) to 7 (Almost entirely female students) scale. Students reported demographics (sex, age, ethnicity) and their current college GPA.

Results

Gender Differences in STEM

As expected, students in STEM majors perceived significantly less females (relative to males) in their major courses ($M = 3.77$, $SD = 1.13$) than students in non-STEM majors ($M = 4.71$, $SD = 1.08$, $t(633) = -10.63$, $p < .001$). Also, STEM majors expected

---

[16]Students were also asked to report their normative performance expectancies as compared to male and female students in the same major, in the same discipline (STEM or Non-STEM) excluding their major, and to students from another discipline (STEM or Non-STEM).

to perform better compared to their female classmates ($M = 5.03$, $SD = 1.20$) than compared to their male classmates ($M = 4.95$, $SD = 1.23$, $t(374) = 2.53$, $p = .012$), suggesting a negative female stereotype. There was a significant interaction of students' sex-specific performance expectancies and students' discipline (STEM, Non-STEM), suggesting that this negative female stereotype is unique to STEM majors ($F(1, 631) = 20.14$, $p < .001$) (Figure 1). In fact, Non-STEM majors had a reverse positive female stereotype, expecting to perform better than compared to their male classmates ($M = 5.17$, $SD = 1.14$) than compared to their female classmates ($M = 5.02$, $SD = 1.10$, $t(259) = -3.49$, $p < .001$).

Sex of Comparison Targets

A significant student sex by sex-specific comparison interaction revealed that students compare more to their same sex ($F(1, 631) = 17.36$, $p < .001$) (Figure 2). This interaction did not differ between STEM and Non-STEM majors ($F(1, 631) = .009$, $p < .925$).

Goal-Specific Sex Comparisons

To test for differences between which comparison sex was selected between goals, we tested a within-subject interaction between sex-specific comparison and performance goal type. The interaction between sex-specific comparison and goal type was non-significant but trending ($F(1, 631) = 2.87$, $p = .091$). Exploratory analyses revealed that females were used
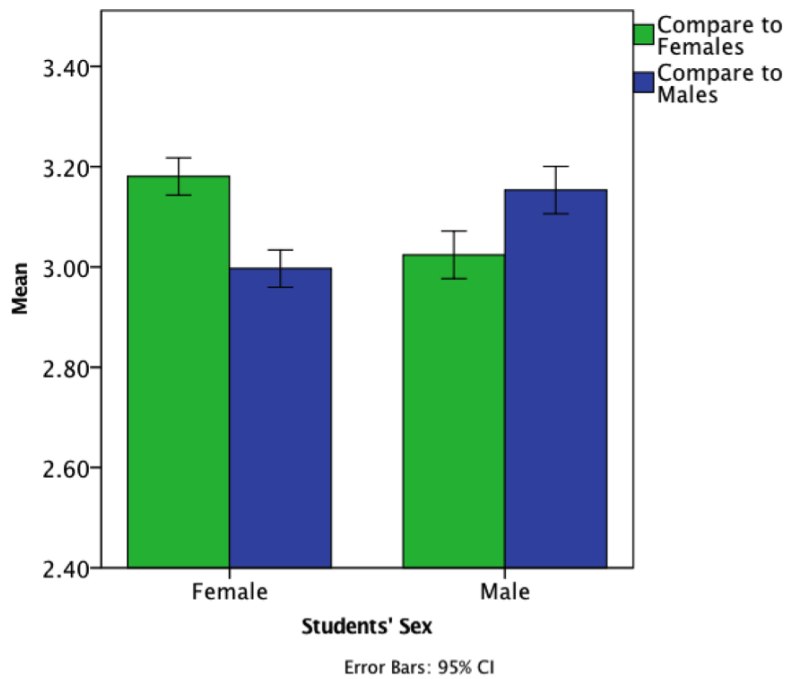
Figure 2. Mean ratings of sex-specific comparison between male and female students.

as the comparison group for performance-avoidance goals more than for any other performance goal comparison sex (ps <.001) (Figure 3).

Performance Goal Adoption

Sex-by-discipline two-way ANOVAs were conducted for both performance-approach goals and performance-avoidance goals. Discipline significantly predicted goal adoption, with students in STEM majors reporting more performance-approach ($F(1, 637)$ = 7.35, p = .007), and performance-avoidance goals than non-STEM majors ($F(1, 638)$ = 12.92, p = .016). Neither students' sex nor an interaction between sex and discipline (STEM, Non-STEM) predicted goal adoption.

Next, each performance goal was regressed onto expectations compared to males (mean-centered) and expectations compared to females (mean-centered) controlling for the other performance goal and for sex (-1 female, 1 male)[17]. The expectations compared to males significantly predicted goal adoption whereas expectations to females did not. The more students expected to perform better than their male classmates the more they adopted performance-approach goals (ß = .152, $t(630)$ = 2.74, p = .006) and the less they adopted performance-avoidance goals (ß = -.166, $t(630)$ = -2.88, p = .004).

The same regression models were then re-run after replacing sex-specific performance expectations

with sex-specific comparisons[18]. When controlling for the influence of performance-avoidance goals, performance-approach goals were significantly associated with sex (ß = .072, $t(630)$ = 2.45, p = .015), suggesting male students adopt more. Also, when students compared themselves more to female classmates the more likely they were to adopt performance-approach goals (ß = .103, $t(630)$ = 2.32, p = .021). Comparisons to males did not predict performance-approach goals, but they were the only predictor of performance-avoidance goals, with the more students compared to male classmates the more they adopted performance-avoidance goals (ß = -.092, $t(630)$ = 2.05, p = .041).

Academic Performance

Influence of goals and gender differences between majors on GPA.

To examine the influence of goal adoption, GPA was regressed onto both performance goals simultaneously to examine each goal's unique contribution with students' sex (-1 female, 0 male) and discipline (-1 Non-STEM, 1 STEM) included as controls. As expected, performance-approach goals were positively associated with GPA (ß = .302, $t(571)$ = 5.41, p <.001), and performance-avoidance goals were marginally, negatively associated with GPA (ß = -.102, $t(571)$ = -1.83, p = .067) (Table 1). To

---

[17]No significant interactions with sex emerged so interactions were excluded from the regression models.

[18]Again, no significant interactions with sex emerged so all interactions were excluded from the regression models.
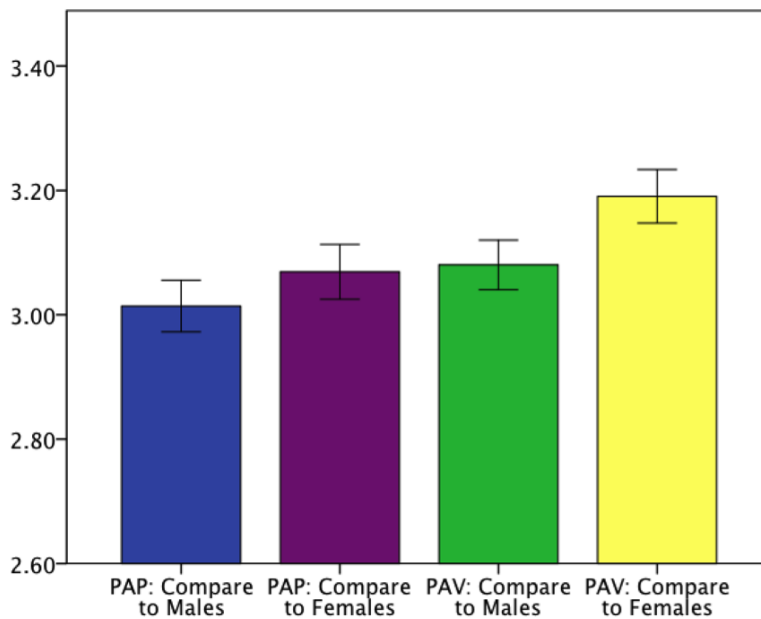
Figure 3. Mean ratings for goal-specific sex comparisons.Error bars represent 95% confidence interval.

empirically counter gender stereotypes, a sex by major interaction was then added to the model and it did not significantly contribute to the model (ΔR^2 <.001, $F(1, 570) = 0.28$, $p = .597$) and was not retained.

Influence of goal-specific sex comparisons on GPA.

All four goal-specific sex comparisons were added to the previously described model (GPA regressed onto sex, STEM, performance-approach, and performance-avoidance goals), but these comparisons did not significantly contribute to the model ($\Delta R^2 = .01$, $F(4, 567) = 0.01$, $p = .468$) (Table 2). However, these goal-specific sex comparisons were retained in the model to examine interactions between these with sex and major. Interactions between each of the four goal-specific sex comparisons with sex were tested in the next step. Interactions between goal-specific sex comparisons and sex did not significantly contribute to the model ($\Delta R^2 = .01$, $p < .10$) and were removed from the model. Interaction between goal-specific sex comparisons with discipline (STEM, Non-STEM) marginally contributed to variance in GPA ($\Delta R^2 = .02$, $F(4, 563) = 2.36$, $p = .052$). There were significant interactions of discipline with sex comparisons for both performance-approach goal comparisons (ßs >.17, ps <.05 (Table 2)[19]. Calculation of simple slopes revealed that sex comparisons for performance-approach goals were not significant for Non-STEM majors, ps >.25. However, for students majoring in

STEM disciplines, trying to outperform male classmates was positively associated with GPA (ß = .239, $t(563) = 2.16$, $p = .031$) and trying to outperform female classmates was negatively associated with GPA (ß = -.286, $t(563) = -3.03$, $p = .003$).

Conclusions

Our findings reveal that despite opposite gender stereotypes, students generally compare to classmates of their same-sex regardless of being in a STEM or Non-STEM discipline. However, the more that students in STEM compared themselves to male classmates for their performance-approach goals (i.e., the degree to which they are trying to out perform their male classmates) the higher their GPA, and the more that students in STEM compared to female classmates for their performance-approach goals (i.e. degree they are trying to outperform their female classmates), the lower their GPA, even after controlling for their performance goal adoption. Because sex comparisons were naturally occurring and not experimentally manipulated, we cannot infer causality nor direction of the relation between goal-specific sex comparisons and GPA. It is possible that student's previous performance (GPA) influences who they compare themselves to—a student who is doing poorly in STEM may choose to "aim low" by shifting their attention to outperforming the presumed high-performers to only outperforming the presumed low-performers (i.e. the female classmates who are negatively stereotyped). However, students who believe that females perform worse may set themselves up for greater discouragement when they don't reach

---

[19]Three way interactions were tested to see if sex interaction with discipline by goal-specific sex comparisons and none were significant, ps >.25.

Table 1.Summary of multiple regression analysis on GPA.

Variable	B	95% CI	β	Sig. (p)	B	95% CI	β	Sig. (p)
		Step 1				Step 2		
Sex (-1 female, 1 male)	-.021	(-.060, .017)	-.045	.279	-.022	(-.061, .018)	-.046	.279
Discipline (-1 Non-STEM, 1 STEM)	.039	(.002, .076)	.085	.041	.042	(.004, .079)	.092	.029
Performance-approach goals (PAP)	.101	(.064, .137)	.302	< .001	.100	(.063, .137)	.298	< .001
Performance-avoidance goals (PAV)	-.030	(-.063, .002)	-.102	.067	-.031	(-.064, .001)	-.105	.061
Comparison to males for PAP					.022	(-.037, .081)	.058	.468
Comparison to females for PAP					-.047	(-.100, .006)	-.125	.082
Comparison to males for PAV					-.016	(-.080, .047)	-.041	.615
Comparison to females for PAV					.045	(-.012, .101)	.114	.123

Note. R^2 = .07 for Step 1, R^2 = .08 for Step 2.

Table 2.Summary of multiple regression analysis with interactions on GPA.

Variable	B	95% CI	β	Sig. (p)	B	95% CI	β	Sig. (p)
		Step 1				Step 2		
Sex (-1 female, 1 male)	-.022	(-.061, .018)	-.046	.279				
Discipline (-1 Non-STEM, 1 STEM)	.042	(.004, .079)	.092	.029				
Performance-approach goals (PAP)	.100	(.063, .137)	.298	< .001				
Performance-avoidance goals (PAV)	-.031	(-.064, .001)	-.105	.061				
Comparison to males for PAP	.022	(-.037, .081)	.058	.468				
Comparison to females for PAP	-.047	(-.100, .006)	-.125	.082				
Comparison to males for PAV	-.016	(-.080, .047)	-.041	.615				
Comparison to females for PAV	.045	(-.012, .101)	.114	.123				
Discipline x Comparison to males for PAP					.069	(.011, .128)	.179	.021
Discipline x Comparison to females for PAP					-.068	(-.121, -.014)	-.175	.013
Discipline x Comparison to males for PAV					-.054	(-.118, .009)	-.131	.095
Discipline x Comparison to females for PAV					.026	(-.031, .083)	.065	.364

Note. R^2 = .08 for Step 1, R^2 = .11 for Step 2. All predictors used in interactions were mean centered.

what they mistakenly presumed was an easy target since as supported herein, females performed equally well as males. Alternatively, a third variable such as self-efficacy beliefs may explain the relation between goal-specific sex comparisons relations and students' GPA. Understanding the causal chain is critical for designing educational interventions that can target causal factors, thus future research is needed.

The current findings have both theoretical and indirect practical value. The results contribute to theory by demonstrating the predictive utility of linking performance goals with sex-specific comparison groups. For instance, results found that the type of performance goal adoption was predicted by sex-specific comparisons—across all students, greater comparisons to females was associated with performance-approach goal adoption and greater comparisons to males was associated with performance-avoidance goal adoption. Furthermore, the goal-specific sex comparisons contributed unique predictive utility above goals alone. These findings offer indirect practical value by calling attention to social comparison as a potential arena for performance goal and school performance intervention.

References

Alper, J. (1993). The pipeline is leaking women all the way along. Science, 260, 409-411.

Elliott, A.J., & Harackiewicz, J.M. (1996). Approach and avoidance achievement goals and intrinsic motivation. Journal of Personality and Social Psychology, 73, 171-185.

Elliot, A.J., & McGregor, H.A. (2001). A 2 × 2 achievement goal framework. Journal of personality and social psychology, 80, 501.

Elliot, A.J., McGregor, H.A., & Gable, S. (1999). Achievement goals, study strategies, and exam performance: A mediational analysis. Journal of educational psychology, 91, 549-563.

Elliot, A.J., & Murayama, K. (2008). On the measurement of achievement goals: Critique, illustration, and application. Journal of Educational Psychology, 100, 613-628.

Festinger, L. (1954). A theory of social comparison processes. Human relations, 7, 117-140.

Halpern, D.F., Benbow, C.P., Geary, D.C., Gur, R.C., Hyde, J.S., & Gernsbacher, M.A. (2007). The science of sex differences in science and mathematics. Psychological Science in the Public Interest, 8, 1–51.

Hulleman, C.S., Schrager, S.M., Bodmann, S.M., & Harackiewicz, J.M. (2010). A meta-analytic review of achievement goal measures: Different labels for the same constructs or different constructs with similar labels?. Psychological bulletin, 136, 422-449.

Lochbaum, M., & Gottardy, J. (2015). A meta-analytic review of the approach-avoidance achievement goals and performance relationships in the sport psychology literature. Journal of Sport and Health Science, 4, 164-173.

National Science Foundation. Division of Science Resources Statistics. (2009). Women, minorities, and persons with disabilities in science and engineering: 2009 (NSF 09-305). Arlington, VA.

Suls, J., & Wheeler, L. (2014) On the Trail of Social Comparison. In The Oxford Handbook of Social Influence.

Wheeler, L., Martin, R., & Suls, J. (1997). The proxy model of social comparison for self-assessment of ability. Personality and Social Psychology Review, 1, 54-61.

The Teaching and Learning of Scientific Controversies in an Online Setting

Kathleen Hefferon

Department of Food Sciences, Cornell University
klh22@cornell.edu

Abstract

Information (or misinformation) regarding agricultural biotechnology has resulted in a renewed interest by educators in the teaching of complex socio-scientific issues surrounding contentious topics. Classroom discussions can help students develop the skills necessary to improve the quality of their arguments, to critically analyze material in a variety of formats to generate a defendable position, and to communicate this position effectively to one another. A Cornell course that covers the socio-scientific controversies surrounding the topic of genetically modified organisms (GMOs) has recently been turned into a Massive Open Online Course (MOOC). The MOOC provided videos, a discussion board, live interviews and quizzes, and is available to anyone around the world who has access to the Internet.

This paper examines how the MOOC instructors addressed the issue of 'instructor bias' in an online version of the course, "The Science and Politics of the GMO". This was achieved by interviewing the MOOC instructors with a series of carefully crafted questions to identify how they intended to provide a balanced viewpoint when teaching a controversial topic. The paper also follows the evolution of individual opinion towards a controversial topic through an analysis of student contributions to the MOOC discussion boards using both EdX and Facebook formats. Coding trees were developed based on the data collected from both students and instructors.

The results of this study demonstrate that instructors could avoid bias by directly challenging one another online and by developing discussions around topics for which there is no one right answer. Analysis of the discussion boards indicated that EdX and Facebook formats yielded argumentative structures which were vastly different. The study found that the MOOC was largely successful; by experiencing many different aspects of the GMO controversy, students left the course more nuanced in their thinking and appreciative of the complexity of the topic. While the online discussions were integral to the course's success, the type of technological format plays an important role with respect to the level of student engagement and the quality of their dialogue.

Introduction

It is striking that only about 1% of space in science textbooks is devoted to any discussion of science controversy and that the majority of the public hold the perspective that scientific knowledge is unchanging (Driver et al., 1996). For scientists working in fields where controversy is common, such as global warming, renewable energy, and genetically modified organisms (GMOs), the public debate is now challenging the legitimacy and credibility of their efforts. Scientists feel compelled to defend themselves and engage in a struggle over the public view of science. The March for Science, held in April 2017 in Washington, DC, has been taken as a battle cry for the scientific community to demonstrate and 'fight back' against poorly constructed policies based on ignorance and misinformation.

Certainly, one of the most controversial biotechnology policy issues today concerns the development and consumption of genetically modified foods (Legge Jr. and Durant, 2010). Public discomfort or outright hostility regarding this controversy has been interpreted by many scientists as merely a matter of scientific literacy, with the implied assumption that improved scientific communication would lead to a better-informed public that is better able to understand and address the issues surrounding GMOs. The American Association for the Advancement of Science describes a scientifically literate person as one who 'uses scientific knowledge and scientific ways of thinking for individual and social purposes' (1990: ix). The scientific community has thus made an effort to reach out and engage with the public regarding current issues of contention in science. For example, in October 2011 41 Swedish scientists from seven universities sent an open letter to politicians and environmentalists that encouraged a revision of European legislation based on scientific assessments

of GM technology. The letter was endorsed by more than 450 scientists from Europe (Moloney et al., 2012). Ethical risks concerning the creation of new crop traits not found in nature by genetic modification, and the precautionary principle regarding too little information on the impact of GMOs on human health and the environment, are favored by the European regulatory system and are some of the reasons why GMOs remain under a moratorium in most European states.

Some of the arguments for genetic modification include the development of high-yielding, pest and disease resistant crops that can feed a growing world population while also mitigating new stresses created by climate change, including drought and flooding. More nutritious, environmentally friendly food is also considered to be a benefit of genetic modification. Opponents to GMOs on the other hand offer a variety of arguments such as health and environmental risks, intellectual property rights, and global justice issues. It is the collective complexity of these social and scientific issues which makes the topic of GMOs contentious and important to include in scientific literacy goals at all levels of education.

The arguments for and against GMOs are often brought up in discussion groups at many universities as a means to help students develop the skills necessary to negotiate and develop informed opinions about GMOs and other areas of scientific controversy such as climate change. Students are typically required to understand the scientific content while also being mindful of its moral and ethical ramifications, and then develop their own defendable position on the issue of GMOs. This combination of improved communication skills, understanding of core scientific concepts, in this case plant genetics, and critical analysis through evidence-based reasoning will better position students to make informed decisions when confronted with other controversial science issues in the future.

Study Context

I became interested in performing a study of how instructors and students deal with controversial issues for a number of reasons, both personal and professional. My PhD project involved the generation of genetically modified potato plants that were resistant to virus infection. Although the project was a success, my research was never translated into crop cultivars that benefited anyone. This was due to a combination of tight regulations and poor public opinion regarding GMOs. As far as I know, the seeds of my perfectly good plants remain in a freezer at the University of Toronto, waiting for a chance that they can be utilized in the future. It was and still is very frustrating to see all of my work condemned.

Years later, I have been involved as a guest speaker in the course "The GMO Debate" offered at Cornell University. The course is organized as outlined above, where students can be involved in a discourse of the social and scientific issues regarding GMOs and ultimately write an evidence-based report of their final viewpoint on the controversy. A drawback to the course is that most of the instructors have been proponents of genetically modified foods. The paucity of information on arguments against GMOs led to the launch of an anti-GMO course known as "GMO WTF" run by a small group of students. This informal course ran parallel to "The GMO Debate" and was led by a number of anti-GMO activists from the surrounding Ithaca area. I found it fascinating that a single course offered at Cornell could drive such an anti-GMO movement that students were impelled to begin their own course with their own carefully selected GMO opponent speakers.

This past year, several instructors were involved in developing a Massive Open Online Course (MOOC) entitled "The science and politics of the GMO". The MOOC provided videos, discussion boards, online feedback, and quizzes, and was offered for free or for a minimal certification charge to anyone with access to the Internet. I have taken several MOOCs myself in the past and find them a useful and convenient way to familiarize myself with a variety of topics. The MOOCs offer a different kind of teaching and learning environment and experience. I became interested in how this online community differs from that found in a face-to-face classroom and the way that this may impact how students develop informed decision-making processes surrounding the GMO controversy.

Problem Statement

I am interested in the pedagogy involved in teaching controversial topics in the context of a MOOC, such as "The Science and Politics of the GMO". Scientific rationale alone often lacks the ethical, social, and economic connotations which play a significant part of the controversy itself. In addition, instructors have their own biases. I am interested in how conversation develops between instructors and students in the MOOC setting. What changes and what stays the same throughout the duration of the course?

Research Questions

I wanted to look more critically at the stance of the instructor in the teaching and learning of controversial issues: how to be aware of and critical of

'instructor bias' when teaching a controversial topic, from both the instructor and student perspectives. How can instructors use pedagogical tools to successfully prevent their own bias surrounding a controversial topic from pervading the classroom without also losing their personal credibility? This question is intended to respond to student criticisms of the GMO Debate class, that all the instructors are pro-GMO. Hence, in the anti-GMO WTF class last fall, instructors displayed extreme mistrust of course content covered in the GMO Debate class. My methodology involves interviewing the MOOC instructors with a series of carefully crafted questions to identify how they intend to provide a balanced view toward a controversial topic that they are teaching.

I was interested in following the evolution of individual opinion toward a controversial topic: How and when does an individual develop or change an opinion, and what are the driving forces involved in this change within the context of a MOOC discussion board?. How do different discussion board formats affect student argumentary structures? My methodology for these questions would be an analysis of student contributions made on the MOOC discussion boards.

Literature Review

Emerging scientific areas such as biotechnology cannot follow the same paths of instruction as more traditional areas of science. In addition to scientific knowledge, courses in these areas require an examination of ethics, politics and social issues. Interest in biotechnology is important to capture and reinforce because society as a whole will need to better understand the potential benefits and risks that it offers. The outcomes of public decision making surrounding biotechnology will influence future governmental policies. Thus it is important to examine how biotechnology is taught and evaluated, and what kinds of argumentative structures are most helpful to assist in this decision making process (Hilton et al., 2011).

My literature search focused on the following three themes: the role of instructor bias in the teaching of controversial issues, how students evaluate conflicting data, and the impact of the online learning environment on class discussions.

The Role of Instructor Bias in Teaching Controversial Issues

I decided to begin my literature search by looking for examples of possible instructor bias and how that influenced the classroom experience and course outcomes. The literature on this topic ranged from the statistical results of questionnaires offered to large

numbers of instructors to the transcripts of individual interviews. Topics varied from genetically modified foods to climate change. I found that experiences regarding instructor bias observed by other researchers in the published literature closely resembled my own. These themes aligned with criticisms heard regarding 'The GMO Debate' course offered at Cornell, including the development of a GMO counter-culture as an unexpected outcome.

Unlike traditional disciplines in science, biotechnology is strikingly multidisciplinary in nature, and thus would benefit from a different format of teaching. How do instructors approach the topic of biotechnology in the classroom? Fonseca et al. (2012) surveyed 150 teachers from 20 secondary schools on their belief systems regarding biotechnology and biotechnology education. They found that teachers largely relied on their core belief systems rather than on science knowledge when teaching, and that they had to improve their information management skills in order to feel they could approach the subject adequately. Teachers seemed to have a limited awareness of resources available for education about biotechnology, yet were positively receptive of any alternative sources that were offered, such as animations and other material. The study concluded that efforts must be made to ensure that instructors are aware of resources available online and elsewhere that could make biotechnology as a topic for discussion more amenable to the classroom.

One way to examine controversies in science is through carefully constructed discussion formats. "Learning to argue as a biotechnologist: Disprivileging opposition to genetically modified food" (Solli, Bach, and Kerman, 2014) describes how groups of students learn to argue between themselves and develop their own viewpoints on issues concerning GMOs. The study was performed to gain insight on how students and professors justify and evaluate knowledge claims about GMOs within the context of a biotechnology course. More specifically, the authors wanted to know how the participants characterize and evaluate the arguments of those who oppose GMOs. The series of decisions that led these students to identify with and become part of the 'pro-GMO community' was also of interest.

The study was performed at a Swedish technological university in 2009. The students attended lectures and met in supervised, structured discussion groups. The discussions were video recorded so that arguments opposing GMOs could be fully evaluated. At the end of the course students prepared a written report of their summary viewpoints on GMOs and what had led them to make their decisions.

Results suggest that opinions made about those opposing GMOs centered on misunderstandings of scientific knowledge. For example, students brought up the point that they had heard GMO opponents insisting that "they do not want to eat food that has any genes in it". Although they find humor in the ridiculousness of this statement, these GMO proponents do not describe any of the writings of any nutrition researchers, environmentalists or economists who also oppose GMOs, all professionals in their perspective disciplines. In this case, the argument source that is least rational is brought to the forefront, while more reasonable arguments made by experts in the field are at best marginalized or at worst ignored completely. The decisions made regarding anti-GMO views were disconnected, therefore, with legitimate concerns held by substantial portions of the public.

Furthermore, whether students really expressed their own thoughts or only what they thought the teacher expected of them remained in question. The university where the study took place has a heavy biotechnological slant and graduating students would most likely pursue a science career. The authors brought up the point that agreeing with anti-GMO concerns could potentially slow or stop the development of biotechnology, thus impacting future job security for many of the students who were taking the course. This implies that there is more at stake for stakeholders than merely competing scientific theories. The authors conclude that the lack of self reflection identified here is problematic for universities who promote genetically-modified crops.

The interpretations described in a later paper, published by a different author, using the same research study and data of Solli et al .were quite compelling. In "Biotechnology education as social and cultural production/reproduction of the biotechnology community", Maria Andre'e (2014) provides an alternate interpretation of the data generated by Solli et al. in their 2014 study. She examines how both students and instructors deal with the dialectic of individual and collective transformation on the controversial topic of GMOs. Andre'e argues that the more specialized functions a society requires, the more specialized the educational system will need to become to meet the need for societal reproduction. However, she sees potential for the development of a counter-culture of non-critical reasoning, reminiscent of the GMO WTF anti-GMO course established at Cornell. This author reinforces the point that the instructor frames GMO opponents as being both ignorant and laughable while ignoring any possibly well-grounded opposition from scholars of other disciplines. The instructor uses phrases that

indicate suspicion about the ulterior motives of the opposition, indicating that GMO opponents are 'very skillful' at 'propaganda', while scientific literature is more trustworthy because it is carefully reviewed by the scientific community. Andre'e concludes that alternative viewpoints regarding GMOs must be legitimized and taken seriously in class to prevent the establishment of a counter-culture that may eventually obstruct agricultural biotechnology from moving forward.

How Students Evaluate Conflicting Data

When it comes to the controversy of GMOs, personal and social values often overcome knowledge-based science. It has been discussed that the deep unease that people feel toward biotechnology is caused by the ways that it challenges people's sense of existential order, and how society handles the responsibility of utilizing biotechnology in a safe way (Deane-Drummond et al., 2014). Little is known about the different emotions that are triggered when learning about the GMO debate. Media reports on GMOs often report emotions of worry, anger, and fear among the public. To examine the possibility that students also harbor these emotions Sorgo et al. (2011) surveyed over 500 students from both secondary schools and universities. The questionnaires were designed to allow students to indicate their emotional response to a series of statements about GMOs. In general, the study showed that students expressed more positive emotions than negative ones.

Albe (2008) examined the argumentation patterns of a number of students who were provided with short texts regarding a controversial issue and then asked to discuss them. Students were also given an opportunity to personally react to the issues raised. The author used role play to show that there is a complex series of factors which influence students' argumentation processes on a socio-scientific controversial issue. Importantly, scientific knowledge played only a minor role in the argumentation process. The role of the scientific community in providing input to a socio-scientific issue in the public domain requires further definition.

Sadler et al. (2004) investigated students' interpretation and evaluation of conflicting evidence regarding a socio-scientific issue. In their study, eighty-four students were each presented with a fictitious 'Science Brief' reporting opposing views of the issue of global warming written by the same author. One statement, entitled 'Global Warming: An Impending Environmental Crisis', reported that global warming is caused primarily by humans and a very real threat to the environment. The other statement, entitled

'Global Warming Myth: Evidence Against Environmental Crisis', presented evidence suggesting that the global warming is a natural event and poses no real threat to the environment. All of the students were asked to fill out a questionnaire based on their analysis of the statements. Interviews were also conducted on a subsample of thirty students.

When students were asked to report which article they found more convincing and why, students cited personal relevance as a primary issue which affected their decision, followed by the quality of information that each article presented. Many students felt that they were most convinced by the position that complemented their own personal beliefs. This factor seemed to be even more important than scientific merit. Similar studies have also found this to be true. For example, a study performed by Zeidler et al. (2002) found that students often compartmentalize scientific knowledge versus personal opinion.

This study reveals that students enter the classroom with a diversity of viewpoints about the issue of global warming. While many students displayed a general understanding of the science, just under one-half of the students remained naive of and could not describe the data they were presented with in the reports they were given. In fact, many of the students valued opinions equally with data. The authors conclude that socio-scientific issues such as climate change and GMOs are often reasoned out primarily based upon a prior belief system which aligns with the information they are provided with. In addition to this, students' decisions on controversial topics are heavily influenced by the personal relevance of the topic, as it may discuss consequences which they could personally relate to. The task of the instructor, therefore, is to challenge students to consider alternative viewpoints and dissect the rationale of their prior opinions. The authors point out that students readily dichotomize personal beliefs and scientific knowledge, with 40% of respondents claiming that the article with the most scientific merit was less convincing than the other article, which they found more personally relevant.

In another study Seethaler and Linn (2014) examine best practices in supporting students' learning, such as scaffolding, as they compare and combine 'pieces' of knowledge about an issue to form an integrated understanding about complex multidisciplinary socio-scientific issues such as the controversy surrounding GMOs. They analyze outcomes of a course where 173 students were asked to sift through and synthesize what they have learned to take a position on what agricultural method they think should be used in their geographical region, while providing appropriate evidence for their decision. This requires students to make use of evidence to defend their position. The authors used the Scaffolded Knowledge Integration (SKI) Framework to design the GMF curriculum (Linn & Hsi, 2000). Students were also given pre-course and post-course surveys to determine how their attitudes toward GMOs had changed. The authors found that students' responses became more normative and less vague throughout the duration of the course, suggesting that they better understood GM crops and agriculture in general. It is interesting to note that nearly one-third of students described in their papers information that they learned from their peers' presentations, and that most students wrote about the advantage of using genetic engineering to improve food taste and quality. The authors concluded that students need time and supportive guidance to engage in the act of reflective thinking in order to construct arguments with evidence to support their positions. Students also benefit by exploring the evidence counter to their positions, which shows that they understood both risks and benefits of agricultural biotechnology. Providing the students with tools to apply what they know made them more autologous as learners.

Wu (2013) generated 'flow maps' based on transcripts from students interviewed regarding GMOs. Each student's flow map illustrated the individual's knowledge structure and could be employed to analyze the student's usage of different information processing modes. Students who more frequently used higher-order processing structures were found to have the best strategies for organizing information and had a greater tendency to achieve a higher-level informal reasoning quality. These students were able to more easily present arguments from multiple perspectives and could achieve a higher informal reasoning quality.

Online Learning, MOOCs, and Their Impact on Class Discussions

Online education has experienced a period of exponential growth over the last decade. Over six million students are currently enrolled in at least one online course and the growth rate of online courses is greater than that of the overall student population in higher education today (Allen & Seaman, 2011). In 2016, one in four students were enrolled in online courses (Online Learning Consortium, 2016).

Online courses provide convenience, flexibility and easy access to education for many students. However, the drop rate for online courses is generally much higher than it is for face-to-face courses. One reason for this difference could be that feelings of isolation and disconnectedness are created by the phys-

ical and psychological distance between instructors and students in online courses (Hart, 2012). However, a recent study by Bowers and Kumar (2015) suggests that the perception of online instructor presence that is encountered by students does not radically differ from that of students in a more traditional, face-to-face classroom setting. The authors examined the performance and responses of thirty-four undergraduate students enrolled in an online course and twenty-nine students enrolled in a traditional classroom. The course was identical, taught by the same instructor and covering the same content, with the same syllabus, assignments, rubrics, and discussions. The authors employed a mixed-methodology approach, using both survey and open-ended question formats. Students perceived a stronger instructor presence in the online course, which offered video lectures, carefully designed interactions that were carried out in trusting environments through discussion boards, graduate student contact, and instructor feedback. Students were provided with the opportunity to experience and improve their abilities to reflect upon the knowledge they that they built on in the course through the use of discussion questions based on higher order thinking skills. The learning community was further enhanced by the quality of instructor presence during the course. The instructor was able to effectively communicate well- defined goals and learning outcomes and engage with students personally and in a timely manner. The quality of instructor interaction, combined with a sense of belonging and perceived learning, motivated students to persist and be successful with the online version of the course. In fact, the average for students completing the course was 93%, compared with 87% for students who took the course in a face-to-face format (Bali, 2014; Hart, 2012). This implies that online discussions can be just as effective, if not superior, formats for students to explore their critical thinking and decision-making processes through interaction with others.

Online discussions about controversial topics such as GMOs are gaining in popularity and offer a unique experience for the participant. For example, Online Partnerships in Education (SCOPE) set up an online genetically modified food (GMF) discussion board that offered various materials to help visitors understand the GM controversy (Triunfol & Hines, 2004). The discussion list presented multiple positions and included 20 commentaries written by scientists. The Virtual Panel Discussion highlighted the divergence in opinions among scientists of different disciplines. People with various levels of understanding, backgrounds, and points of view and who wished to debate the controversies surrounding genetically modified foods could join the discussion list. The site became very heavily used; for example, during the first quarter of 2002 traffic to the SCOPE-GMF site averaged 1,070 hits per day.

In another example, Walker and Zeidler demonstrated that computer-based activities can provide a dynamic learning environment by which students can analyze current scientific controversies. The Web-based Science Environment (WISE) curriculum enables students to pull apart their own pre-instructional beliefs and compare them with different features of argumentation utilized by other students who take the course.

Online courses allow students to spend ample time reading about and evaluating a socio-scientific issue that has been well prepared by an instructor, instead of surfing the web for information that can sometimes be misleading. Argumentation skills can be developed through a variety of online activities that promote interaction between both instructors and students. Students' reactions to these activities can be traced and monitored through discussion boards to determine the role of instructor bias in the argumentative process. My study explores how these elements work together to change or reinforce opinions of the GMO controversy.

In conclusion, these studies suggest that instructors can avoid the appearance of being outwardly biased toward a controversial topic by taking great care in presenting only arguments that are well-founded and justifiable. Carefully constructed discussion formats can ensure that only valid forms of arguments are presented. Pre-existing beliefs regarding conflicting scientific data can significantly influence students' decision –making processes. In addition, online learning activities can positively influence how students build argumentative structures within a 'safe' learning environment.

Provided with this background, I wanted to find out how the instructors of the Cornell MOOC had intended to navigate students through the considerably contentious topic of genetically modified foods. I was also interested to determine how the students evaluated what they were learning with respect to both self-reflection and their interactive discourse with other students. I found that instructors were largely successful in achieving their goals. The mode of online discussion format that was used made a big difference with respect to how well the students could engage with the course material. This literature review provided additional insight and was instrumental for the formation of my experimental design.
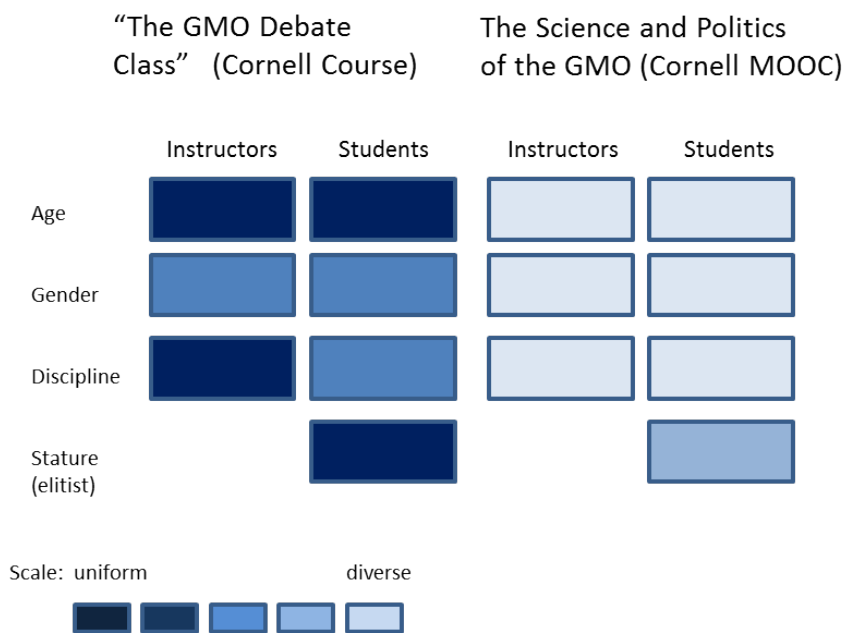
Figure 1: Classroom dynamics.

Methods

All of the work described here was first successfully filed for IRB exemption by Cornell University Office of Research Integrity and Assurance.

In the first part of this project, the four instructors of the MOOC "The Science and Politics of the GMO" were interviewed. Consent was provided by each instructor through a consent form that is listed as a supplement to this paper. Each instructor was prompted with a series of open ended and/or follow up questions. The interviews lasted for approximately one hour per instructor. All of the instructors' answers were recorded and transcripts were analyzed.

The questions were:

1. In your opinion, What do you feel is the role of the instructor with respect to discussing controversial issues?

2. In your opinion what do you think the role of the institution (meaning Cornell) should be with respect to the teaching of controversial issues? Do you feel that you as an instructor have adequate support from your institution with respect to exploring new teaching strategies to help you address controversial topics in the classroom?

3. How do you avoid perceptions of 'instructor bias' with respect to teaching controversial issues ? In what ways do you think your own viewpoint might affect classroom dynamics? Or individual student beliefs/feelings?

4. What pedagogical methods can be employed in the classroom to approach controversies?

5. What are the pros and cons of teaching a controversial topic online as opposed to in person? Do you think you conduct yourself differently? Do the students respond differently in an in-person versus an online setting? If so, in what ways?

6. Can you identify what driving forces might be responsible for any changes in opinion (or lack of change in opinion) toward a controversial topic that you have personally seen in your students over the duration of the course?

The second part of the project involved the analysis of two discussion board formats within different offerings of the MOOC, one which ran in the Fall of 2016 and one which ran in the Spring of 2017. Discussion board participants have already provided consent to have their comments used when registering for this MOOC. There are approximately 700 MOOC participants, all between the ages of 18- 65. MOOC participants could originate from all over the world. The duration of the course was 5 weeks. The Fall version of the MOOC used an EdX discussion board format, whereas the Spring version used a Facebook format. Participants were followed along the discussion board; their interactions with both instructors and other students, along with their own personal reflections on the material covered, were monitored each week. The development of selected participants' argumentary structure over the duration of the course was analyzed. The two discussion board formats were compared with respect to a series of criteria including evidence-based discussion and the types of supporting files uploaded, tone and depth of engagement with other students, level of sophistication with respect to discourse and the evolution of opinion regarding a controversial topic. Data collected from both stu-
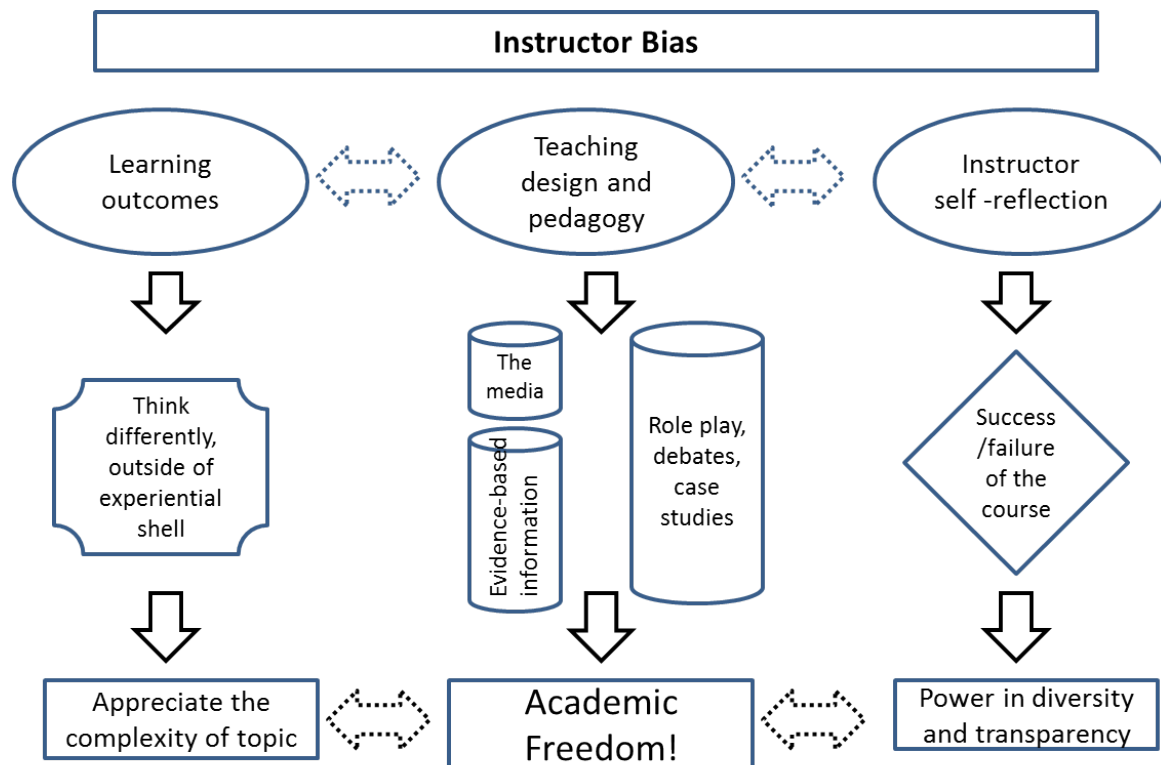
Figure 2: Coding tree for instructor interviews.

dents and instructors was codified using the method of Bryman, (2012) and coding trees based on the major themes presented were developed.

Results

The original GMO Debate course, held in a classroom setting, consisted of five primary instructors and a number of invited guest speakers each year. The majority of instructors were older professors and most were scientists by discipline. Gender was equally divided among the instructors (Figure 1). The student body consisted of Cornell University undergraduates, many of them from agricultural backgrounds and completing science majors that aligned with this discipline. Others were majoring in social science. For the MOOC, the instructor and student body changed dramatically. The MOOC instructors represent a wide spectrum of ages, genders, and disciplines, including science, communication, social science, behavior economics, and political science. The student body enrolled in the MOOC was also more diverse (Figure 1).

Instructor Interviews

Instructors were each asked a series of six questions and their answers were coded using the method of Bryman, 2012. The coding tree revealed that instructors all felt that maintaining academic freedom was a core component of their pedagogical strategies. Instructors described the importance of evidence-based information and its role in interpreting news, misleading or not, that is presented by the media (Figure 2). The instructors were unanimous in their responses that controversial topics such as GMOs should be addressed in the classroom using role play, case studies, and classroom debates aimed at developing and refining students' argumentative structure. The diversity of academic backgrounds provided a sense of security among the instructors that their motivations were transparent and open to challenge by both their peers as well as the student body. For the course to be considered a success, the instructors felt that it was more important that students completed the course with a new appreciation of the topic's complexity, rather than influencing student opinions about GMOs. The results of these interviews are summarized below and were formatted into a coding tree structure (Figure 2). Themes and associated quotes from the interviews are depicted in Table 1.

Summary of Responses to Interviews

What is the role of the instructor when teaching controversial issues?

An attractive feature of the MOOC format is the fact that the instructors can use the collective diversity of their expertise to interrogate, engage, and debate with each other. While most of the instructors of this MOOC recognized that it is next to impossible to remain neutral on a controversial topic, they

Table 1: Themes from instructor interviews.

Questions for instructors	Notable quotes
Q1 What is the role of the instructor with respect to teaching controversial topics?	"when you have multiple instructors, you do not have control over where the narrative goes"
Q2 What is the role of the institution with respect to teaching controversial topics?	"My peers will see this; every word counts"
Q3 How do you avoid perceptions of instructor bias?	"Don't bend yourself in a way that denies your academic responsibilities"
Q4 What pedagogical methods were employed?	"Introduce the topic in a way that diffuses all politically incorrect bombs"
Q5 What are the pros and cons of teaching online versus in person?	"Online discussion board gives people the time to react and the opportunity to wordsmith"
Q6 What changes in opinion did you see over duration of the course?	"People open their minds, become more nuanced in their thinking"

all were cognizant of the others' strengths and how each shaped the narrative in different ways. For example, several of the instructors focused on empirical evidence and the critical analysis of information for its validity. Others welcomed controversy and instead played a role of navigating the students through the vast array of information surrounding the science that constitutes the GMO Debate, such as risk perception, trade, ethics, and the role of the media. In fact, several instructors pointed out that scientific structure doesn't seem to make some people feel welcome, and went out of their way to provide alternative approaches that expose students to productive ways to evaluate information sources and the way they are transmitted.

What is the role of the institution?

All of the instructors felt strongly about maintaining academic freedom at the university, so that students feel that their opinions are valued. Instructors recognized that while the university should prevent instructors from becoming advocates for a favored position, the freedom to express one's opinion should not be used as a vehicle to descend into, as one instructor put it, "pure quackery." One instructor did mention that from a more extremist point of view, the course could be seen as privileging an elite populace. This opinion was shared by one other instructor,

who questioned student motivations for enrolling in this particular MOOC.

How to avoid instructor bias.

The instructors held firmly to the belief that transparency about their own bias was critical for the success of the MOOC. They all recognized that they are in a position of power and made efforts to equalize that dynamic by offering provocative questions that were challenging not only for the students but for the instructors themselves. The instructors could directly challenge one another, and in doing so created a space where people felt comfortable challenging others including instructors. All sides of the GMO debate were discussed week by week by a diverse teaching team.

Pedagogical methods to approach controversies.

Controversial topics can be introduced by getting the students to talk openly about their prior beliefs in the class in a manner that is inoffensive. The details of the controversy can be slowly introduced as students become comfortable with the course format. An example of this can be found in the labeling exercise of week three. The students were asked to go out and find examples of food products that they regularly buy at the grocery store and analyze whether the labels represent (or misrepresent) appropriate claims

for the product. This exercise got the students to critically analyze what they were eating, without necessarily bringing GMOs into the equation, and gave them a more complex understanding of the labeling of food products.

The concept of incorporating real life scenarios that do not have straightforward solutions is also a part of the MOOC. This is accomplished via interviews with people who are impacted by GMO regulations, such as farmers from Uganda, Hawaii or Bangladesh. A mixture of film clips representing real farmers were provided as well as activities that the students can perform in their own grocery store. Students were able to use these tools to gain a better understanding of how far-reaching consumer choices can be, and correspondingly, their global impact on producers.

The instructors all described the benefits of role play and case studies as pedagogical methods that can be used to better understand controversies. By adopting positions that they may not particularly favor and challenging each other, the students learn to look more intensely at all sides of an argument. In the MOOC students are given a hypothetical scenario (should GM mosquitoes be used to prevent a fictitious disease from spreading in a fictitious town) and are randomly given the role of one of the stakeholders (government, scientist, public). The students were then asked to debate from the perspective of their assigned stakeholder roles. This in turn helps them to consider the complexity of an issue such as the GMO Debate.

The students were also provided with opportunities to contribute to both live sessions as well as through written comments. This provided an easier format for those who are more comfortable with speaking rather than writing. Student assignments were peer-evaluated so that feedback from multiple sources and points of view could be provided to such a large class size.

Teaching online versus in person.

In general, the instructors welcomed the online nature of the MOOC as a way to reach a bigger audience; however, several did mention that the MOOC students represented more of an upper class community compared to the in person course. Instructors noted that a lack of confidence that they were able to reach as broad an audience as had been wished for. Two of the younger instructors mentioned that when constructing their video lectures they were extremely conscientious that future employers and anti-GMO activists could see their presentations. This made them cautious about what words and gestures they included. These instructors missed the luxury

of teaching a live class where they can convey the subtleties found in person-to-person relationships and implement quick checkpoints to ensure that the students are taking in and processing the information accurately. Being unable to see the students and watch their reaction through words and gestures was an impediment for some, and several commented on their familiarity with the Cornell student body compared to the large diversity of age, gender, ethnicity, and socioeconomic status found in the MOOC audience. The fact that the online course is not for university credit, as opposed to a live Cornell course that would count toward a student's degree, also altered the dynamics of the student responses.

Other instructors found the online setting greatly advantageous. From their perspective, their initial time and effort invested into the MOOC led to a high return of investment with respect to the number of students who could be reached and who would benefit from the course. They were thrilled that the work they presented could reach a classroom of 10,000 or more, rather than the few dozen they would regularly teach on campus. The videos, animations, and other formats prepared for the MOOC could be used by other instructors who would like to bring this topic into the classroom.

Interestingly, both of the male professors commented on how students were far less likely to be bullied by a professor online as compared to in person. The distance provided by the online format allowed people the freedom to disagree, take their time to react, and think things through before stating their true opinions. Indeed, one instructor commented that he enjoyed the opportunity to wordsmith and neutralize off-putting comments through the online format. Students in turn would offer ideas online that they may never say face to face. The instructors found the online discussion to be respectful. The students used the EdX Discussion Board in the Fall MOOC, which several instructors complained had a clunky and frustrating structure. The online discussion was moved to Facebook for the Spring session of the MOOC; a general problem with Facebook was a lack of organization. For example, discussion threads did not seem to appear in chronological order, rather posts were organized in order of popularity, making it more difficult to follow conversations and access data.

Any changes in opinion?

None of the investigators felt that students seriously changed their opinions about GMOs during the course. However, that was not an outcome they were looking for. What the instructors felt was a success was that they managed to get the students to ques-

Which of the following most accurately reflects your feelings about GMOs?

1. "GMOs are one way to help address food insecurity in a safe and sustainable way"

2. "GMOs have demonstrated significant promise, but require more testing by independent scientists"

3. "I have not yet formed an opinion"

4. "I think tinkering with genetic codes is too risky and unnatural, and requires more scrutiny than scientists are capable of"

5. "Not enough time is spent considering the non-science, social, and economic impacts of this technology"
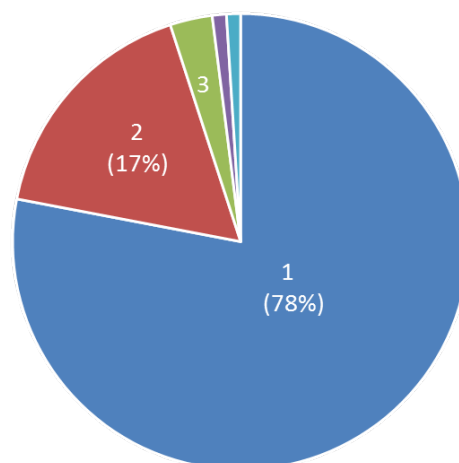
EdX Discussion Board (Fall 2016)
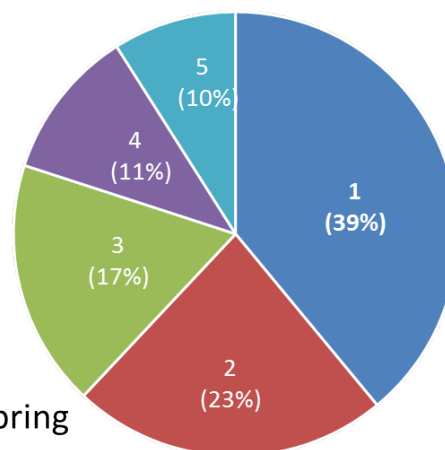
Facebook (Spring 2017)

Figure 3: Students' feelings about GMOs.

tion their own values and think about the topic of GMOs differently, not simply as an abstract scientific concept, but also in a real life setting such as a grocery store. Students who completed the MOOC were aware of the opposing arguments and were provided with enough comprehension of the topic that they can move the dialogue forward in a meaningful way. Students were provided with the opportunity to get out of their experiential shells and speak to actual people whose lives are impacted by the GMO Debate, such as the farmers themselves. By opening their minds to all aspects of the controversy, the students left the course more nuanced in their thinking and appreciative of the complexity of the topic.

Analysis of Student Activity on Discussion Boards

Depending upon when they enrolled for the course, students were expected to contribute their thoughts and comments to either the EdX Discussion Board or Facebook group. Since its inception at MIT and Harvard in 2012, EdX has been well established as a MOOC provider of university-level courses in a wide range of disciplines. Today over 70 schools, nonprofit organizations, and corporations provide over 1,270 EdX online courses to over 10 million students worldwide. Facebook, on the other hand, was launched in 2004 by Mark Zuckerberg as an online social media service out of California. Originally designed for university students, Facebook is wildly popular among people of all ages for its easy status updating, message sharing, friending searches, and photo uploading functions.

Students who enrolled in the course were expected to fill out a survey regarding their attitudes about GMOs prior to beginning the first module. The demographics of the student body differed between each of the two times that the course was offered. Students who took the course in the fall were far more comfortable with the concept of GMOs than students who took the course in the Spring (Figure 3). This change in student body may be the result of how the course was marketed the second time it was offered.

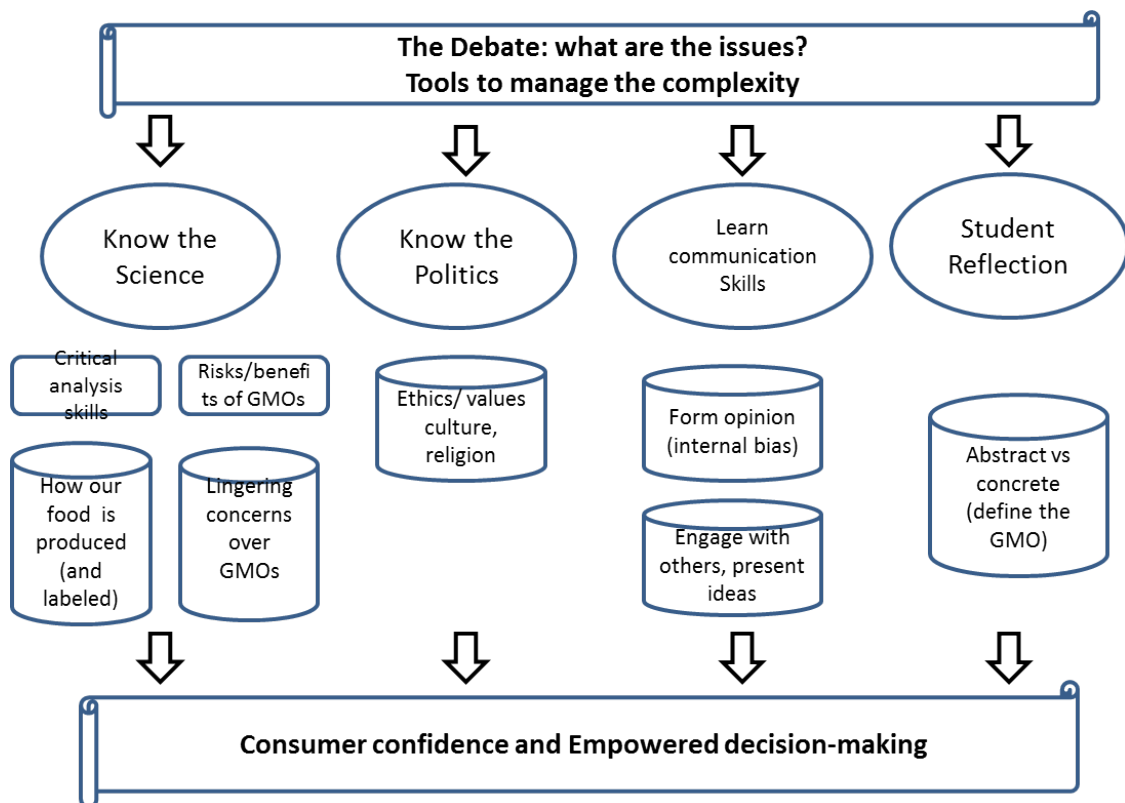During the first offering of the MOOC in Fall 2016

Figure 4: Coding tree for student discussion board, Fall 2016.

the student body prepared a series of responses to carefully designed questions (eg. define what is a GMO?) and crafted their answers in the form of one to several paragraphs. By reading approximately 30 of these entries, trends in the student thought processes began to surface. Most students felt that the course had provided them with a set of tools by which they could manage the complexity of the GMO controversy. These tools could be further pared down to understanding the science, the politics, and the skills needed to communicate effectively. The students were given opportunities through the discussion board to engage with their peers and exchange ideas. Finally, the students were encouraged to reflect on their prior bias and preconceived notions of what a GMO is in all of its abstract and concrete forms. A coding tree structure based on discussion board responses is provided in Figure 4. Themes and associated quotes from students who contributed to the discussion board are listed in Table 2.

EdX Versus Facebook as a Format for Student Discussion

The EdX Discussion board format was replaced with Facebook for the Spring version of this MOOC. The EdX discussion board was utilized by students primarily as a formal academic space, and this was reflected in the tone set by the students. In general, the discussions were dry in quality, carefully crafted

and well thought out, with only one or two responses (if any) to a particular post. The narrative was scientific, even-tempered and well constructed. The discussions on Facebook were of a different nature. These were high in emotionality, more contentious, and often highly personal. For example, one student spoke quite ardently about his battle with celiac disease on Facebook, how this impacted his feelings about food labeling, and how he wished that GMOs had been produced which could address his condition. Another student was more than happy to upload a photo of herself at the Capital Buildings in Washington, DC, protesting the lack of GMO labeling on food. It was clear that the student body present on Facebook was more willing to challenge the course material, was less focused on the science of GMOs, and placed more emphasis on the politics instead. Unlike the EdX discussion board, there were often as many as 30 interactive responses to a Facebook conversation about a particular MOOC post, with as many as 20 different respondents. The comments on Facebook were more impulsive, more informal, and less-evidence-based. It was apparent that the students were having fun as well. The post about food labeling, for example, led to students loading a series of photos of mislabeled food products from their local grocery stores, in an effort to outdo each other (Figure 5). The free-for-all was confounded by the fact that the Facebook threads lacked chrono-

Table 2: Selected discussion board comments.

Comments from the Discussion Board
"I learned about the arguments of those opposed to GMOs and the reasons behind them. It opened my eyes to the other side of the argument"
"the material gave me a great background and access to a plethora of resources. The course offered me a whole new set of tools to use when engaging in conversations about science"
"I learned that I cannot expect all people to make decisions simply based on science. I need to understand that some people don't think like me"
"throughout the course, I looked at the definition of the GMO itself, and even that is something that is not clear."
"the course made me really dig deep into my firmly held biases regarding food production"

logical order, making it difficult for anyone to follow the story line and join in for the moment. While this format has its advantages, it was difficult for the MOOC instructors and webmasters to provide appropriate feedback and encouragement along the way. For example, Figure 6 illustrates a conversation initiated by two MOOC instructors regarding food labeling. A new person enters the thread and changes the conversation to one on the ethics of scientists accepting funding from large multinational companies. This was followed with at least 20 more entries by 11 different students conversing about the ethics of scientists. Such a change of course and emotional outpouring was never observed on the EdX Discussion board.

The results of this study indicate that the best Facebook exchanges center around broader topics (e.g. food labeling, biopiracy and large corporations), rather than simply on science. EdX discussions, on the other hand, were more science focused and evidence-based. Students on Facebook are more colorful, visceral and lively. The discussions were also more emotional and personal than those found on EdX.

Discussion

The Role of Instructor Bias in Teaching Controversial Issues

It was interesting to familiarize myself with the literature concerning instructor bias prior to interviewing the instructors. I found myself finding insight and appreciating their answers more resoundingly, as I linked their responses to what I had been reading and thinking about. With respect to instructor bias, I feel that this group of faculty did their utmost to present a fair and uniform depiction of the GMO controversy without enabling students to go down dark and twisted paths of junk science and media misrepresentation. I think that incorporating real problems concerning GMOs that are legitimate and valid, rather than ignoring them as described in several papers outlined in the above literature review, provided authenticity for the students. The points made by several instructors that they are conscientious that their words and gestures could be evaluated by future employers kept a tighter structure to the debate than one may find in a live classroom setting, for better or for worse. The fact that both instructors and students alike could take pause, plan, and wordsmith their responses online also kept inflammatory exchanges to a minimum.

Facebook Example 1.

MOOC participants are asked to find a food product in their local grocery store and upload a photo of the labeling. They all went to town on this assignment, trying to outdo each other for the best case of mislabeling! The EdX students politely discussed labeling, but did not add photos. I don't think it was as much fun for them.

Student X

Almonds on the left have the Non-GMO label. This is an example of a product that is not available as a GMO. The label is a marketing tool and implies that GMO products are 'bad' in some way.

Student Y

I was surprised to see all this information in this chips bag. It may be counterproductive to have so many pieces of information that can mislead you to believe you are getting the most nutritional healthiest product. It irritates me when the product has no wheat, barley or rye that it comes with the "gluten-free" logo.

Figure 5: Facebook Example 1.

Facebook Example 2

Labeling is a good topic to debate on because you can be pro or anti-GMO and still be on either side of the argument.

The instructors here are helping the students ask themselves deep philosophical questions about the ethics of food labeling.

However, one student responds to the question of ethics by moving onto a tangent about the ethics of scientists, funding for research....and selling out to industry!

- **Instructor A:** What would a meaningful label look like? What types of information do you think that consumers would like to know and how can that be best presented on a label? Any advice for the USDA as they continue to deliberate on mandatory labels? (3/11/2017)

- **Student 1:** How would labelling of GMO work in reality, if we can't even agree what a GMO is - as the first unit of this course clearly showed? Could we find an international consensus?

- **Instructor B:** Ola, a great point you make here. But also, in rural societies, and small towns all over Asia, there are no labels on anything fresh: it's hard to imagine labeling in systems other than those controlled by formal distribution and marketing networks.

- **Student 2:** Money in science is such a huge issue! But it does seem to me that when a scientist makes the transition to lobbying and promoting AND being paid for it, s/he loses her integrity as a research scientist....

20 more exchanges (with 11 different facebook friends) later....(4/8/2017)

Figure 6: Facebook Example 2.

I appreciated the fact that the instructors each acknowledged their own bias and made efforts to remain transparent as they taught the course. Fonseca et al. (2012) identified that teachers maintain their core belief systems when teaching a controversial topic. The MOOC instructors were able to both acknowledge and circumvent this problem by challenging each other online and developing assignments for which there is no one right answer and that transcended the GMO controversy, such as food labeling or the role of patents in agriculture. Students were provided with case studies that did not create a dis-privileging environment as described by Solli Bach and Kerman (2009).

How Students Evaluate Conflicting Data

As mentioned above, student responses differed depending on whether the EdX university discussion board or the Facebook format was used. The emotionality of GMOs was addressed in the MOOC, and even their impact on students' religious views was discussed. While worry, fear, and anger came through in the Facebook posts of several students, much optimism was also presented, similar to the work of Sorgo et al., (2011). I was left to conclude that the Facebook responses focused less on the science and that many other factors such as politics, marketing, and even personal preference were considered. This observation might suggest that the cohort of students who used Facebook held science as less important in their decision-making process than did the edX cohort. However, distinction of traditional uses of the Facebook versus edX online environments, including differences in uploading and posting personal photos, websites, etc., could also explain the changes in argumentary structure between the students who used each format. The edX students made a greater effort to present an evidence-based argumentative structure, but unfortunately these turned out to be mostly summaries that others did not engage with. The Facebook respondents' argumentative structures were not as refined in general, however, a few did come through with poignant comments that made me thankful for having read them. Since Facebook is a popular social media site where people regularly upload topics ranging from their response to the recent US election to events concerning family and friends, it may be difficult for some to consider it an appropriate online space for more sophisticated discourse. Students may associate the Facebook site as less of an academic space than the edX discussion board. It is also important to note that far more students who enrolled in the Fall and used the edX site were pro-GMO to begin with. The nature of their arguments possibly reflected their comfort with the scientific aspect of the controversy. I felt less confident that the science behind GMOs had been fully and adequately digested by the Facebook student body, due to the less scientific nature of their postings.

Online Learning, MOOCs, and Their Impact on Class Discussions

Bowers and Kumar (2015) discussed the importance of student perception of instructor presence in online courses. I am aware that the instructors made efforts to stay on top of the entries posted in both discussion boards. The Facebook site was not maintained in chronological order, making it difficult for instructors to keep track of what was being said. The situation was exacerbated by the fact that the Spring MOOC had no teaching assistant. However, the comments offered by the end of the course indicated that while the Facebook discussions were somewhat fractured, the students were highly satisfied with their MOOC experience and felt that they had broadened their minds, which was what the instructors had sought to achieve.

What are the lessons to be learned for other instructors regarding teaching controversial issues online? Both formats can work, depending on what an instructor would like to accomplish. Facebook is a well-established format for rants and tirades! EdX would be considered to be best from an instructor's perspective for grading, but some of the color of the discussions will be lost.

The use of a diversity of instructors and media, as demonstrated in this course, made a substantial impact for keeping instructor bias to a minimum. The students' confirmation that they left the course with a heightened sense of the complexity of the GMO controversy indicates that the course, in general, was a success. Online discussion can help with this process, but care must be taken to ensure that the students are provided with a technological format that helps them to engage with each other while improving their ability to build and present argumentative structures that are both sophisticated and well thought out.

References

Albe, V. (2008). When Scientific Knowledge, Daily Life Experience, Epistemological and Social Considerations Intersect: Students' Argumentation in Group Discussions on a Socio-scientific Issue Res Sci Educ 38, 67–90

Andre'e, M. (2014). Biotechnology education as social and cultural production/reproduction of the biotechnology community Cult Stud of Sci Educ 9, 25–30.

Bowers, J. and Kumar, P. (2015). Students' Perceptions of Teaching and Social Presence: A Comparative Analysis of Face-to-Face and Online Learning Environments International Journal of Web-Based Learning and Teaching Technologies (IJWLTT) 10(1).

Deane-Drummond, C., Grove-White, R. and Szerszynski, B. (2014). Genetically Modified Theology: the Religious Dimensions of Public Concerns About Agricultural Biotechnology Studies in Christian Ethics, 14(2), 2001, pp. 23-41.

Fonseca, M.J., Costa P., Leonor Lencastre, L., Tavares, F. (2012). Disclosing biology teachers' beliefs about biotechnology and biotechnology education Teaching and Teacher Education 28 368e381.

Hilton, A., Nichols, K., Kanasa, H. (2011). Developing tomorrow's decision-makers: opportunities for biotechnology education research Aust. Educ. Res. 38, 449–465.

Online Learning Consortium. (2016). Report: One in four students registered in online courses. Retrieved from `https://onlinelearningconsortium.org/news_item/report-one-four-students-enrolled-online-courses/`

Seethaler, S. and Linn, M. (2004) Genetically modified food in perspective: an inquiry-based curriculum to help middle school students make sense of tradeoffs, International Journal of Science Education, 26:14, 1765-1785.

Sadler, T.D. , Chambers, F.W. and Zeidler, D.L. (2004) Student conceptualizations of the nature of science in response to a socioscientific issue, International Journal of Science Education, 26, 4, 387-409.

Solli, A., Beach, F. and Kerman, B.A. (2014). Learning to argue as a biotechnologist: Disprivileging opposition to genetically modified food Cult Stud of Sci Educ 9, 1–23.

Šorgo, A., Dolinšek,J.A., Tomažič, I., Franc Janžekovič, F. (2011). Emotions Expressed Toward Genetically Modified Organisms Among Secondary School Students and Pre-Service Teachers Journal of Baltic Science Education, Vol. 10, No. 1.

Triunfol, M.L. and Hines, P.J. (2004). Dynamics of list-server discussion on genetically modified foods. Public Understand. Sci. 13, 155–175.

Walker, K.A. and Zeidler, D.L. (2014). Promoting Discourse about Socioscientific Issues through Scaffolded Inquiry International Journal of Science Education 29 (11), 1387-1410.

Wu, Y.T. (2013). University Students' Knowledge Structures and Informal Reasoning on the Use of Genetically Modified Foods: Multidimensional Analyses Res Sci Educ. 43, 1873–1890.

Habitus in Big Classrooms

Jing-Mao Ho

Department of Sociology, Cornell University
jh2268@cornell.edu

Introduction

Sitting in a big lecture hall and listening to an instructor speak is a common experience of undergraduate students. It is usually assumed that classrooms are places where student learning occurs. However, is learning always the only thing that students engage with in big classrooms? Investigating classroom dynamics that are directly or indirectly related to this question is an important area of research in education. For example, studies of classroom composition have mainly explored how the student composition of a class, in terms of ability, race, gender, or other factors, impact student learning. More recently, with the advancement and proliferation of computer technology, educators and researchers have started to debate the influence of the use of computers or other electronic devices in classrooms on teaching and student learning.

This paper starts with a simple but fundamental question: What do students do in big classrooms? As mentioned above, most previous research on classroom activities primarily focuses on the relation between students' behavior in classrooms and learning outcomes. The pedagogical concern underlying this literature is the question of how to improve teaching and learning in classrooms. In order to improve teaching and learning, teachers and researchers tend to collect data on a variety of factors that are relevant to student learning. The hope, for them, is to tease out cause and effect relationships and, based on their understanding of causality, to make changes to their teaching approaches that will facilitate student learning. This is in line with the tradition of behavioralism, which emerged in American psychology and economics.

The philosophy behind behaviorism is that human behavior can be altered by giving people different incentives and instructions. However, many social scientists, especially sociologists, have pointed out that this epistemological framework, transplanted from natural sciences, is problematic. Although human behavior can be changed, as all sociologists would agree, behavioral change is usually caused by long-term, structural processes. Education and socialization serve as good examples of these processes. Therefore, if teachers and researchers wish to improve teaching and learning in classrooms to? make a difference in the long term, it is imperative to first understand student behavior and how it comes about.

To understand student behavior in classrooms, this paper explores culture—more precisely, habitus—in big classrooms through sociological approaches. This means that I assume students' habitus in big classrooms is to some degree shaped by pre-existing social structure. More concretely, I focus on three dimensions: gender, race, and laptop use. By employing observational methods, I document students' distributions of seating and laptop use in big classrooms.

This paper is organized as follows. First, I review literature on student learning in classrooms. Second, I introduce the theoretical framework that I use to interpret my findings. Third, I briefly discuss the methodological approach to my research. Fourth, I present major findings that are based on my research. Finally, I discuss implications for teaching and future research.

Literature Review

Previous studies have shown that classroom space matters to student learning. For instance, in Classroom Spaces that Work (2001), Clayton highlighted the influence of classroom environments on pedagogical and learning outcomes. This study indicated that, to better understand student learning, it is crucial to place students' behavior in the context of classroom spaces. In other words, student learning can be thought of as the consequence of the interaction between students, instructors and classroom spaces. This perspective, therefore, is well connected with my research question: What are students doing during instruction within large lecture classrooms?

Literature on classroom composition is primarily concerned with the effects of physical classroom arrangements on student learning. For example, Leiter (1983) examined the effect of "ability composition" on learning, finding that the higher the ability level in a class, the greater the learning gains. More specifically, to test the hypothesis about the ability composition, Leiter used three variables as measures of ability composition: the ability homogeneity of the class, the achievement level of the class, and the interaction term between homogeneity and the achieve-

ment level. However, Leiter also added "proportion of white students in class" and "proportional of female students" as additional measures of ability composition. These measures are problematic in that race and gender are hardly "abilities," although they possibly can be proxies for variation in abilities. Moreover, Leiter's research design becomes more problematic when race and gender were also included as control variables. Leiter's results indicated that students in high achievement level classrooms performed significantly better in reading and mathematics. In addition, white students performed better in both reading and mathematics when compared with non-white students. By and large, Leiter's findings indicate that classroom composition influences student achievement.

Similarly to Leiter's work, Moss et al. (2014) found that classroom composition exerts significant influence on educational effectiveness. In this study, race is also included as one of the measures of classroom composition. These two studies indicate the importance of classroom composition for student learning.

Classroom composition is also employed to refer to other cultural dimensions in classrooms. Gender is a prominent one. For instance, Crombie and Gale (1999) compared all-female and mixed-gender classrooms of computer science students. Finding that students from the all-female classroom perceived greater teacher support than their counterparts, the authors concluded that gender composition does matter to female students' learning. Wolfe (1999) studied gender variation in learning experiences in computer-mediated classroom environments, finding that female students are less likely to be heard even though male students speak up less than females.

In addition to gender, racial, and ethnic compositions, there are also critical dimensions to consider when it comes to classroom dynamics. Comparing student achievement in racially integrated and segregated schools is an important strand of the sociology of education (e.g., Entwisle and Alexander 1994). Previous studies also examined the effect of classroom ethnic composition on student lives (e.g., Bellmore et al. 2007). With regard to the effect of racial or ethnic composition on student learning, Caldas and Bankston (1998) found that racial concentration—a phenomenon where students are racially self-segregated—has significantly negative effects on students' achievement. Moreover, classroom racial composition matters to interactions among students. Hallinan and colleagues (1982; 1985) found that classroom racial composition plays a crucial role in students' friendships and school life. Hoxby (2000)

evaluated how both gender and racial composition contribute to peer effects in classroom environments, finding that positive peer effects are stronger at the intra-race level.

Based on this literature review, gender, racial, and ethnic composition in learning environments is a prominent area for further exploration. However, the concept of physical classroom "space" seems to be largely ignored in that literature. Space, I argue, serves as an infrastructure for student learning and interaction. This means that students interact with or respond to not only other students but also the spaces of a classroom. As mentioned earlier, sitting in a big lecture hall is a unique experience for college students. In this sense, the questions of how students interact with or respond to a large, physical space should be given more attention.

Classroom seating serves as a good proxy for studying the culture of student learning in big classrooms. Becker and colleagues (1973) conducted three case studies studying whether class size and arrangement influence student participation. The settings they examined were a traditional lecture class, a laboratory setting, and a large lecture. They found that the sizes of classrooms did indeed matter to student participation. More specifically, in the third setting (big lecture hall), seating and position influence not only students' participation but also their interest in the course subject. Montello (1988) confirmed this association between seating locations and student participation, although he found no evidence suggesting an effect of seating on students' grades. Marx and colleagues (1999) also found a positive association between seating and participation in classroom discussion. Another relevant study, however, argued that students' achievement motivation influences their choices of seats as well as their participation in classrooms(Burda and Brooks 1996). In sum, based on my review of the literature, I focus on the relationship between classroom seating and student learning as one aspect of my observation of big classroom settings.

The other focus of my observations pertains to students' use of laptops and cell phones in big classrooms. Laptop and cell phone usage has been a prevalent, and sometimes problematic, cultural phenomenon of student life in college (Jenaro et al. 2007; Lepp, Barkley, and Karpinski 2014). Moreover, in large spaces like big lecture halls, students are likely to have more leeway to use laptops or cell phones. In this sense, the question of how laptop use affects student learning is important to answer. Fried (2008) found that the association between laptop use on campus and student learning is negative. Stu-

dents with their laptop out in class spend much more time multitasking and have more sources of distraction. Another study found that both students and faculty perceive cell phones as sources of distraction (Burn and Lohenry 2010). Tindell and Bohlander (2012) found that cell phone use, particularly texting, is prevalent among college students and most faculty are not aware of it. These results suggest that laptop and cell phone use is an important dimension for understanding students' learning culture in big classrooms.

Theoretical Framework

To investigate collective behavior, culture is a common term that has been used in sociology or social science more broadly for decades. The "tool kit" or "repertoire" theory of culture, originally proposed by Swidler (1986), has been enormously influential and popular in sociology in the past two decades. Based on a critique of Weberian and Parsonian value-deterministic models of culture's effects on action, Swidler argued that culture, rather than determining people's action, provides actors with toolkits or repertoires of strategies for constructing actions. Culture does not affect social actors by determining their goals, but enables them to build a variety of actions in response to changing environments. A slightly different take on Swidler's emphasis on culture as resources is Sewell's (1992) theory of structure that builds structural power into the concept of culture. Schemas (cultures) are a constituent part of structure, thereby empowering and constraining social agents. Swidler's and Sewell's theories of culture, although having different emphases, both perceive that culture offers agency to social actors. Nonetheless, the term "culture" is static, as many scholars have pointed out. People tend to think of culture as a static attribute of a certain group. For example, it is common to hear people use "White culture" and "Black culture" to describe different racial groups and these terms are usually derogatory and stereotypical. In order to avoid this problem, I use the concept of "habitus," created by French sociologist Pierre Bourdieu (1999), to describe students' collective behavior in big classrooms. Habitus is "a structuring structure, which organizes practices and the perception of practices" (Bourdieu 1989). More specifically, it refers to a "system of acquired dispositions functioning on the practical level as categories of perception and assessment or as classificatory principles as well as being the organizing principles of action" (Bourdieu 1989). Simply put, habitus shapes the way we interpret and engage in the world that we live in. In the context of my study, three types of habitus are

of particular focus: gender, race, and students' use electronic devices in big classrooms.

Operationally, these three forms of habitus are examined by observing seating in large lecture halls. I wanted to examine the ways that seating in classrooms may represent college students' perceptions of, and engagement in, gender, race, and the use of laptops.

Methods

I approach student habitus in big classrooms by using non-participant observational methods. That is, I went to big lecture halls on campus to make observations. The theoretical foundation of this paper's method is grounded theory (Corbin and Strauss 1990). Grounded theory is a sociological approach, which requires researchers to observe social phenomena without any pre-existing research question. In other words, observations are not guided by specific intention. This methodology is opposed to quantitative approaches, which intentionally direct researchers' observation of data collection.

Initially my observations were not guided by any particular theory. However, as this study unfolded, specific theoretical concerns emerged and I shifted my observations to certain dimensions of student habitus. Gender, race, and the use of electronic devices are three dimensions that crystallized as having particular importance as the study progressed. I specifically examined students' gender and racial habitus in big classrooms by documenting students' seating distribution. I assumed that the choice of a seat reflects a student's unconscious behavior, which is shaped by society. The choice of seating also serves as a proxy for understanding interactions among students. Thus, seating distribution at the macro level represents student habitus. The limitation of using non-participant observational approaches is that gender and race are not self-reported and not always obvious which is a weakness of this approach.

I also documented students' use of laptops in big classrooms, focusing on two dimensions of laptop use. First, I documented whether or not a student was using a laptop in class. Second, I observed what they were doing with their laptops (to the degree that I could determine it from my vantage point). I also informally chatted with students about the issue of laptop use in classrooms.
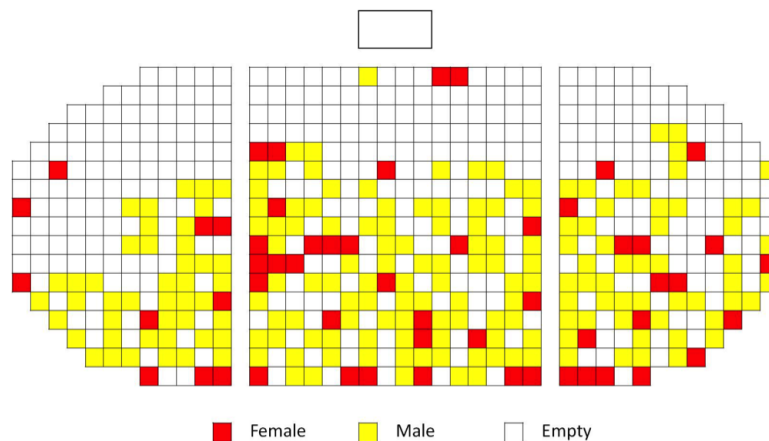
Results

Gender

Figure 1 represents gender distribution of individuals attending a social science class on March 15,

1: Self-selected seating arrangement by gender of students in a social science classroom.

2: Self-selected seating arrangement by gender of students in a natural science classroom.

2016. A square denotes a seat; this classroom contains nearly 600 seats. As the figure shows, students occupied less than one-third of the seats. One can see that the majority of students in this class are women. In this class it seems to be common for female students to sit together. In contrast, male students, who are in the gender minority, tend to sit alone as well as sit on the classroom margins. Other seating charts of the same class present a similar pattern of Figure 1 (not shown here). Overall, these findings demonstrate an intriguing example of gender segregation in big classrooms.

Figure 2 demonstrates another gender distribution in the same classroom but for a class in a different course, held March 2, 2016. In this natural ("hard") science course there are many more male students than their female counterparts. Interestingly, the gender composition of the class shown in Figure 2 is dramatically different from Figure 1, in which female students outnumbered their male counterparts in a social ("soft") science class. In Figure 2 female students tend to sit alone or sit with other female students and very few women sit with men. This pattern suggests another illustrative instance of gender segregation in classrooms.

In sum, Figures 1 and 2 present seating patterns by gender in big classrooms. The size of a classroom matters to student habitus because when there are much more seats than the total number of students, students have more choices of seats, which is more likely to reflect their unconscious behavior. This might not be observable in small classrooms where students have only limited options of seats. This privilege of observation in big classrooms makes it possible and more meaningful to document the distribution of seating in classrooms. Gender segregation is an emerging, significant pattern shown in Figures 1 and 2.

Race

Figure 3 represents the seating distribution by race and ethnicity in a social science class on March 21, 2016. Like previous figures, a square represents a seat. There are 424 seats in this classroom and about half of the seats are typically occupied by students. In this class, White students are the majority and other races are minorities. With regard to the seating patterns, Figure 3 suggests that Black students tend to sit either alone or with other Black students. Also, those Black students who sit alone occupy seats on the margins of the classroom. Asian and Hispanic students do not occupy as many marginal seats as Black students, while they tend to sit with students of the same race as well.

Figure 4 represents another example of seating distribution of race in a "hard" science class held on March 16. Note that this classroom is different from the one shown in Figure 3. Figure 4 shows that the majority of students in this class are White and the minority of students are of other racial groups. Asian students are the second largest group in this class and Black and Hispanic are very few. Similar to Figure 3, Figure 4 presents that students tend to sit with others of the same race. But unlike Figure 3, all the Black students in this class tend to either sit alone or with students of different race, which does not show a clear sign of self-segregation. Another difference between Figures 3 and 4 is the number of non-White students. In Figure 4, which is a hard science class, one can see that there are fewer Black students than Figure 3, which is a social science class. However, there seem to be more Asian students in the hard science course than in the social science one. A caveat is that this comparison might not be accurate because these two seating charts were drawn from different classrooms. But it is clear that in the social science class, the pattern of racial segregation in seating is more obvious than in the hard science class.

In summary, Figures 3 and 4 present seating patterns by race in big classrooms. Similar to gender segregation patterns shown in Figures 1 and 2, racial/ethnical segregation is also found in big classrooms.

Use of Laptops

Figure 5 represents students' use of laptops in a social science class. As the figure shows, most of the students who sit in the back of the classroom have their laptops out. By contrast, students who sit in the front of the classroom are less likely to use a laptop in class. These differential patterns of laptops use are very consistent over time during the semester.

According to my observations, students use laptops in class for a variety of purposes. Students use laptops to take notes, to do assignments, check email, chat with friends, or do on-line shopping. Some students use laptops to only take notes. But more students juggle note-taking with web browsing, checking email, or chatting online. In other words, it is common that students switched between note-taking and other tasks. In addition, it was observed that some students always looked at their screen and rarely (or never) looked up.

Discussion

Sitting in big classrooms is a common experience for college students. Big classrooms are also places where educators expect to see student learning and engagement happening. If making big classrooms more inclusive and engaging for student learning is a common goal for instructors, then it is crucial to understand student behavior, particularly unconscious behavior, in big classrooms in order to know how to construct inclusive learning environments. In other words, it is important to understand classroom dynamics before making changes.

First, my findings show that there are more male students than female counterparts in large lecture-style natural science classes. This is consistent with findings in research on gender segregation in college majors. England and Li (2006) argued that gender segregation in college majors was declining from 1971 to 2002, but the gap still exists and the trend of desegregation has been stalled. This suggests that gender inequality in higher education is a large-scale and structural phenomenon. Gender segregation in college majors can hardly be discovered if researchers only focus on a single class or correlations between teaching strategies and learning outcomes. More importantly, the issue of gender segregation in college majors matters to teaching because research on gender inequality in labor markets has consistently shown that there are "male" and "female" jobs (England 2005). This means that certain jobs are more open to a specific gender and students of the other gender will likely face barriers to entry into the field after they graduate. This structural limitation of gender segregation dramatically influences student learning (so called "stereotype threats", see Steele and Aronson 1995). If educators do not remain cognizant of this structural pattern of gender segregation, institutes of higher education will only reproduce existing gender inequality.
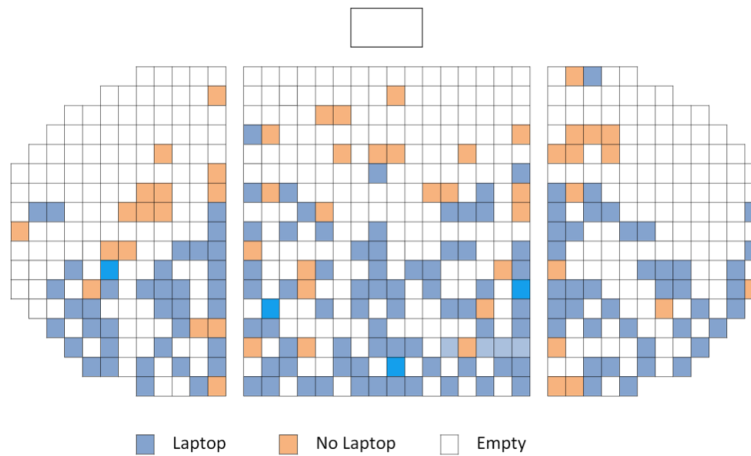
Second, similar to gender segregation, my findings show how students racially self-segregated in big classrooms. Racial segregation is a salient and
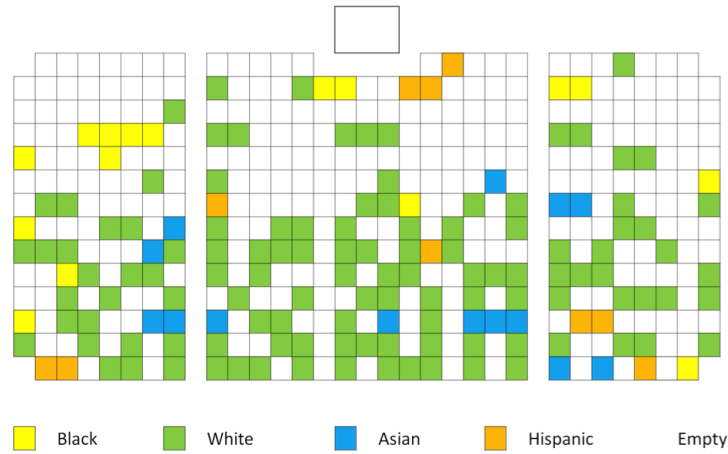
3: Self-selected seating arrangement by ethnicity of students in a social science classroom.
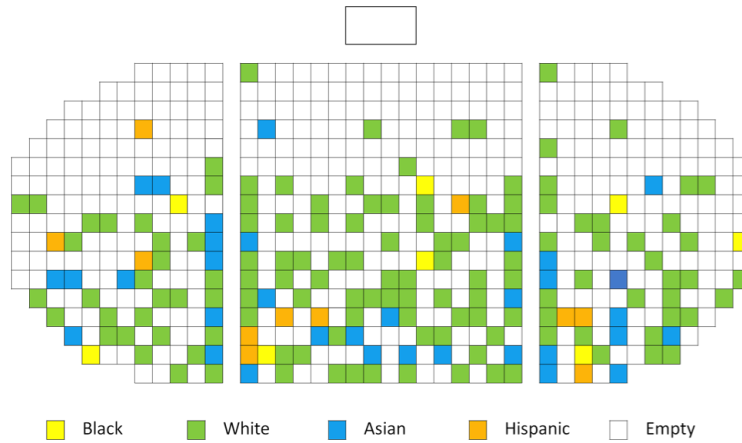
4: Self-selected seating arrangement by ethnicity of students in a natural science classroom.

5: Use of laptops in a social science classroom.

durable phenomenon in American society and one of the most important issues in sociology and social sciences more broadly (Charles 2003). In this sense, my finding that students tend to sit with people of the same race is not only consistent with previous studies of racial segregation more generally, but also suggests a fact that student behavior in particular is the mirror of social collective behavior in general. Racial segregation is particularly important in higher education because college has been shown to serve as a field for maintaining and reproducing social inequality (Armstrong and Hamilton 2013). For most educators, inclusiveness and diversity are two fundamental principles of teaching in college. It is, therefore, critical to pay more attention to the issue of racial self-segregation in classrooms.

Third, I have found a heterogeneous pattern of laptop use in classrooms—students use laptops for a variety of purposes. Many teachers wonder if the use of laptops has negative effects on student learning in classrooms. Even though this paper cannot answer this question, it is reasonable to speculate that laptop use has mixed effects on learning. If an instructor plans to create a laptop-free environment, the best strategy might not be banning laptops in classrooms. Rather, building a learning environment where students need not depend on or have any opportunity to use laptops is a more effective policy. In addition, the issue of laptop use in classrooms speaks to student culture on campus more generally. Students are not isolated from society, so their behavior in classrooms to a significant degree reflects their daily lives. Hanging out with friends and surfing on the Internet are essential part of daily life for college students. Before make any policies about laptop use in classrooms, it is important for teachers to learn what students do with laptops.

Finally, this paper suggests that future research on student learning can benefit from paying more attention to students' habitus in classroom. Even though habitus in classrooms may differ in significant ways based on schools, courses, and instructor-student interactions, understanding student habitus serves as a first and foremost step to better understand student learning. Just like other social actors, students come from different backgrounds, which dramatically shape their habitus. Student habitus also shapes ways in which they approach to higher education and learning. Educators have focused on the importance of learning outcomes in higher education. Learning outcomes are goals for the short term and it is not obvious whether or not these short term outcomes will influence students' lives in the long term. Although teachers can benefit practically from investigating "causal" relations between teaching and learning outcomes, those "causal relations" are likely to fade away after students go out of school. In this regard, more attention to student habitus is the first step for educators to construct a learning environment that is not only truly relevant for learners but is also more likely to make a difference in student learning in the long run.

References

Armstrong, Elizabeth A. and Laura T. Hamilton. 2013. Paying for the Party. Harvard University Press.

Becker, Franklin D., Robert Sommer, Joan Bee, and Bart Oxley. 1973. "College Classroom Ecology." Sociometry 36(4): 514–25.

Bellmore, Amy D., Adrienne Nishina, Melissa R. Witkow, Sandra Graham, and Jaana Juvonen. 2007. "The Influence of Classroom Ethnic Composition on Same- and Other-Ethnicity Peer Nominations in Middle School." Social Development 16(4): 720–40.

Bourdieu, Pierre. 1989. "Social Space and Symbolic Power." Sociological theory 7(1): 14–25.

Bourdieu, Pierre. 1999. Language and Symbolic Power. Harvard University Press.

Burda, Jeffrey M. and Charles I. Brooks. 1996. "College Classroom Seating Position and Changes in Achievement Motivation over a Semester." Psychological Reports 78(1): 331–36.

Burns, Shari M. and Kevin Lohenry. 2010. "Cellular Phone Use in Class: Implications for Teaching and Learning a Pilot Study." College Student Journal 44(3): 805–10.

Caldas, Stephen J. and Carl Bankston. 1998. "The Inequality of Separation: Racial Composition of Schools and Academic Achievement." Educational Administration Quarterly 34(4): 533–57.

Charles, Camille Zubrinsky. 2003. "The Dynamics of Racial Residential Segregation." Annual Review of Sociology 29(1): 167–207.

Clayton, Marlynn K. 2001. Classroom Spaces That Work. 1st edition. Greenfield, MA: Center for Responsive Schools, Inc.

Corbin, Juliet M. and Anselm Strauss. 1990. "Grounded Theory Research: Procedures, Canons, and Evaluative Criteria." Qualitative Sociology 13(1): 3–21.

Crombie, Gail and Patrick Ian Armstrong. 1999. "Effects of Classroom Gender Composition on Adolescents' Computer-Related Attitudes and Future Intentions." Journal of Educational Computing Research 20(4): 317–27.

England, Paula. 2005. "Gender Inequality in Labor Markets: The Role of Motherhood and Segregation." Social Politics: International Studies in Gender, State & Society 12(2): 264–88.

England, Paula and Su Li. 2006. "Desegregation Stalled The Changing Gender Composition of College Majors, 1971-2002." Gender & Society 20(5): 657–77.

Entwisle, Doris R. and Karl L. Alexander. 1994. "Winter Setback: The Racial Composition of Schools and Learning to Read." American Sociological Review 59(3): 446–60.

Fried, Carrie B. 2008. "In-Class Laptop Use and Its Effects on Student Learning." Computers & Education 50(3): 906–14.

Hallinan, Maureen T. 1982. "Classroom Racial Composition and Children's Friendships." Social Forces 61(1): 56–72.

Hallinan, Maureen T. and Stevens S. Smith. 1985. "The Effects of Classroom Racial Composition on Students' Interracial Friendliness." Social Psychology Quarterly 48(1): 3–16.

Hoxby, Caroline. 2000. Peer Effects in the Classroom: Learning from Gender and Race Variation. National Bureau of Economic Research. Retrieved March 14, 2016 (http://www.nber.org/papers/w7867).

Jenaro, Cristina, Noelia Flores, María Gómez-Vela, Francisca González-Gil, and Cristina Caballo. 2007. "Problematic Internet and Cell-Phone Use: Psychological, Behavioral, and Health Correlates." Addiction Research & Theory 15(3): 309–20.

Leiter, Jeffrey. 1983. "Classroom Composition and Achievement Gains." Sociology of Education 56(3): 126–32.

Lepp, Andrew, Jacob E. Barkley, and Aryn C. Karpinski. 2014. "The Relationship between Cell Phone Use, Academic Performance, Anxiety, and Satisfaction with Life in College Students." Computers in Human Behavior 31: 343–50.

Marx, Alexandra, Urs Fuhrer, and Terry Hartig. 1999. "Effects of Classroom Seating Arrangements on Children's Question-Asking." Learning Environments Research 2(3):249–63.

Montello, Daniel R. 1988. "Classroom Seating Location and Its Effect on Course Achievement, Participation, and Attitudes." Journal of Environmental Psychology 8(2): 149–57.

Moss, Brian G., Ben Kelcey, and Nancy Showers. 2014. "Does Classroom Composition Matter? College Classrooms as Moderators of Developmental Education Effectiveness." Community College Review 0091552114529153.

Sewell, William H., Jr. 1992. "A Theory of Structure: Duality, Agency, and Transformation." American Journal of Sociology 98(1): 1–29.

Steele, Claude M. and Joshua Aronson. 1995. "Stereotype Threat and the Intellectual Test Performance of African Americans." Journal of Personality and Social Psychology 69(5): 797–811.

Swidler, Ann. 1986. "Culture in Action: Symbols and Strategies." American Sociological Review 273–86.

Tindell, Deborah R. and Robert W. Bohlander. 2012. "The Use and Abuse of Cell Phones and Text Messaging in the Classroom: A Survey of College Students." College Teaching 60(1): 1–9.

Wolfe, Joanna L. 1999. "Why Do Women Feel Ignored? Gender Differences in Computer-Mediated Classroom Interactions." Computers and Composition 16(1): 153–66.

Experts in the Classroom: Informing Undergraduate Career Choices in Biomedical Engineering

Yehudah Pardo[1,*] and Kimberly Williams[2]

[1] Meinig School of Biomedical Engineering, Cornell University
[2] Center for Teaching Innovation, Cornell University
[*] Corresponding Author: yap6@cornell.edu

Abstract

Biomedical engineering (BME) consists of an amalgam of fields focused on a single goal of improving medical technology rather than a single approach or set of techniques. Education of undergraduate BME students about the various career options possible in the field of BME is therefore essential to ensure that the students can pursue the opportunities and training that support their goals. We organized expert testimonials from professionals in BME in an introductory BME course to inform the students about each different career path. We surveyed the students before and after the testimonials and discovered that, following the testimonials, approximately 15% of students changed their career aspirations. Furthermore, qualitative analysis of open-ended survey questions and observation notes from a focus group revealed that the testimonials encouraged the students to critically think about their future career plans. Overall, we found that the testimonials provide an effective method of introducing students to a wide variety of career paths in a manner accessible to a wide variety of classroom environments. Our study could potentially act as a framework and inspiration for future studies to develop creative methods for informing undergraduate career selection, even beyond BME.

Introduction

Biomedical Engineering (BME), also known as Bioengineering, is perhaps unique among engineering fields in its lack of definition (Lisenmeier, 2003). Indeed, working definitions of BME are general and serve to highlight the diverse nature of activities considered aspects of BME. For example, Imperial College in London defined BME as: "...a rapidly evolving interdisciplinary field that applies engineering principles and technology to medical and biological problems. With an ageing population and advances in technology, bioengineering plays an integral role in global issues such as healthcare, energy and the environment" ("Imperial College London Department of Bioengineering," 2017).

Similarly, a working definition of bioengineering provided by the National institute of Health (NIH) demonstrates the diverse nature of activities considered as aspects of the field of BME: "The application of concepts and methods of engineering, biology, medicine, physiology, physics, materials science, chemistry, mathematics and computer sciences to develop methods and technologies to solve health problems in humans" ("Science Education Glossary of Terms," 2017).

The broad and ill-defined nature of BME is a significant challenge for engineering educators who are tasked with preparing the next generation of biomedical engineers for the work force. In an ideal world,

BME students should graduate university with the same or a similar set of basic skills ensuring that the possession of a Bachelor's degree in BME is indicative of the possession of a certain basic skill set. Unfortunately, this does not seem to be the case. In fact, there is no single BME course that is required by all current BME programs (Linsenmeier, Harris, & Olds, 2002). This lack of clarity poses a problem both for employers as well as the students themselves. Employers cannot know what to expect of graduates of an ill-defined BME program, while the students may be left with a feeling of being a "jack-of-all-trades, yet master of none" and a general lack of direction for selecting career paths.

BME is no better defined outside of university than it is as a major. Biomedical engineers can follow a diverse array of career paths requiring many different sets of skills and knowledge. Engineering educators are therefore faced with the dual challenge of creating an educational system with a consistent result while simultaneously preparing students for a wide variety of career options.

A Broad-Scale Approach

One approach toward mitigating the lack of consistency associated with BME education is to identify a set of skills critical for biomedical engineers to possess and redesign curricula to emphasize these skills. The VaNTH engineering research center (ERC) was organized with this idea in mind. The VaNTH ERC

is a consortium of a few main educational institutions (Vanderbilt, Northwestern, University of Texas at Austin, and Health, Science, and Technology at Harvard/MIT) with several other participating institutions that together have attempted to define BME curricula as well as teaching methods.

Rather than define a specific set of courses that should be adopted by the different BME programs across the country, the VaNTH ERC created a set of core content that could be fulfilled by a diversity of classes (Linsenmeier et al., 2002). In particular, they identified the following eight categories of BME related knowledge:

1. Signals and systems analysis
2. Instrumentation and electronics
3. Imaging
4. Chemical physics and thermodynamics
5. Materials
6. Mechanics
7. Fluid mechanics and transport
8. Applied mathematics, e.g. statistics & modeling

In addition, they proposed that these "traditional engineering" topics be presented in the context of biological systems and focus on biological applications. They cited anecdotal evidence that such an approach tended to motivate BME students and enhanced their ability to apply their knowledge to biomedical problems (Linsenmeier et al., 2002).

The VaNTH ERC also focused on implementing new teaching methods to better transmit the core subjects to BME students. An innovative pedagogical framework known as "How People Learn" (HPL) was developed in support of the core BME curriculum (Bransford, Brown, & Cocking, 1999). The HPL framework is based on four basic principles: learning should be knowledge centered (new knowledge is focused on applicable skills), assessment centered (formative and summative assessments are applied to track student progress), learner centered (takes advantage of student prior knowledge), and student exercises should be based on real-world problems. In an effort to evaluate the effectiveness of this approach in BME education, a study was constructed around a class on biotransport, a typical required course in BME (Martin, Rivale, & Diller, 2007). Learning outcomes after implementing HPL were directly compared to traditional teaching and the results quantitatively assessed through pre- and post-tests. Interestingly, both methods seemed to result in similar knowledge growth; however, in the HPL class students improved in innovation as well, while in the tra-

ditional class a decrease in innovation was observed.

The effect of the HPL framework was not limited to students. Faculty and postdocs participating in the VaNTH ERC were surveyed to determine how HPL affected their teaching as well as their own career decisions (Cox, 2011). Participation impacted a number of aspects of the faculty members' careers, including their research, their student interactions, and their teaching. They found themselves employing the HPL principles with a greater frequency and generally reported a positive experience.

The studies of the VaNTH ERC illustrate that broad scale organization of BME educational practices can and does have a positive impact on students and faculty. However, the VaNTH program is designed to improve the overall quality of the BME workforce rather than inform individual student career decisions or help students achieve their personal career aspirations. In order to accomplish these equally worthwhile goals, a bottom-up approach is needed, starting from the students themselves rather than institutional-scale reorganization.

Early Career Preparation: Pre-College Programs and Activities

Although the present study focuses on an intervention applied to college-age students, a significant amount of research has focused on pre-college programs and activities and their corresponding effect on student career decisions (Sadler, Sonnert, Hazari, & Tai, 2012; Savelsbergh et al., 2016). These studies can inform the development of programs designed for college-age students and improve their quality and efficacy.

In one study, out-of-school experiences of high school students were investigated relative to their ability to predict the engineering discipline chosen by the student in college (Godwin, 2016). The authors of the study surveyed students in mandatory first-year English classes to measure their participation in experiences such as tinkering with mechanical or electrical devices, experimenting with chemical sets or kitchen chemistry, astronomical observation, and caring for animals, among several others, during their high school years. They found that many of the experiences predicted not only the choice to enter the engineering major but the specific type of engineering as well. In addition, the authors examined gender differences and found that participation in more stereotypically "female" activities tended to predict a choice of major that has a relatively higher proportion of females (Bioengineering, environmental engineering, chemical engineering).

Classroom experiences can also influence student

attitudes toward science. An innovative approach toward science learning, known as STS (science, technology, and society), makes science learning a student-centered experience (Akcay et al., 2010). Students generate their own questions and are guided through an inquiry process enhancing their interest in the scientific field. In a study of 6th through 9th graders, STS was found to significantly improve attitudes toward science (Akcay et al., 2010). Furthermore, the STS approach was found to work equally well for both male and female students, an important result given the historic underrepresentation of females in the STEM fields.

Both of these studies share a common theme: hands-on experience. Practical experiences outside of a formal educational setting had a significant impact on student career decisions. Likewise, a classroom setting that simulates the scientific method (inquiry generation and investigation) enhanced student interest in science as a subject. The logical conclusion is that students are more influenced by practical experiences with subject matter than knowledge transferred in a more abstract manner. This suggests that undergraduate programs focused on practical exposure to engineering careers may be more effective than other methods of introducing students to different career options.

Experiential Learning in BME

In fact, experiential learning, rather than learning solely via knowledge transmission in the classroom, has been shown to have direct and measurable benefits for engineering students in a college setting. Blair and Millea (2004) showed that engineering students who participated in cooperative internships (co-ops) had both a higher GPA and a higher starting salary than students who did not participate in such programs. These results were corroborated by Dressler and Keeling (2004) who found that in addition to higher GPA, students participating in cooperative education programs displayed improved analytical thinking and better problem solving skills.

Co-ops and Internships

The course load that is typical of engineering degree requirements hampers incorporation of co-ops into BME curricula (S. Krishnan, 2010). Simply put, it is difficult to find the time to establish a worthwhile internship while keeping graduation time to the four years expected in the U.S. higher education system. Professor Shankar Krishnan proposed a number of models for how such programs could be included in BME education (S. Krishnan, 2010; S. M. Krishnan, 2015; S. M. Krishnan & Asee, 2014). Co-op programs begin with appropriate collaborators that can host students and provide sufficient mentorship. Krishnan recommends that multiple collaborators be simultaneously included to take into account the limited resources of each collaborator, differences in student training and interests as well as providing a diversity of experiences in the various areas of BME. Such collaborators might include: teaching hospitals, medical equipment manufacturers, biomedical research institutes, and regulatory agencies, among others (S. Krishnan, 2010).

In a later paper, Krishnan describes four distinct models for incorporating co-ops into semester-based degree programs, quarter-based degree programs, and even international five-year programs (S. M. Krishnan & Asee, 2014). In the first model, co-ops are restricted to summer sessions. While this approach avoids taking away time from required coursework, collaborating institutions receive less benefit from the student interns, as their time at the institution is limited. Model 2 splits the co-op experience across three different semesters. Both the students and the cooperative institutions receive additional benefit from the increased time spent on the internship. However, some coursework has to be relegated to the summertime, putting increased pressure on the students. Models 3 and 4 are for quarter-based systems. Model three relegates two consecutive quarters for the co-op while model 4 is designed for a five-year degree program and splits the co-op across a total of six quarters. The various models for co-op experiences were implemented at a variety of universities beginning in the state of Illinois but later expanding across the U.S. and to schools in Singapore, Germany, Canada, India, and France.

The effectiveness of the co-op programs was evaluated through three sets of student interviews at the beginning, middle, and end of the internship. Overall Krishnan and colleagues (2014) reported that greater than 90% of students found that the co-ops "enhanced their overall learning". In addition, he found that students who participated in co-ops found permanent jobs faster and more easily than their peers upon graduation.

Clearly, cooperative programs have enormous potential to not only inform student career decisions but also enhance their career success. There are however limitations to this approach beyond the time commitment. First, the quality of the student's experience is highly dependent on the site of the co-op or internship. The quality of the mentorship and type of institution hosting the co-op are critical factors in determining the benefit to the student. In addition, a given internship typically provides insight into only

a single career path. Thus, students gain in depth knowledge in a particular area, but not the broad experience necessary to make a choice between different career options. Finally, many of the benefits of co-ops are most relevant to industry related careers. Industry partners benefit from hiring student workers and students gain valuable connections by working at companies. Krishnan et al. (2014) noted that many co-op students chose to work at the companies with which they interned. While for students interested in industry such opportunities are highly beneficial, for students interested in pursuing the other aspects of BME, industry co-ops may be of limited value.

BME Experience in the Clinic

Alternative programs have been developed to expose students to BME careers beyond industry. A program initiated at the University of Pennsylvania (U Penn) was designed to provide BME undergraduate students with an in-depth insight into the clinical aspect of BME (Davies & Litt, 2006). The clinic was chosen because the authors felt that such direct interaction with clinicians and patients is particularly rare at the undergraduate level. The program was implemented as a sixteen-week course for upper division BME students. The first eight weeks consisted of lectures given by various faculty to introduce students to different clinical areas and to prepare them for the final ten weeks of the course (which overlapped with the final two weeks of the lectures) during which students participate in a clinical preceptorship. This is a mentored experience in which students were paired with a clinical faculty member. During the preceptorship students were exposed to patients, clinical practices and clinical equipment and technology. The program began in 2000 with a few students and expanded to 48 students by 2005. In order to keep track of student progress and accomplishments, midterm exams and a final presentation were assigned. In addition, informal interviews were conducted with both students and faculty. The faculty reported a positive experience in general and that students obtained "satisfactory breadth and depth of understanding of the lecture material and considerable depth in the preceptorship specialty". The authors indicated an interest in following graduates of the program to see how it affected their future career paths; however, at the time of the study insufficient data was available to make any conclusions. They did note that in the beginning approximately 25% of students intended to go to medical school post-graduation and in later years of the program, the authors expected that number to increase to 40%.

The preceptorship program at U Penn provides a unique opportunity for students to gain insight into the clinical aspects of the field of BME. Although a significant percentage of the participating students intended to go on to medical school, such a program has value for students intending to work in the medical device industry or as biomedical engineers in a clinical setting.

BME Experience in Research

The U Penn program arose as a collaboration with the teaching hospital associated with the university. The University of Texas M.D. Cancer Center (UTMDACC) developed its own internship program that gave students a more research-oriented experience (Wright, 2007). The UTMDACC is an independent research and clinical institution that does not grant degrees of any sort. Student interns come from a variety of educational institutions both within the United States and internationally. Over the first ten years since its inception, survey and demographic data were collected from the student participants and then evaluated to determine the effectiveness of the internship program. Likert scale-based responses showed that overall the students found that the internship enhanced their skills in six critical areas: analytical skills, communication, creative problem solving, research skills/ technical competence, self-learning, and teamwork. The authors noted however, that of these areas creative problem solving received the lowest score (4.04 out of 5) suggesting that perhaps it merited additional attention in the future. From the more open-ended questions in the survey, the researchers discovered that the internship was able to clarify career decisions for many of the students; helping them to make decisions such as whether to pursue a research career or attend medical school.

Areas for Future Work and Our Approach

Biomedical engineering educators have a responsibility to BME students to provide the training they need to succeed in their future careers. Unfortunately, it is difficult to define what that training entails due to the cross-disciplinary nature of the BME field. Furthermore, no single skill set is likely to be applicable to all of the BME career paths available. The VaNTH engineering research center has taken a top-down approach toward solving the former problem by defining a general curriculum and implementing novel teaching methodologies in BME. In addition, programs have been developed to offer students opportunities to gain hands-on experience with different BME career paths, allowing them to focus on developing the skills needed for a particular career. The value of such programs is undeniable. Researchers

have reported improved analytical and problem solving skills, a greater depth of understanding of BME applications and careers, and even increased likelihood of finding a permanent job immediately after graduation.

However, these programs are typically focused on a single in-depth experience. Furthermore, the time commitment required by such programs limits the ability of students to participate in more than one or two experiences. Thus, there is a need for educational programs that inform students about the breadth of career options available within the BME field. Such a program would act as a primer for more in-depth experiences, allowing the students to intelligently select the experience that best fits their interests and ultimately achieve their career aspirations.

In this study we present a novel approach to inform undergraduate career choices within the confines of a typical classroom setting. We focused on developing a method that could be implemented with minimal time commitment, yet achieve maximum benefit and provide freshman-engineering students with a broad perspective on BME career options. We implemented our approach and analyzed its effectiveness in a single introductory biomedical engineering class at Cornell University. However, the implications of this study reach beyond this particular example. We present lessons learned from our study that will hopefully inspire and inform future studies to provide lower division university students with the resources necessary to make informed decisions in choosing their future careers.

Literature Review

Difficulties in Teaching Biomechanics

Huston (1999) articulates many of the issues that make biomechanics difficult to teach. From my perspective, the two most critical aspects are: (1) students not having sufficient prerequisite background when starting a biomechanics course, and (2) students not understanding the expectations on them when they "discover how broad biomechanics is" leading to them "often [becoming] discouraged." Class time should be set aside as needed to cover any remedial review: either through an intense and thorough review of mechanics or basic biological concepts that students might be missing entirely. Huston also suggests that the main application areas of biomechanics are performance, injury, and rehabilitation. Therefore, learning objectives should coincide with these areas. Yet a key issue with teaching biomechanics is the integral connection to basic mechanics. Knudson (2016) observes: "Teaching Newtonian mechanics in physics and an introductory biomechanics class can be difficult given that many students fear the course and often avoid it as long as possible." Framing problems in a biomechanics classroom is difficult, especially in classes that are required to cover multiple biological systems. For instance, in the case of understanding the physics of walking, one could decide to lecture on the physics of joints, tendons, and muscles, or to focus on the kinematics and general forces of moving a leg. Since an entire class could be devoted to this subject and interrelating the different functional scales, problems must be framed clearly to avoid distractions and must clearly identify an objective, such as understanding injury, that can be realized within the limited framework.

Another key aspect in engineering is estimating quantities that are not known exactly. Such estimations are usually made using information from available evidence and literature but may also need to be done rapidly in the field with incomplete information. Shakerin (2006) provides activities to help assess and practice estimation problems for engineering students. In this study, students demonstrated a poor ability to make estimations. Students struggled due to two elements of estimation: (1) by default these are situations where they do not have all information required, and (2) there is not necessarily a right answer to get the question "correct". When students were tasked with estimating quantities for which known values are available they are usually not accurate. Estimations are often required to form the boundary conditions of biomechanical problems. Therefore, during my module, students need to estimate quantities while they learn to setup and solve complex mechanics.

Specific Teaching Methodologies

Scott, Volz, and Athanasiou (2006) present a course project that heavily uses a teaching as learning strategy. Students must learn a mechanical model and explain it to the class using a classic lecture style: with a whiteboard or chalkboard and colored markers or chalk. The obligation to "prepare others to use information immediately to solve problems" is cited as a main driver of student learning over simply giving a PowerPoint presentation, which may cause students to they think they are simply being evaluated (Scott et al., 2006). Overall, students found this experience worthwhile, even when surveyed one year later. Students state that the main benefit was learning how to teach a lecture and how much work is required to do so, with one student explicitly stating that it forced him to learn material taught in the course. The effectiveness of student teaching can be capitalized in any classroom without needing to flip the classroom

completely. Students can act as teachers when class time is provided for think-pair-share (TPS) activities. In this case, students are aware during the "think" stage that they will need to present their approach to a partner and not only work on the problem but think about how they can articulate their own successes and confusions. An appropriate time to use TPS could be at the start of solving a problem, where students can complement each other by inherently using different methods to reach the same end or by getting stuck at different parts of a problem.

Another challenge in teaching biomechanics is that often our final models contain many complex parts that would be overwhelming to learn all at once. An equally competing problem is that smaller aspects of the model system may seem meaningless without this context. Clyne and Billiar (2016) state: "In problem based learning, the problem is introduced first, and it then becomes the motivation for learning the subject material." Their study focuses on a student-centered approach to problem based learning where the professor is a facilitator to a larger discussion. Their study states that faculty who are accustomed to lecturing can slowly add active learning sections to their lectures. Presenting the fully realized problem at the start of a class or course module is a method itself and does not necessitate an active learning strategy. However, assessing students' initial understanding of the problem can help the instructor use the remaining lecture time more efficiently, in addition to the use of TPS.

Savinainen and others (2013) investigate different methods of representing physical problems. Their study also identifies the value of pairing students, which allows them to feel comfortable making mistakes in front of their peers. Overall, students should be encouraged to draw depictions of physical systems that are most appropriate for representing that system, such as using free body diagrams that show forces or kinetic diagrams that show accelerations.

Methods

The goals of this study were threefold. First, we hoped to provide freshman with an opportunity to learn about the breadth of different career paths in biomedical engineering. Second, we hoped to facilitate the students' choice of career path by providing in depth information about different BME career paths. Finally, we hoped to create a program that could be integrated into a typical curriculum with minimal time commitment for maximum benefit. Taking these goals into consideration we designed a one-week module for an introductory, freshman-level BME course during which professionals in vari-

ous BME fields presented their work to the students in an interactive question and answer format. These interactive "testimonials" thus provide direct insight into the day-to-day aspects of BME careers. The first author of this study was a guest lecturer in this course and was therefore able to observe the proceedings without inducing student angst over grades. In order to incorporate a breadth of knowledge, professionals in medical, research, and community relations-based BME careers presented to the students. Furthermore, the two testimonial sessions were each organized around a specific theme, either cancer or solid organ transplantation. Thus, each session provided significant depth into a particular topic. Finally, the fact that the entire experience was condensed into two lessons makes this format accessible to nearly any BME class. While unlikely to provide the same benefit as a co-op or internship, this experience requires far less time commitment or administrative organization than such formal programs.

Data Collection

The Cornell University Institutional Review Board approved all procedures. We employed both quantitative and qualitative methods to analyze the effectiveness of our program and the degree to which the testimonials fulfilled our goals. Primarily we relied on written surveys filled out by the students a few days prior and a few days after the testimonial sessions. The surveys were intended to uncover two main categories of information. The first several questions were designed to measure how the testimonials affected the students' career aspirations, while the final few questions were directed toward measuring the students' perceptions of their career aspirations and the career selection process.

In order to determine if the testimonials had any effect on the students' career aspirations, we asked them to select their intended field and intended job both before and after the testimonial sessions. The students were provided with a list of twelve STEM fields from which to choose and fifteen job types. In addition, students were asked to rate the degree to which they felt certain about these choices both before and after the testimonials.

The majority of the survey questions were quantitatively rated on a Likert scale. However, at the end of each survey qualitative questions were included to gain a greater depth of understanding into how the testimonials were perceived by the students. The survey questions are provided in the supplementary methods section.

In addition to the surveys, a focus group was conducted after the completion of the testimonials to

gain additional qualitative insight. To avoid leading the students to any one perception or conclusion, the focus group was conducted as an informal discussion. Questions were prepared ahead of time to help direct the conversation, but there was no rigidly followed script. Finally, during the testimonials themselves, observations were made to understand the degree and type of student interaction with the professionals and overall engagement during the session.

Data Analysis

We applied a variety of statistical tests to analyze the quantitative data. Tests implemented in this study included permutation chi-squared tests, paired t-tests, ANOVA, and linear correlation. The results section specifies the particular test applied to each quantitative data set. All statistical analyses were performed using the open-source software R. Qualitative data such as responses to open-ended survey questions and observer notes taken during the focus group were coded and systematically analyzed to identify recurrent themes.

Results

Changes in Student Career Aspirations

We were interested in answering three questions: (1) how did the testimonials affect the students' career aspirations?, (2) How did the testimonials affect student attitudes toward career selection?, and (3) What were student attitudes towards the testimonial format itself? To answer the first question, we asked students to identify their career aspirations from a defined list of possible choices both before and after the testimonial sessions. Rather than define a single set of career options, we split career aspirations into field of interest (e.g. BME, Medicine, Physics, etc.) and job type of interest (e.g. Medical professional, Professor, Industry, etc.) (Fig. 1A and Fig 1B). We compared the student selections in the pre and post-testimonial surveys to see if their career aspirations had changed either with regard to either field or job type. We separately counted the number of students who selected a different choice in the pre and post surveys for field and job type and calculated Clopper-Pearson confidence intervals for the percentage of students that changed their choice (see Fig 1C). On average 17.2% (95% CI: 5.8% to 35.8%) selected a different field and 14.8% (95% CI: 4.2% to 33.7%) of students selected a different job type after the testimonial sessions. It should be noted that although there were 30 respondents to each survey, some students selected more than one choice for either field or job type and consequently had to be eliminated from the analysis. This left 29 students for field and 27 students for job type.

Different students were expected to be more or less certain of their career selections. We asked students to rate their certainty regarding field choice and job type choice separately on a scale of 1 to 5, with 1 being completely unsure and 5 being 100% certain (see Fig. 2). In order to determine if student certainty changed following the testimonials we performed a two-way ANOVA, including field/job as well as pre/post testimonial sessions as factors. We found that certainty did not significantly change following the testimonial sessions. However, students were significantly more certain of their field choice than their selection of job type ($p=0.008$).

Although there was no significant change in mean certainty for either field or job choice before and after the testimonial sessions, we hypothesized that the certainty would be negatively correlated with the students' likelihood to change their career selection. In order to test this hypothesis, we performed a permutation chi-squared test to compare the likelihood that a student would change his/her field or job selection given the student's category of certainty (1 through 5). We found that student certainty significantly affected the likelihood of changing field selection ($p = 0.0025$), while certainty did not significantly affect the likelihood of changing job selection ($p=0.127$).

In addition to certainty, we asked students to identify the time at which they originally made the career choices indicated on the survey (Fig. 3). Students were provided with four possible options: still unsure, during college, during high school, or before high school. We hypothesized that: (1) the earlier the decision was made the more certain the students would be of their decision, and (2) the less likely they would be to change. To test the latter hypothesis, we again performed a permutation chi-squared test, treating each selection time (college, high school, etc.) as an independent category. We found that decision time did not significantly affect the likelihood of changing either field or job type. To test the former hypothesis, we treated both certainty and decision time as continuous variables (the higher the number the greater the time in the past that students made the decision). We performed linear correlation and found that a modest positive correlation existed between certainty and decision time for both field and job type (Pearson's correlation coefficients of 0.57 and 0.51, respectively). In both cases the correlation coefficients were significantly different from zero, thus we concluded that there was a general tendency for students to be more certain of their choice with increasing time since the decision was originally made.
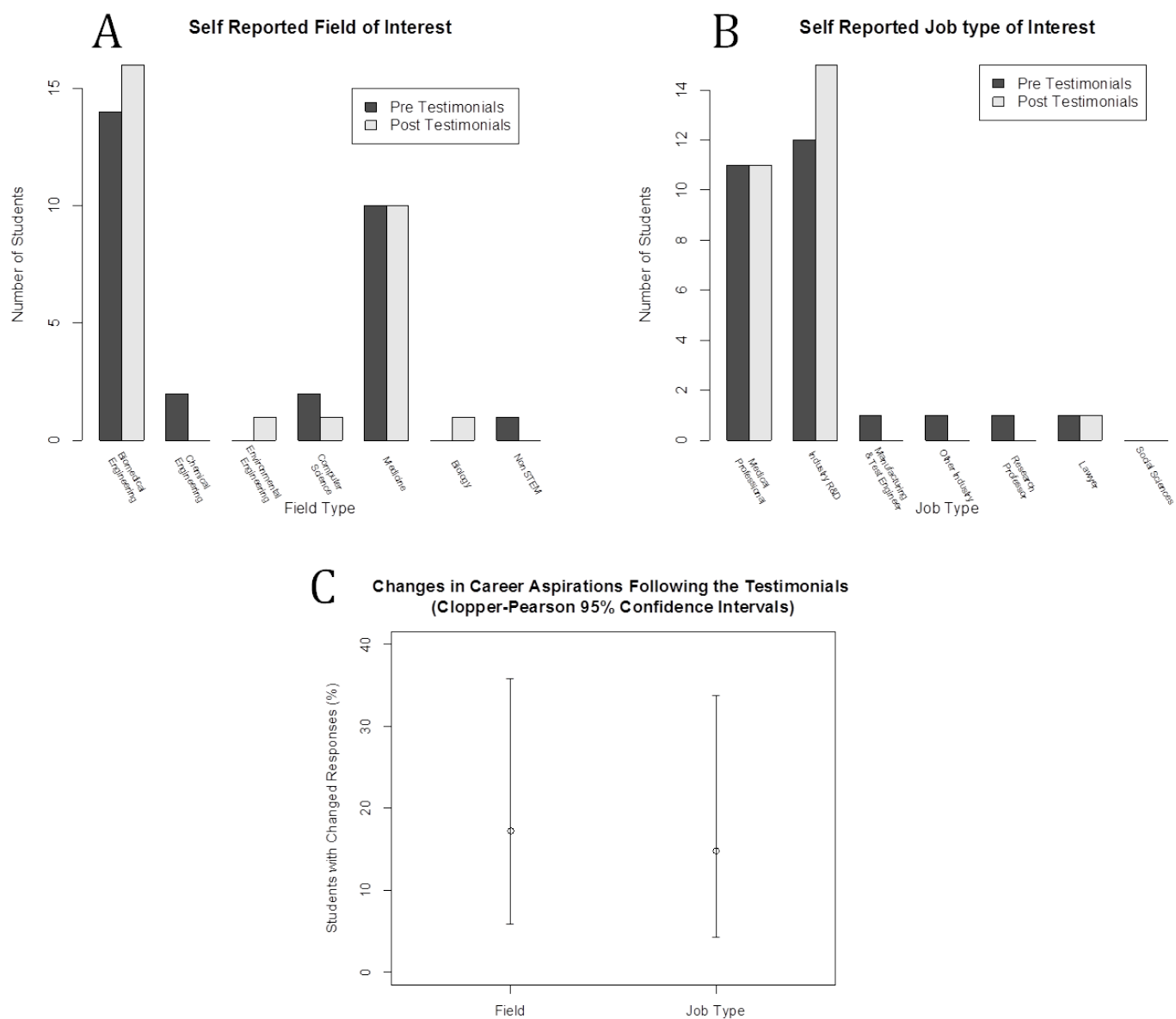
Figure 1: Effect of the testimonial sessions on student career aspirations. (A) Self-reported field of interest before and after the testimonials. (B) Self-reported job type of interest before and after the testimonials. (C) Clopper-Pearson confidence intervals for the percentage of students that changed with field of interest or job type of interest following the testimonials.
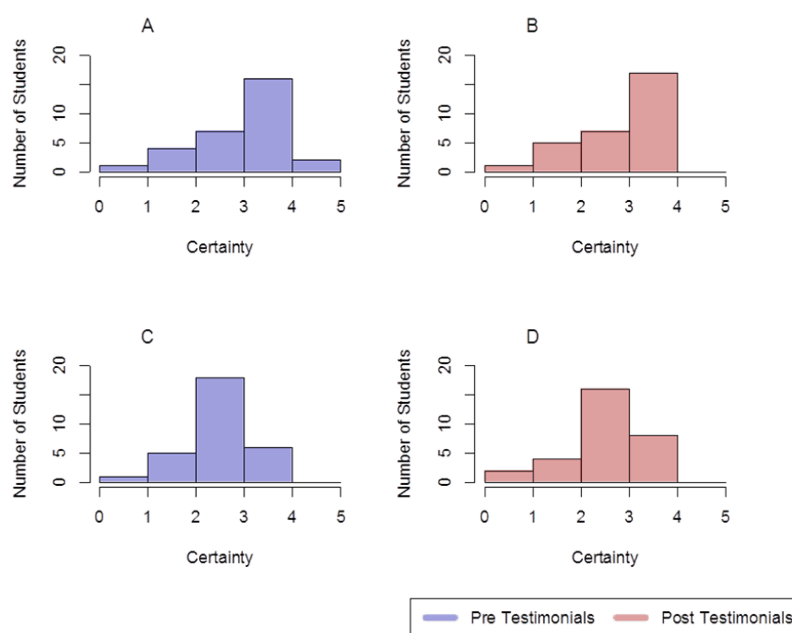
Figure 2: The distribution of student certainty before and after the testimonials for both field (A, B) and job type (C, D).

Finally, we asked if students had decided upon their field of interest at a different time than job type. We compared mean decision time for field and job type using a paired t-test and found that students selected field earlier than job type (p=0.007).

Student Attitudes Towards Career Selection

Having measured the effect of the testimonials on the students' career aspirations, we next considered student attitudes towards selecting future a career, and how the testimonials might have affected these attitudes. Both before and after the testimonials, we asked students to indicate the degree to which they agreed with the following statements: "I feel well prepared to select a career.", "Practical (i.e. hands-on or direct engagement with professionals) experience is necessary to choose an appropriate career.", and "I believe that learning about a field in school is sufficient to select a career."

Students indicated their agreement on a scale of 1 to 5 with 1 indicating, "completely disagree", and 5 indicating "completely agree". For all three questions, there was no significant difference in mean agreement before and after the testimonials (p= 0.541, p=0.698, p=1, respectively). However, the general distributions of the responses are informative. The responses to the first question (preparedness for career selection) followed a normal-like distribution with a mean of approximately 3. This suggested that the students did not feel fully ready to select a future career at their current stage of education (second semester, freshman year of college). By contrast, the

students' responses to the question about the necessity of practical experience were very asymmetrical. The vast majority of students selected either a 4 or a 5 indicating that there was a strongly held belief among the students that practical experience is indeed necessary for career selection. The responses to the third question ("is learning about a field in school sufficient?") were relatively speaking more uniformly distributed. While the mean response did not significantly change, it should be noted that there were more responses of 2 (somewhat disagree) after the testimonials than before. This might suggest that there was some degree of shift in student opinion regarding the effectiveness of school learning at guiding career decisions, albeit one that was not statistically significant.

Attitudes Towards Testimonials

To inform future studies that might adopt a similar model to the testimonials described here, we surveyed student opinion of the testimonials after the testimonial sessions were complete (Fig 5). The majority of students felt that the testimonials fulfilled their expectations (Fig. 5A) and nearly every student felt that they had learned something new from the testimonials (Fig. 5B). However, students did not feel strongly that the testimonials helped clarify their career choice (Fig. 5C). The majority of students were neutral on the issue and about as many students felt that the testimonial sessions did not help as those who felt that it did help.

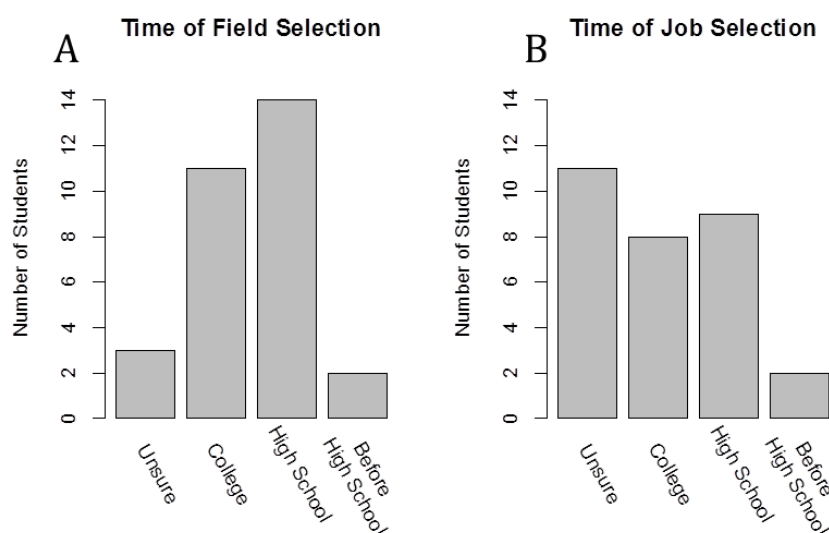In determining student attitudes towards the tes-

Figure 3: Student reported time of selection of both field (A) and job type (B). Students were given four distinct categories (still unsure, during college, during high school and before high school). These categories were also assigned to quantitative values (1-4) with increasing value indicating increasing time since the decision was made.

timonial sessions we found that quantitative survey questions could only provide a limited amount of insight into student opinions. To gain a deeper understanding of how students viewed the testimonials and whether they believed the testimonials to have been worthwhile, we asked several qualitative questions in the surveys in addition to the quantitative ones. A systematic analysis of the student responses revealed several recurrent themes. Phrases such as "choosing a career", "learn about different career paths", and "how students our age can decide [on a career]", suggested that students emphasized the value of the testimonials in affecting their career selection. This idea appeared both in responses suggesting that the testimonials were helpful in this regard as well as in responses indicating that the testimonials did not provide sufficient help in this area. For example, one student wrote: "[The testimonials] helped me think about potential careers." Another student clearly felt that the testimonials could have done more to aid in his/her career choice, writing: "I was really hoping more personal advice on choosing a career would be given." The frequency with which this theme appeared suggests that students considered the testimonials an opportunity to help them decide on a future career path, not just an interesting guest lecture. In fact, one student specifically mentioned this distinction stating, "I thought it would be just another guest speaker experience... [b]ut it actually exceeded my expectations."

Another commonly occurring theme was recognition of the diversity of fields existing in BME. Stu-

dents noted the "different aspects of BME", the "many career paths" available or "different areas in the medical world". This recognition suggested that by the end of the testimonials, the students were aware of the many options available in the BME field regardless of whether the testimonials had actually narrowed their decision.

Finally, students seemed to be particularly interested in the personal aspects of BME careers rather than technical knowledge. More than twice as many students referred to a desire to learn about "day-to-day" aspects of the speakers' careers, as well as their "life experiences" and how they made the decision to pursue their current career compared with students who referenced the technical information in the testimonials.

Focus Group

The poor attendance of the focus group (two out of thirty students) makes it difficult to draw any conclusions from the focus group alone. However, it should be noted that many of the ideas reflected in the surveys were reiterated in the focus group. The students who attended the focus group suggested that in future years, additional emphasis should be placed on industry, and how professionals identified their career paths. In addition they expressed an interest in hearing more about the "daily life" of professionals. When asked to identify the best part of the testimonials, the most common response was the question and answer sessions, an idea that was also discussed in the focus group. The overall response to the testimonials was positive with one student expressing that "I feel like
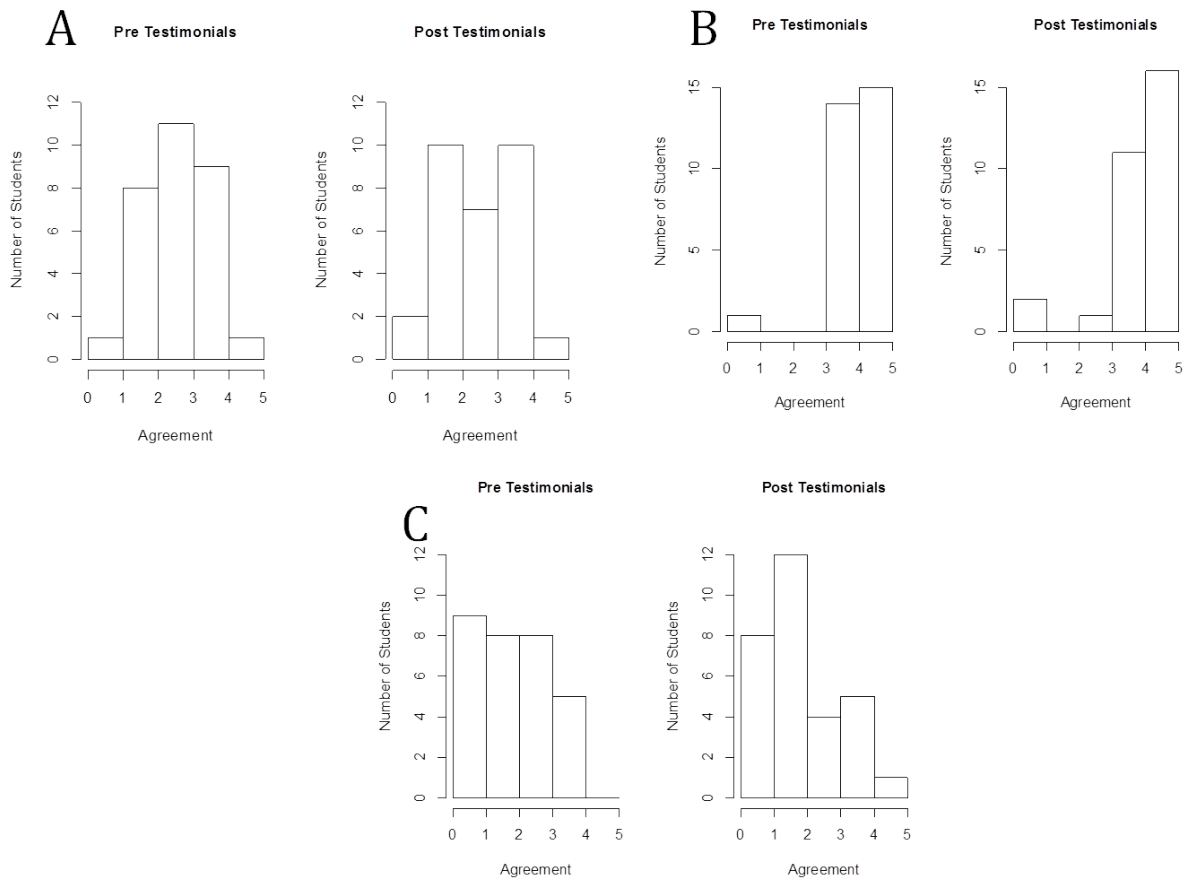
149

Figure 4: Student attitudes towards career selection. Students were asked to rate their agreement with three different statements on a scale from 1 to 5 with 1 being "completely disagree" and 5 being "completely agree". The statements were: (A) "I feel well prepared to select a career.", (B) "Practical (i.e. hands on or direct engagement with professionals) experience is necessary to choose an appropriate career.", and (C) "I believe that learning about a field in school is sufficient to select a career."

this is something that should happen in every major".

Discussion

In this study we developed and tested a new approach to provide undergraduate BME students with "practical" knowledge of the different career options in BME within the confines of a classroom environment. We hoped not only to present our approach using professional testimonials but also to inspire future studies on methods for informing undergraduate career selection in BME.

The most striking effect the professional testimonials might have had would be to directly influence or alter student career decisions. In our study we found that 15% of students did indeed change their career selection following the testimonial sessions. While our study design cannot support the conclusion that this change was definitively caused by the professional testimonial experience (as opposed to some other unmeasured factor), these data are suggestive that this type of testimonial can indeed influence student career choices. Perhaps more interesting are the specific choices of field and job type that students made. Unsurprisingly, the majority of students selected either biomedical engineering or medicine as their field of interest both before and after the testimonials. However, the number of students indicating an interest in biomedical engineering increased slightly. While this change is not statistically significant, student responses to the more open-ended survey questions suggest that the testimonials provided a new perspective on what a career in biomedical engineering is like and what it involves. Several students commented that the collaboration between BME researchers and patients was particularly intriguing and made the work that BMEs do seem more relevant and important.

The specific choices that students made regarding job type are as informative as their choices of field. The most commonly chosen job type was research and development in industry. Furthermore, two additional students selected other industry job types (manufacturing & test engineer as well as other industry position). However, there were no profession-
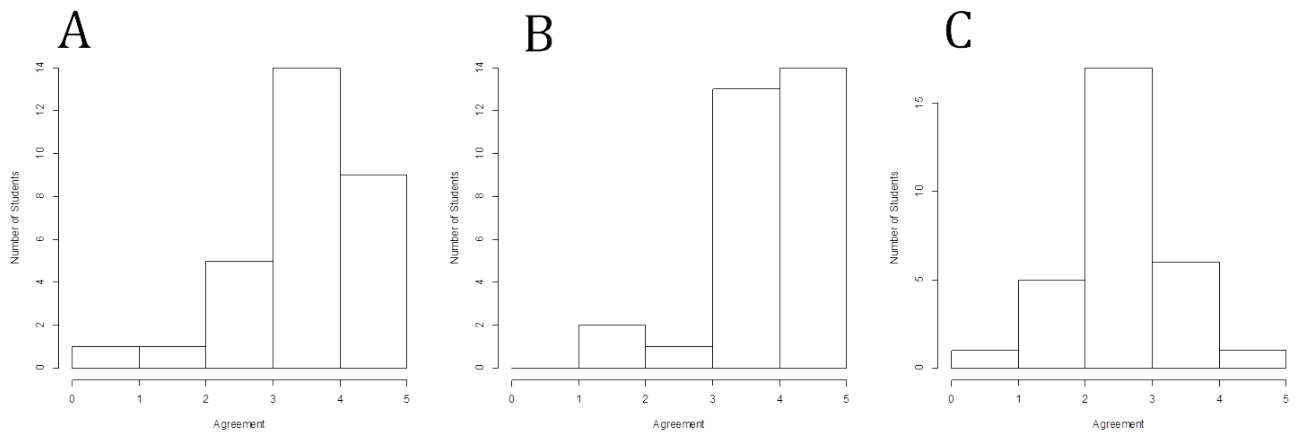
Figure 5: Student attitudes towards the testimonials. Students were asked to rate their agreement with three different statements on a scale from 1 to 5 with 1 being "completely disagree" and 5 being "completely agree". The statements were (A) "The testimonials fulfilled my expectations.", (B) "I learned something new from the testimonials", and (C) "The testimonials helped me clarify my career choice."

als from industry who participated in the testimonials. Despite this, the number of students interested in R&D in industry actually increased following the testimonial sessions. One may speculate that the increase may be due to either an increase in interest in the research aspect of R&D, which was emphasized in the testimonials, or a decreasing interest in the other professions that were specifically addressed during the testimonials, or some combination of the two. Regardless of the reasons behind it, the particularly high interest in industry careers suggests that future studies should place emphasis on industry careers. Alternatively the students could be surveyed ahead of time to identify the areas of interest to the specific class of students.

Unsurprisingly, a large proportion of students identified medical professional as their job type of interest. It should be noted that there was one student who selected medical professional as job type of interest, but selected BME as field of interest rather than medicine. Consequently, the number of students interested in medicine is exceeded by the number of students interesting in some sort of medical profession by one. It is not clear whether this discrepancy is due to a misreading of the survey by the student (e.g. assuming that field was equivalent to college major) or the student is interested in working as a BME in the medical space (rather than as a doctor or nurse).

The data on student certainty and time of career selection combined illustrated the importance of distinguishing field from job type. Student certainty was significantly greater for field choice than job type choice and students selected their field of interest sig-

nificantly earlier than they did their job type. We speculate that this may be in part due to the categorization of higher education into majors that are each associated with a particular field. Since students come into college already thinking about different possible major choices, it is more natural for them to select a field of interest, than a particular job type. Especially in a field such as BME in which the number of different job types is particularly high, this can pose an issue for college graduates. Not only are graduates faced with a difficult decision upon leaving university, but also different career options might require specific skill sets that are not necessarily integrated into the curriculum (e.g. undergraduate research builds critical skills for a research-related career). Without prior knowledge of the job types available within a given field, students may not seek the right opportunities during their undergraduate experience for the job type they select upon graduation. In such cases, students may be forced to pursue additional years of training before continuing along their chosen career path. By informing students about the available job types as well as possible fields for future careers early in their college experience, the students gain valuable time to pursue the right opportunities for their personal career paths during their undergraduate education.

The results of our study support the idea that the students themselves recognize that there is a lack of information provided to undergraduates regarding various possible career paths. While students did not feel strongly that they are either well- or-ill prepared to select a career, they felt that practical experience was an important part of making such a decision.

Furthermore many students did not feel that such practical experience could be easily attained in an academic setting. Taken together these results emphasize the need for providing students with practical engagement with professionals rather than information transmitted in a purely academic setting.

The data generated in this study are not sufficient to conclude that the testimonial sessions alone significantly improved the students' ability to make an informed career decision. However, it is clear that the testimonials were viewed as an opportunity to learn about different career options in BME rather than a lecture about BME research. Students engaged in critical thinking about what different career paths would be actually entail, expressing significant interest in the daily activities involved in different career paths. Given that the testimonials centered on lectures regarding specific technical topics in BME (solid organ transplantation and cancer), the students' perception of the testimonials as an opportunity for gaining information about various career paths is both surprising and encouraging.

Conclusions and Perspectives

As with any small-scale study it is difficult to conclude whether the testimonial sessions are likely to have a lasting impact on the participating students. However, the testimonials did have some degree of influence over student career decisions in the short term. Furthermore, these testimonials are only one example of method for informing undergraduate career selection. We envision that future approaches could include field trips to industry, medical, or research facilities, Skype sessions with professionals unable to travel to speak to the students in person, as well as a wider array of expert testimonial sessions in the classroom. One of the key benefits that these testimonials sought to provide was career advice within the confines of a typical introductory college class. While there are many approaches that might provide a greater insight into various careers such as internships or co-op programs, the logistic challenges involved in implementing such programs reduces the number of students that can benefit. The lessons learned from our study will hopefully inspire and inform future studies that can have a broader impact, reaching a wider population of undergraduates.

References

Akcay, H., Yager, R.E., Iskander, S.M., & Turgut, H. (2010). Change in student beliefs about attitudes toward science in grades 6-9. Paper presented at the Asia-Pacific Forum on Science Learning & Teaching.

Blair, B., & Millea, M. (2004). Quantifying the benefits of cooperative education. Journal of Cooperative Education, 38(1), 67-72.

Bransford, J.D., Brown, A., & Cocking, R. (1999). How people learn: Mind, brain, experience, and school. Washington, DC: National Research Council.

Cox, M.F.C., McNeil, N., Cekic, O., & Frye, M. (2011). Assessing the Pedagogical Impact of the VaNTH Engineering Research Center on Faculty and Postdoctoral Professionals. International Journal for the Scholarship of Teaching and Learning. doi:10.20429/ijsotl.2011.050219

Davies, P.F., & Litt, M. (2006). Interdisciplinary BME education: a clinical preceptorship course for undergraduate bioengineering students. Ann Biomed Eng, 34(2), 276-281. doi:10.1007/s10439-005-9028-x

Dressler, S., & Keeling, A. (2004). Student benefits of cooperative education. International Handbook for Cooperative Education, 217-236.

Godwin, A.S., & Daler, P.M. (2016). Disciplinary Differences in Out-of-School High School Science Experiences and Influence on Students' Engineering Choices. J-PEER, 6(2). doi:10.7771/2157-9288.1131

Imperial College London Department of Bioengineering. (2017). Retrieved from https://www.imperial.ac.uk/study/ug/courses/bioengineering-department/

Krishnan, S. (2010). Collaboration for cooperative work experience programs in biomedical engineering education. Conf Proc IEEE Eng Med Biol Soc, 2010, 2955-2958. doi:10.1109/IEMBS.2010.5626260

Krishnan, S.M. (2015). Embedding Internship Programs to Augment BME Education. In I. Lackovic & D. Vasic (Eds.), 6th European Conference of the International Federation for Medical and Biological Engineering (Vol. 45, pp. 761-764). Berlin: Springer-Verlag Berlin.

Krishnan, S.M., & Asee. (2014). Development of Experiential Learning Models in Biomedical Engineering Programs for International Implementation. 2014 Asee International Forum, 10.

Linsenmeier, R A., Harris, T.R., & Olds, S.A. (2002). The VaNTH bioengineering curriculum project. Second Joint Embs-Bmes Conference 2002,

Vols 1-3, Conference Proceedings, 2644-2645. doi:Doi 10.1109/Iembs.2002.1053470

Lisenmeier, R.A. (2003). What Makes A Biomedical Engineer? Defining the Undergradaute Biomedical Engineering Curriculum. IEEE Engineering In Medicine and Biology Magazine.

Martin, T., Rivale, S.D., & Diller, K.R. (2007). Comparison of student learning in challenge-based and traditional instruction in biomedical engineering. Annals of Biomedical Engineering, 35(8), 1312-1323.

Sadler, P.M., Sonnert, G., Hazari, Z., & Tai, R. (2012). Stability and Volatility of STEM Career Interest in High School: A Gender Study. Science Education, 96(3), 411-427. doi:10.1002/sce.21007

Savelsbergh, E.R., Prins, G.T., Rietbergen, C., Fechner, S., Vaessen, B.E., Draijer, J.M., & Bakker, A. (2016). Effects of innovative science and mathematics teaching on student attitudes and achievement: A meta-analytic study. Educational Research Review, 19, 158-172. doi:10.1016/j.edurev.2016.07.003

Science Education Glossary of Terms. (2017). Retrieved from `https://www.nibib.nih.gov/science-education/glossary-g-42691`

Wright, A.S.W., Frye, C.A., Mathur, A.B., & Patrick Jr., C.W. (2007). A Ten-Year Assessment of a Biomedical Engineering Summer Research Internship Within A Comprehensive Cancer Center. Journal of STEM Education, 8(3), 28-39.

Voluntary Assessment in a Calculus Classroom: Student and Instructor Perspectives

Kristen Pueschel

Department of Mathematics, Cornell University
kpueschel@math.cornell.edu

Abstract

In this paper we use student surveys and data on student computer use to probe how students engaged with a low-stakes but mandatory assessment. We use instructor interviews and data on instructor computer use to probe how instructors engaged with the same low-stakes assessments. We find that student participation rates are steady and high. Although a segment of the student population spent very little time on the assessments, there were still many students spending at least a moderate amount of time on them. Student surveys showed that many students reacted to getting incorrect answers. For example, they reported reading the responses that appeared when they chose an incorrect answer. Instructor use of these assessments was very limited, with only a few instructors looking at even 10% of the assignments during the semester. We found that most instructors did not think of these assessments as part of their classes and most reported that they did not incorporate the assessments into their class in any way. We conclude that although students tend to participate and engage with assessments if they are graded and will assess their understanding if it is convenient to do so, instructor buy-in to pedagogical features of coordinated classes is not automatic.

Introduction

In a 17 section calculus course, with a total of 11 instructors, students were given participation-only quizzes online. In our department, large courses of this nature have a coordinator each term that sets some of the policies of the class. The decisions they make include producing common homework assignments, common exams (with other instructors offering input), a common schedule that is used both by instructors and students, and common online quizzes, graded only for participation, that were due twice a week throughout the semester. These quizzes covered material before it was scheduled to be presented in class. In this study, we explore two main topics:

1. In low-stakes assessments where engagement is voluntary, how do students behave? What is their use like? What motivates them? How do they use these low-stakes assessments, if at all?

2. How do instructors engage with low-stakes assessments, if at all, when they have control over what happens in their classroom? How do instructors incorporate these assessments into the classroom, if at all?

To better understand how the online systems were used by students and instructors we examined the logs of the online system which recorded information such as time spent on each quiz and number of attempts. This metadata provided a supplement to traditional surveys of students and interviews of instructors.

I came to this research as a graduate student instructor in this course. I took particular interest in the participation-only online quizzes, as I was interested in encouraging students to read before class. My goals were two-fold: by encouraging reading before class, I hoped that students would not be seeing (or recalling) the material for the first time in the classroom. My classes incorporated frequent 'check points' for partnered discussion, so it was important to me that my students were equipped to think about and discuss the material in class. My second goal was to foster transferable skills such as study skills and technical reading skills. The quizzes seemed aligned with my reading goals, but casual feedback from students made me wonder how this pedagogical tool actually worked in practice such as in dorm rooms, libraries, and cafeterias. How were students doing these quizzes? Did they find them useful, or understand how they fit into the course? My impression that my students did not find the quizzes to be very useful made me wonder how other instructors were incorporating this component into their classrooms. Did my students seem to feel they were unhelpful because I was not using them effectively? Did other instructors use the quizzes to further different goals?

During the course of this study, themes of balance arose from interviews with instructors. Questions that we did not explicitly examine, but were important themes in instructor responses to questions about the use of informal quizzes in their classes included the following:

- How do instructors negotiate their power in the classroom with the restrictions of common course elements?
- How do instructors handle pedagogical decision-making in a large group?

From instructor interviews, we are able to propose answers to these questions from graduate student and postdoctoral instructors' perspectives with respect to this particular course element.

Literature Review

This study was designed around understanding how students interact with low-stakes pre-testing, with particular interest in how this is related to reading compliance. We begin by discussing reading compliance, the factors that effect it, and common strategies to increase it.

Studies of reading compliance rates have shown that although students tend to complete most of their reading before being examined on the material, much of the reading is not done before the class in which it is taught. In Clump, Bauer, and Bradley (2004), 423 students in an introductory psychology class filled out anonymous surveys reporting their reading compliance. Students reported completing, on average, only about 27.5% of the assigned reading before class, and about 70% of assigned reading was completed before exams related to those readings. This behavior is seen to persist into graduate school, with reported reading rates of 54% before class in a masters level psychology course (Clump & Doll, 2007).

Much has been written on the reasons that students do not read before class and ways to increase compliance. For example, Lei, Bartlett, Gorney, and Heirschbach (2007) posit that students do not complete reading assignments because they don't have the reading comprehension skills to do so, a lack of self-confidence, disinterest in the course, and underestimation of reading importance, and suggest implementation of pop quizzes to motivate students to read. When students must decide how to spend their time, they may not choose reading before (or after) class. Students may have jobs, social obligations, and other classes to balance. To increase compliance, suggestions include surprise reading quizzes (with proper warning in the syllabus) (Sappington, Kinsey & Musayac, 2002), scheduled reading quizzes, reading guides that are due before class (Stickles & Stickles, 2010), and reading assignments in which students must submit questions on the reading before class (Henderson & Rosenthal, 2006)

Hoeft (2012) conducted a study in which students described themselves as reading compliant or reading non-compliant. The three most common suggestions from non-compliant students for increasing their likelihood of reading were implemented a few years later in three courses. Student suggestions were: frequent reminders of reading assignments, in-class quizzing on the reading, and guiding questions which were implemented in student journals that were graded and commented on weekly. Each class used only one of these strategies, and the students were surveyed twice in each class. Hoeft found that reported reading compliance was highest in the journaling class, with 94% reported compliance, followed by quizzing with 74% reported compliance, and extra reminders with 46% reported compliance. The compliance level when the course was taught with no particular motivators was similar to that when extra reminders were given, and significantly less than in classes implementing journaling or quizzes. Interestingly, Hoeft found in the first study that basic comprehension of the reading was demonstrated by only about half of the students reporting reading compliance, suggesting that they may have not actually completed the reading. Sappington, Kinsey and Musayac (2002), found similar mismatches of student reporting of reading and actual reading compliance. In their study, students were directed to read the syllabus, and then were polled the next day about whether they had read it. While 93% of students reported reading the syllabus, only 22% followed the directions in the syllabus that would earn the student a bonus point, namely, emailing the instructor. They conclude that students do not accurately report their own reading. Hoeft's result is startling because students were polled anonymously. In our study, students were polled anonymously after the course was over, so there is no possibility of punitive action based on their answer, but her study suggests that students may still not provide accurate data in this situation.

Discussion centered on the teacher's use of the readings in class is mostly absent in the papers above; they instead focus on external motivators that can be used to get students reading. Indeed, most of these studies assume that students lack internal motivation to read, or are un-prepared to handle reading. Instead of trying to address this lack of internal motivation, focus is placed on how well external motivators such as reading quizzes encourage students to read. In a case study, Brost and Bradley (2006) provide additional explanations for student non-compliance that assign some responsibility to instructors. During a course composed entirely of guest-speakers the authors considered the great variety of reading lengths and difficulties assigned by the speakers and reported their observations on the use, in class, of these readings. They found that the idea of an appropriate

reading assignment varied greatly across the lecturers, from assignments of a few pages from a textbook chapter to numerous research papers. Moreover, lecturers ranged from reviewing the readings step-by-step, to building discussion from them, to lectures with no apparent connection to the readings. Brost and Bradley propose that teachers also play a central role in reading non-compliance. Interestingly, the role of the teacher is implicit in many informal articles on increasing reading compliance in introductory mathematics and physics courses. In these articles, instructors commonly report how readings are explicitly addressed during class time, for example in Boelkins and Ratliff (2000) and Amick (n.d), including the procedure for their discussion.

In this study, we consider the implementation of one reading compliance strategy. We look at student use of reading quizzes, student reports on reading behavior, and the way that instructors adopted and interacted with this course component.

As practitioners, we know that grades can be a powerful motivator - ungraded assignments often fall by the wayside. Completion-only grades help to motivate students to complete the assignments, but it is not clear whether or not they are sufficiently motivating to get students to engage with the assignments. Especially in the case of questions about material not yet covered in class, completion-only grades may be the only practical solution, as students may feel discouraged by earning low grades on new material. The literature on optional and completion-only grades is small. Zywno (2007) reported on one attempt to implement optional activities in the setting of a large engineering class. Students were rewarded for participation in optional activities throughout the semester, which included in-class group work, use of online discussion boards, exit surveys, and paired quizzes. The variety of activities was intended to appeal to students with a variety of different learning styles, and the activities coordinated with an active-learning style class. Zywno found that participation in activities ranged from 40% to 70% on average. The median student completed about 40% of optional activities. She found that student participation in optional activities correlated with performance on exams. Her study does not include student reactions to the optional activities, though engagement might be measured by popularity – students might only complete activities that they found engaging. In our study, we get student feedback on the low-stakes assignments, and also look at response times to try to understand whether students were rushing through the assignments or really trying.

Our final interest in this paper is instructor buy-in for course components in large coordinated courses. Kobler (1978) asserts that before there is any hope of national standards for minimal performance in courses being implemented in a meaningful way, instructor teams need to be able to reach consensus about the goals and expectations for their students. He explains that group discussion and consensus is more desirable and easier to evaluate than handed-down guidelines that may not be designed with the institution's students in mind. In an apparent response to Kobler's article, Iyasere (1984) argues for the importance of consensus among instructors on common goals in the structuring and grading of multi-section composition courses. She expands upon the reasons that instructors might oppose efforts to establish common standards and course components in a classroom. The most important of these was the perceived loss of academic freedom – the more commonality between classes, the less room instructors might feel free to teach things the way they like or to structure the course how they think is appropriate. Iyasere argues that through explicit discussion of course goals instructors do lose some freedom but gain greater collegiality, and that instructor effectiveness will increase as focus on the course goals will especially in courses with high turn-over of instructors, such as courses with graduate student TAs. She argues that making goals (learning objectives) for the course explicit is a natural extension of agreeing on a course description. Moreover, Iyasere says that by establishing common standards among instructors teaching the same class instructors get a voice in the process that says what is valued by the department, and thus are more likely to teach in a way consistent with departmental expectations, protecting them in evaluations of their teaching.

In Bullock, Callahan, and Shadle (2015) , mathematicians at Boise State University share how their school moved from many instructors responsible for a single section of calculus to a coordinated, coherent approach. They found that with voluntary buy-in of instructors to coordinating the course and implementing pedagogical reforms that they were able to change the culture of the course, thereby improving outcomes for students and increasing satisfaction of the instructors. They argue that voluntary, discussion-based coordination efforts provide an alternative to top-down efforts in which a single coordinator makes decisions about the course. They argue for the importance of instructor buy-in for the implementation of changes in the classroom. In this study, we examined an example of a common course element that was implemented in a top-down manner and the extent to which it was implemented in

Figure 1: An example Moodle-quiz question and response.

classrooms.

Methods

Data about the use of quizzes implemented through the Moodle open-source learning platform (hereafter 'Moodle-quizzes') were obtained from three sources: surveys from students, interviews with instructors, and data on Moodle usage stored by the system. The student survey was distributed around the same time that I began instructor interviews, and the usage data was analyzed afterwards. My own assumptions about how the Moodle-quizzes were intended to be used influenced the student survey– my questions focused heavily on reading as well as Moodle quiz use, and on their intersection. This is also seen in my language in the survey. Moodle-quizzes were referred to as reading quizzes throughout the surveys. Interviews with instructors showed that this focus was far from universal. It is therefore probable that for questions focusing on reading, it would have been more appropriate to get participants from my own two sections rather than drawing from the entire student body of the course.

Quizzes

In this section we describe the Moodle-quizzes in more detail. All students had access to a course management system, Moodle, an open source analogue of Blackboard. Announcements about the course, practice exams, and quizzes were posted on Moodle. In a typical week, students would have to complete two Moodle-quizzes. These were open for several days, with explicit instructions that they were meant to be completed before class. Students had four to five days during which the quiz was available to be taken. Because of the way classes were scheduled, one quiz each week would close before class, and the other would close after class. Quizzes consisted of between two and nine multiple-choice questions. Grading on the quiz was strictly for completion. Students were allowed one quiz attempt. After completing the quiz, the correct answers were revealed. The course coordinator and his assistant frequently wrote an explanation either of the correct solution or a response to an incorrect selection, that appear if an incorrect answer was selected (Figure 1).

The material tested in the Moodle-quizzes was aligned with the assigned reading schedule. If their instructor was following the coordinated lecture schedule, a student would take the Moodle-quiz on content that had not yet been covered in class. In the end, the 22 quizzes made up 4% of the final grade.

Student Surveys

An approximately twenty question Qualtrics survey was developed to learn about student reading and use of the Moodle quizzes, motivation to complete these assignments, student understanding of the purposes of the quizzes, and their understanding of how the quizzes fit into the course. Because factors like previous exposure to calculus and success in those classes seemed like they could affect the amount of reading needed to complete the quizzes, students were asked about previous calculus coursework, including exam scores on standardized exams such as the International Baccalaureate and Advanced Placement exams. Further, since competing priorities are often a reported reason for student non-compliance with

readings, students were asked to rate the importance of this course compared to their other courses. Some short-answer, open response questions were included, but most questions had a list of prescribed answers to encourage student completion.

The survey was distributed by email halfway through the semester after they took the course. The course coordinator sent an email requesting student participation. Students were told that the survey was voluntary and anonymous and no IP addresses were collected. The course finished with approximately 340 students. We received 32 responses that were more than half complete and 30 'finished' surveys. Participants could progress through the survey without completing all questions. A 'finished' survey was one in which the final screen was reached, that is, the student registered that they were done with the survey. Not all finished surveys had all questions answered. All written responses were analyzed to identify common themes and specific comments of interest. Short response rates were significantly lower than multiple choice response rates, which out of the 30 finished surveys had at least 29 respondents for each question. In contrast, only 17 out of 32 respondents answered the open-response question "If you didn't always do the readings, could you please describe some of the reasons that you didn't always do them?" Due to the small number of survey participants, statistical analysis of the data was not pursued.

Of 29 students, 7 (24%), had taken no calculus previously, 13 (45%) had taken a lower-level high school calculus such as the AB AP Calculus or the standard level IB Mathematics, while 7 (24%) had taken a higher-level high school offering of calculus, such as the BC AP Calculus or higher level IB Mathematics. The two remaining students had taken calculus at the college level previously. An approximately equal number of students reported calculus to be a low, medium, or high priority course when they took it (8/12/10). No demographic data was gathered from respondents.

The surveys provided student accounts (picked out of prescribed choices) of their Moodle quiz usage, including how many of the quizzes they did, what resources they consulted when completing the quizzes, and what they did when they did not get the correct answer. This provides some narrative for the use of these low-stakes assessments.

Instructor Interviews

Interviews with instructors aimed to uncover instructor expectations for their students for reading, the ways that they incorporate this course component into their classrooms, and their conception of the goals of Moodle quizzes. Of the eleven instructors for the course, seven agreed to be interviewed. One lecturer, two post-doctoral researchers, four graduate students, and the first author of this study provided responses. The experience levels of the instructors varied. Two of the interviewed graduate students were relatively inexperienced as instructors, while the remaining three were senior graduate students with experience with this course. Further, the head instructor sent an explanation of the goals that led him to adopt Moodle quizzes for the semester. Interviews were voice-recorded and transcribed and recurring themes were coded.

Moodle Usage Data

Moodle stores activity logs for the system with a Moodle-assigned user id, description of activity, and date and time-stamp. Names and IP addresses were first deleted from these logs. Multiple data were available, including for individual quizzes, the scores and time to completion for each student, and the number of log-ins and activities completed on the website, such as downloading a PDF, reviewing a quiz, etc. Text-processing software was used to get data into useful forms. This database allowed us to view the website usage for all members of the course. Demographic data about the student body for the course were not gathered, but the previous exposure to calculus seen in the survey data agrees with the first author's experience in the classroom - usually all but a few students come into the course with previous exposure to calculus.

The data from the Moodle website provide insight into student use - for example, time taken to complete quizzes and scores. This information is not perfect - a student that starts a quiz and then checks Facebook for 10 minutes before randomly choosing answers counts in the same way as a student that spends that 10 minutes working on the quiz, and certainly, quick question answering does not necessarily amount to guessing. Students that are adept with the material may be able to complete the quiz quickly. Still, this data source helps us to triangulate student usage, and supports or provides contrast to student and instructor reported usage.

The Moodle data gave us a complete snapshot of instructor use of the system. Because there were relatively few instructors, it was practical to look more closely at instructor use of the system.

Results

Students

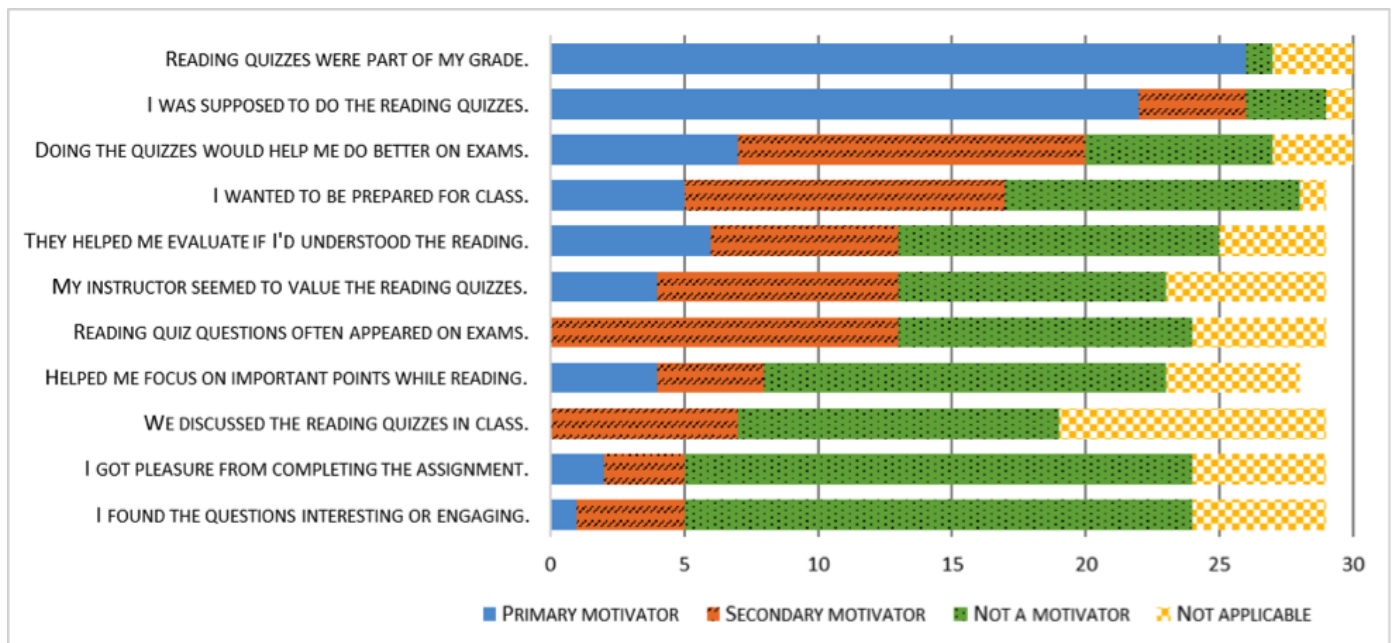Students were polled about factors that motivated them to complete the Moodle quizzes (Figure 2).

Figure 2: Motivators for completing reading quizzes. Bars are ranked by the number of respondents ranking the motivator as at least a secondary motivator.
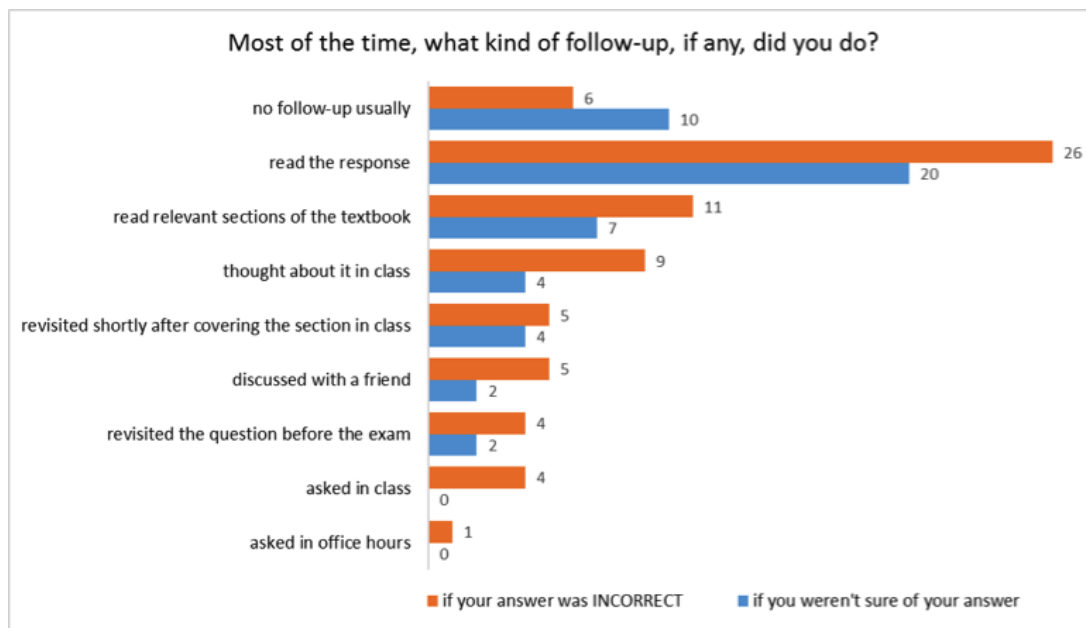

Figure 3: Student responses (n=30) to incorrect answers and uncertainty. Students were allowed to select multiple responses.
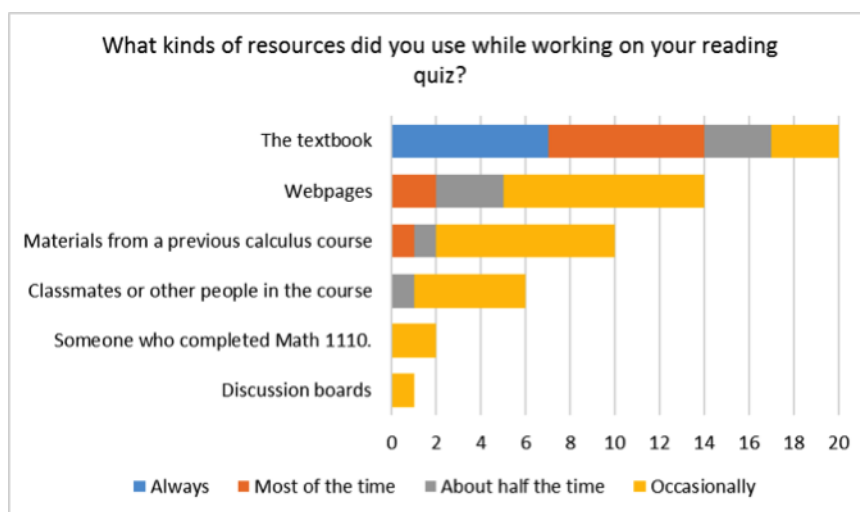
Figure 4: Resources used (n=22) during quiz taking. Eight participants reported never using any resources.
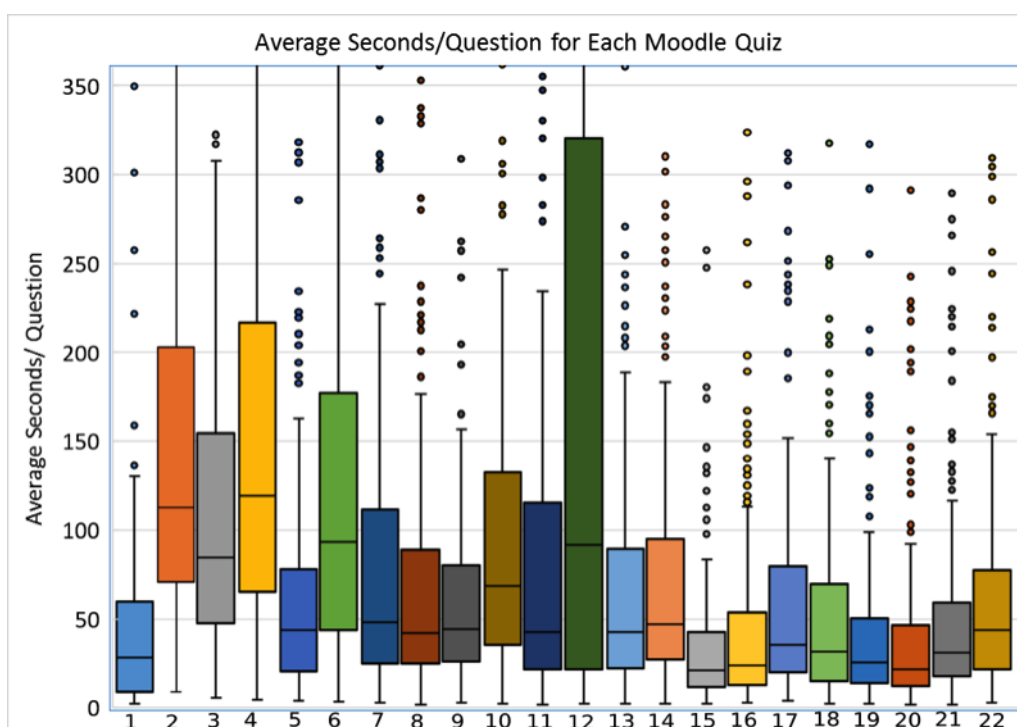

Figure 5: Average seconds/question for each Moodle quiz.

They were able to select as many primary and secondary motivators as they liked. The overwhelming majority of respondents were motivated primarily by the grade and by a feeling of obligation. Notably, almost half of the respondents claimed that they were at least somewhat motivated to do the Moodle quizzes to evaluate their understanding of the readings. In the online survey, no guidance was given to respondents about when to choose that a response was not applicable (NA). The relative popularity of the not applicable option is somewhat puzzling. It was included to allow students to differentiate between potential motivators not being motivating and not being available, in particular for class/instructor motivator. We see that a third of participants reported not discussing the reading quizzes in class.

For the respondent students, it appears that the grade is a strong motivator, even though in total Moodle-quizzes made up only 4% of their final grade. Of the 29 student participants, 23 reported taking at least 75% of the quizzes. In the Moodle system, this completion rate was confirmed - there was attrition from the course, which complicates the measures of participation. However, all quizzes had more than 215 participants, just under 65% of the final number of students, with a median participation of 273 students, or 81% of the final number of students in the course. In the motivation data, I was surprised by how many respondents took cues from their instructors, saying that they were somewhat motivated by the instructor appearing to value this course component.

If students were using the Moodle quizzes as a tool for assessing their understanding we would expect to see reports of reaction when students got a question incorrect or if they were unsure of their answer. Survey participants reported taking action in the case of incorrect answers. An incorrect answer seemed to prompt more action than a correct answer which students were not sure of(Figure 3).

Several students said that they did no follow-up, but also said that they read the response. This might be an artifact of the mobile version of the survey, in which follow-up possibilities were presented one-by-one. The first follow-up option was "no follow-up". This discrepancy was only seen in the first question, in which students would not know what options were available to them.

To better understand how students worked on the Moodle-quizzes, and also to get a measure for their investment in this low-stakes assessment, participants were asked whether they had ever used a resource while working on one of the quizzes. A substantial number of respondents (8 of 30) said that they had not once used a resource. Students that said they had used a resource at least once were presented with a list of potential resources and asked about their frequency of use (Figure 4).

Student use of the textbook was echoed in their response to questions about reading and the Moodle-quizzes. Students were given the prompt "Most of the time I did the reading quizzes..." and were asked to choose all appropriate responses. Students reported high-rates of textbook use while doing their quizzes (Table 1). From this data it appears that the reading quizzes fit into a culture of reading before class.

The Moodle system recorded both how well students did on quizzes and how long it took students to complete them. Because the number of questions per reading quiz was variable we normalized the time taken on the quizzes by taking the average time in seconds to complete one question for each quiz. A box-plot shows the median time to complete the quiz, with the whiskers denoting the first and third quartile for time of quiz submission (Figure 5). The data for time spent on quizzes is simplified in that it marks the time from when the student first opened the quiz to when they submitted it. This may or may not contribute to some of the longer times. For example, some students had the quizzes open for 3 days, but it is impossible to tell how much time they actually spent working on it.

We see that the median time per question hugs closer to the first quartile than to the third quartile in every quiz (Figure 5). If we further collapse these data, by examining the median of the first, second, and third quartiles for average seconds per question across the Moodle-quizzes, we find that the median of the third quartiles across all quizzes was 1 minute 25 seconds, the median of the second quartile across all quizzes was 43.5 seconds and the median of the first quartiles across all quizzes was 22 seconds. This implies that for half of the quizzes, at least 75% of students spent less than a minute and a half on each question. Although students could reasonably read a question and ponder the answer in a minute and a half, these students were probably not reading the book while they worked on their quizzes. This contrasts to student reported reading and quiz-taking habits, in which 11 of 32 respondents reported that most of the time they did some reading activity such as skimming subsections, reading subsections, or reading the entire section. The mean quiz scores range from a high of 7.9 for Quiz 1 to a low of 3.4 for Quiz 15 (Figure 7). It appears that Moodle-quizzes were a challenge to the students. We did not find significant correlation between student success and

Table 1: Reported reading and quiz behavior.

Reading activities done before or while taking quizzes. Multiple responses allowed.

no reading	8
skimming	13
reading relevant subsections	11
reading entire section	14

Most of the time I did the reading quizzes…

without looking at the textbook		8
while doing a reading activity		11
after doing a reading activity		7
BOTH while and after		6
total respondents		32

the median time to complete a question.

When students were asked how the readings or reading quizzes could be improved for future students, we received a wide variety of suggestions. An important theme that emerged was a request for better pay-offs, both with the class and with the exams. Students suggested "incorporat[ing] them into lessons" and "discuss[ing them] in class". One student wrote "The problems those questions addressed was [sic] never really brought up again in classes or tests." Another theme was a suggestion to use Moodle-quizzes not just for evaluating whether or not they had understood the readings, but whether they understood the new material. For example, one respondent wrote "…having the moodle quizzes align with the lectures. I did not like having to take the quiz just based off of readings and not lectures because we have not gotten there yet" and another student suggested "…have people take them after class to see what they really learned". Some suggestions reinforced what was helpful in the reading questions. For example, one participant suggested more detailed responses to incorrect answers.

General feelings towards the quizzes were mixed, with several notable themes. Some respondents (4 of 20) were either confused about how the quizzes were graded or perhaps understood the grading system but found low 'scores' on the quizzes demoralizing. For example, one respondent wrote "they were easy points if you could efficiently learn from the reading. I am not one of those individuals and therefore it was harder for me to do well on the quizzes." Five other respondents mentioned that the quizzes were graded only for completion. Some students expressed tolerance, based on the low-stakes nature: "They were hard but I was appreciative of the fact that only effort counted" while others seemed to prefer higher stakes: "They weren't helpful to me. I just rushed through them because it was a participation grade." Many positive comments were non-specific, such as "good". A recurring complaint was the fre-

quency of the quizzes and difficulty remembering to do them. For example: "it's really easy to forget about them when you have a lot going on" and "had to set reminders to do them".

As a way to triangulate instructor use of the Moodle-quizzes in the classroom we asked survey respondents to describe ways that their instructor used the reading quizzes in class. The question was short answer in case participants had a range of experiences. Disappointingly few respondents answered this question (n=11), but students either said that their instructor did not use Moodle-quizzes in class, or that they could not remember if they had ever been used, or that the instructor answered questions about them if asked. One respondent wrote that it was sometimes used as a springboard for discussion: "sometimes if [the instructor] thought that a question was particularly difficult or noteworthy, [they] went over it in class." Approximately half of the 30 respondents said that their instructors had explained the purpose of Moodle quizzes in class. Those that did not receive explanations were luke-warm about the instructor explanation as motivating: "I did the quizzes because they counted as points so I would have done them with the same amount of effort regardless of an explanation." Those that received explanations had a wide variety of reported instructor explanations. The most popular explanations had to do with preparation, either for class or for the exam, although there were several students who mentioned self-assessment such as "to judge your own comprehension of the reading".

Instructors

The instructor interviews revealed a wide-range of levels of engagement with the Moodle-quizzes. From our interviews, it appeared that instructors had one of three general views of the quizzes: 1) Quizzes as a required part of the student experience, and thus a necessary part of class, 2) Quizzes as a strictly student experience, and not a part of class, but reading quizzes brought some value or served some purpose,
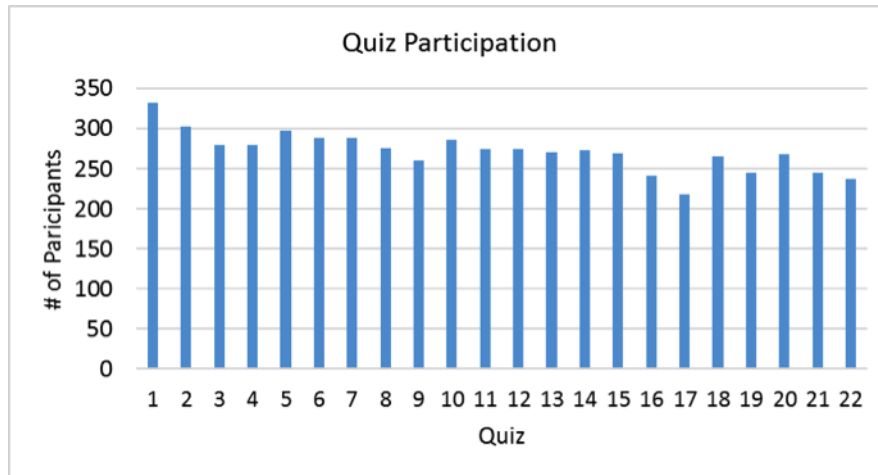
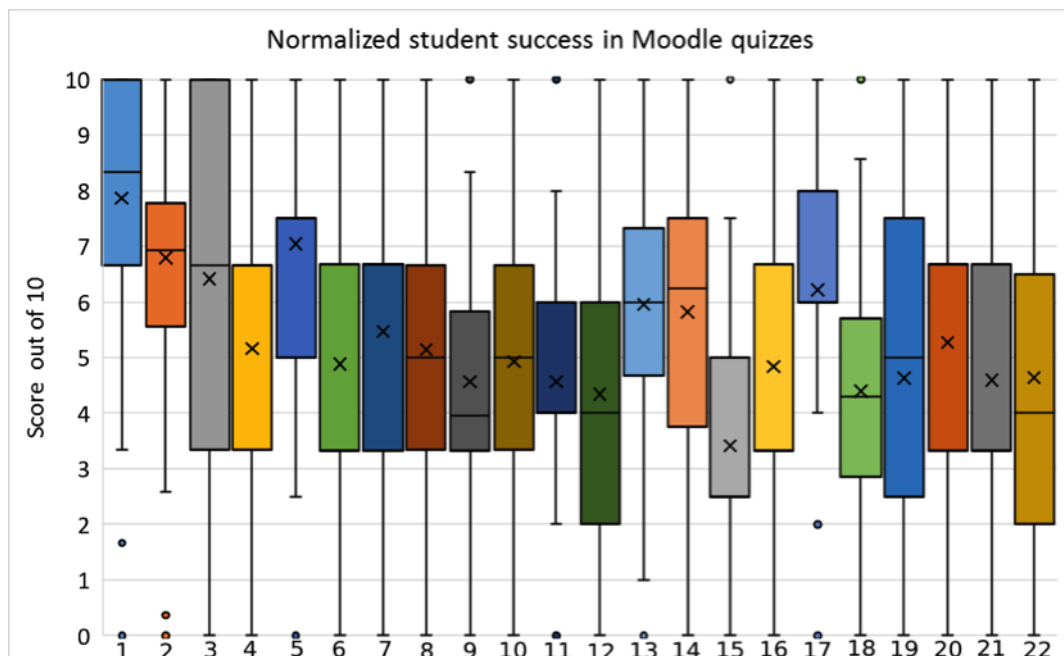Figure 6: Quiz Participation. Number of participants ranged from 218 to 332.

Figure 7: Scores for each Moodle-quiz. The x's denote the mean score.

and 3) Non-engagement, with no instructor buy-in to the value of the Moodle-quizzes.

Instructors in the first category expressed feelings of obligation to tie the quizzes into the classroom experience. For example, one instructor said "I felt like I had to talk about them at least in brief, because [the students] were supposed to do them before every class". In this view, because the students had to do the quizzes, it was up to instructors to make them a useful part of class. "Some of these questions... would end up being true or false for very trivial and vacuous reasons, and uh, without some discussion would be totally meaningless to the student." Confusing questions, or questions with low success rates, were natural starting points of class discussions. Instructors that used the system to look at student success rates on the quizzes reported that they focused on questions that students struggled with. This approach is similar to Just-in-Time teaching, in which an instructor modifies their lessons based on the apparent needs of students as assessed by pre-tests. An important distinction that differentiates this from Just-in-Time teaching is that the instructors do not have control over the content of the pre-tests. Thus for instance, the content of the Moodle-quizzes did not always align with the material, and questions might or might not be appropriate for building a class-period around.

Another kind of usage of the first type was using Moodle-quizzes as an occasional launching board for discussion in class, at the request of students. This puts the onus on students to ask for help when they struggled with a quiz. Instructors who had this kind of usage reported that students did not typically have questions: "the most common situation was that there were no particular questions on it and nothing that they were particularly excited about, nothing that I felt was important to make any kind of big element in a class, and so not too much happened." Instructor goals for their students were centered on encouraging student reading, and exposing students to interesting questions and examples. In contrast, instructors that actively started discussions based upon questions with low success rates reported that in class there were often many students who did not understand the solution or why incorrect answers were incorrect.

Instructors in the second category valued the quizzes in some way, but considered them to be a student-only experience. Some goals associated to Moodle-quizzes by various instructors included:

• Reading compliance: "I did try to design the course to be best-suited to students who had done the reading, especially knowing they had those Moodle-quizzes... [Moodle-quizzes] let me better assume that the students had done the reading."

• Exposure to varied (expert) questions: "[T]hey tended to be kind of pithy... I liked them, so I was glad the students were looking at them." and "It never would have occurred to me to [ask]... some of these things that they obviously found difficult..."

• Keeping pace with other sections: "Points that were coming up on those [Moodle-quizzes] were points that I should make sure my students were properly prepared for."

Several instructors reported having difficulty using the system and expressed interest in greater use if the system was easier to navigate: "I would have liked to be able to find an easy diagram of how the students were doing, and... modified my teaching based on that."

Instructors in the third category did not 'buy in' to the value of reading quizzes. They reported not paying attention to the questions appearing on the quizzes, or to student success or failure on them. Many reported reminding students to take the quizzes, but in all other ways viewed them as completely distinct from their class.

An emergent theme of these interviews was that many instructors considered the quizzes to be an imposed element of the course. For example, after one instructor expressed frustration at not being able to use the system in a way that he would have found useful he asked the interviewer "How are you expected to use them?" Other instructors asked where the questions came from and who wrote them. Finally, even instructors who incorporated the quizzes into class expressed skepticism about the goals and effectiveness of the Moodle-quizzes, which echoes imposition: "I feel like they [students] did no better than 50-50, so on that level, I don't know what the students were doing or getting out of this. Ostensibly it was a check, it was something that would force them to read something in the textbook, and then think about it and answer, but often it was away from what I was doing". The reaction of most instructors to this imposition appears to have been a lip-service alignment with this course element in reminding students to participate. Other instructors more firmly rejected this course component – one recalled "I told them something like 'I don't much care [how you do on the quizzes].'"

In analyzing the Moodle data for instructors, there were two users that were clearly the course coordinator and their assistant. Because they were responsible for developing the Moodle quizzes, their usage history did not seem relevant. From our interviews, the course coordinator included discussions of Moodle-

	In1	In2	In3	In4	In5	In6	In7	In8	In9	In10	# of Quizzes
Day1-Exam 1	1	1	1	5	3	4	11	7	7	17	8
Exam 1- Exam 2	2	2	2	6	9	9	6	10	14	11	8
Exam2-Final	1	3	3	3	5	5	5	8	6	7	6
Total # of Sessions	4	6	6	14	17	18	22	25	27	35	22

quizzes in class, but did not look at student success on quizzes. This left us to examine the usage history of 10 instructors. The Moodle system was a repository for private course information, including practice exams for the students and statistics from exams (mean and median scores, etc) and exam solutions. The system was also used by instructors to look at Moodle-quiz questions, to look at overviews of course-wide student success on the quizzes, and to look at the grades of individual students. The easiest data to retrieve was a simple count of the number of sessions (lumping all accesses in a day together) each instructor had per semester. We broke the semester up into 3 natural divisions: the first day to the day of the first exam, the day after the first exam to the day of the second exam, and the day after the second exam until the day of the final. We then tallied the number of sessions in each of these periods (Table 2).

We then broke down the actions of the instructors based upon the logs. Some functionality was easy to see in the logs, such as looking at practice exams, while other functionality was more difficult to determine, such as looking at announcements. Quizzes were counted based on number of quizzes accessed during this period. The categories of practice exams, exam solutions, and exam statistics were measured by the number of days in which one of these types of PDFs were accessed. Thus a day when five practice exams were accessed counts the same as a day when a single practice exam was accessed. From this data we see that most instructors used Moodle moderately, and primarily as a repository for course files, with Moodle-quiz related usage declining for all instructors except for Instructors 9 and 10. An oddity of the data is that the all accesses of Moodle-quizzes by Instructor 5 occurred on one day in the first period and on another day in the second period. Our best explanation is that Moodle-quizzes were suggested as a source of exam problems. Since exams were written jointly by instructors, this instructor might have been reviewing the quizzes to see if there were any questions they might like to have on the exam. This data agrees with our analysis of attitudes from the interviews: two instructors seemed to use the system frequently, looking both at Moodle-quiz problems and student success. A few instructors used the system initially, but use was sporadic and petered out after the second exam (Figure 8).

Discussion

In this study we aimed to understand how students and instructors interacted with informal readingquizzes given before classes where course material was first addressed. We wished to answer the following questions:

1. In low-stakes assessments where engagement is voluntary, how do students behave? What is their use like? What motivates them? How do they use these low-stakes assessments, if at all?

2. How do instructors engage with low-stakes assessments, if at all, when they have control over what happens in their classroom? How do instructors incorporate these assessments into the classroom, if at all?

We found that student participation in Moodle-quizzes was relatively high and that most students put in enough time to reasonably read each question and consider the answers. About two-thirds of surveyed students reported sometimes using the textbook while working on Moodle-quizzes, although this seems unusual, based on median times for completion. Students in the survey reported that they used the feedback that was given for incorrect answers, which implies that they were using it for self-assessment. Suggestions for future use of these quizzes by students included requests that this feedback be expanded. The high use of this feedback makes expansion worthy of future consideration. In particular, the system often provides little feedback for a correct answer – showing the responses to all of the answers might make the feedback more valuable for the students who guessed when they were not sure and happened to get the answer correct. Further, explicit encouragement of students to use the quiz for self-assessment might help close the loop as opposed to doing the quiz for the vague goal of "preparation for class". Instructors could also be encouraged to further incorporate this course component into their classes- in the survey, students suggested using these questions on exams, discussing them in class, as well as using them for pre- and post-testing to see what they had learned during class.
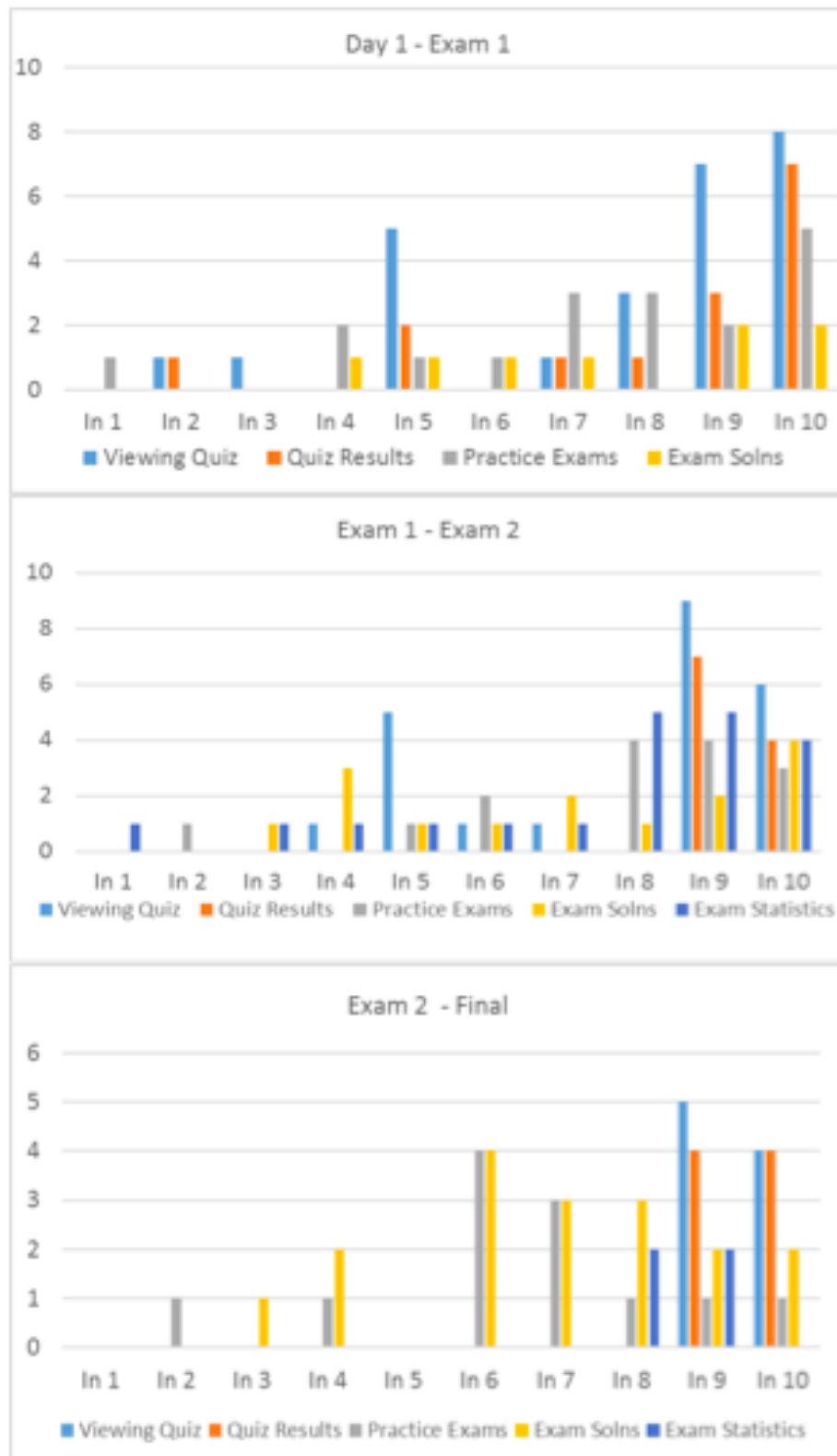
Figure 8: Instructor Moodle usage over the course of the semester.

Factors that might influence the generalizability of our results included the demographic of students: because of our course offerings this calculus class had no engineering students and is a terminal math class for many students who are taking it to fulfill a general math requirement. Moreover, a large proportion of the students had taken calculus in high school. In other schools, calculus students may have other goals in taking calculus and different familiarity with the material that would change how the approached similar low-stakes assessments.

Future work in the student perspective should assess the difficulty of quizzes to estimate a 'reasonable' time frame for doing each quiz to better assess the engagement of students in Moodle-quizzes. Indeed, the length of quizzes varied between two and nine questions, with a range of problems including very basic true-false questions, calculation problems, conceptual questions, and questions that asked students to read examples in the textbook and answer specific questions about them. Longer quizzes often had several simple questions that would push down the average time for completion of a single question, as opposed to shorter quizzes that did not necessarily have 'gimme' problems.

Instructor reactions to Moodle-quizzes suggest that instructors are gate-keepers to their classrooms. Without instructor buy-in, classroom goals such as sparking discussion among students are not likely to be accomplished. Instructors showed a wide range of engagement with this course element, from considering how to make it most useful for the students to not having any knowledge of this course component further than acknowledging that students must do it. One of the reasons that we might have seen such variety of engagement is that these quizzes were a relatively small portion of the grade. In interviews, several instructors reported that students could "do the math" and understood that this course component would not greatly affect their grade. To speculate, it is possible that instructors did not engage because students were graded on participation only and so did not exert any pressure on their teachers to make the course component more directly related to class. Still, this suggests that a 'top-down' approach to incorporating new elements in the classroom may not be successful unless instructors are involved in the discussion of the goals for the element and how to implement it in the course. Additionally, encouraging discussion among instructors and gathering feedback from them on the implementation might help both reveal problems and fix them.

Future work to understand the instructor perspective should again be by interview or group discussion, and should explicitly discuss the emergent themes of issues of pedagogical decision-making in a large group and the balance of controlling the classroom and receiving restrictions from course coordination. In this study these themes emerged in interviews but were not discussed explicitly. Moreover, we wonder how much the experience level and job security of the instructor influences the choice to engage with course components that are implemented from the top-down. Instructor non-engagement was not restricted to any particular experience-level. It would be particularly interesting to better understand the reasons for this non-engagement across experience-level: Are graduate instructors protective of their classrooms and anxious to make it 'theirs'? Are tenured professors settled on how they like to teach?

References

Amick H.L. Math Class- Have you seen the preview?, on Innovative Teaching Exchange on MAA Online (Mathematical Association of America). Available at http://www.maa.org/math-class-have-you-seen-the-preview.

Brost B. & Bradley K. (2006). Student Compliance with Assigned Reading: A Case Study. Journal of Scholarship of Teaching and Learning, 6(2), pp. 101-111.

Boelkins M. & Ratliff T. (2000) How We Get Our Students to Read the Text Before Class, in Innovative Teaching Exchange on MAA Online (Mathematical Association of America). Available at http://www.maa.org/how-we-get-our-students-to-read-the-text-before-class.

Bullock D., Callahan J., & Shadle S. (2015). Coherent Calculus Course Design: Creating Faculty Buy-in for Student Success, American Society for Engineering Education Annual Conference & Exposition, Seattle WA, June 2015.

Clump M., Bauer H., & Bradley C. (2004). The extent to which psychology students read textbooks: a multiple class analysis of reading across the psychology curriculum. Journal of Instructional Psychology, 31(3).

Clump M. & Doll J. (2007). Do the Low Levels of Reading Course Material Continue? An Examination in a Forensic Psychology Graduate Program. Journal of Instructional Psychology, 34(4), pp. 242-246.

Henderson C. & Rosenthal A. (2006). Reading questions. Journal of College Science Teaching, 35(7), pp. 46-50.

Hoeft, Mary E. (2012) "Why University Students Don't Read: What Professors Can Do To Increase Compliance," International Journal for the Scholarship of Teaching and Learning: 6(2), Article 12. Available at: https://doi.org/10.20429/ijsotl.2012.060212.

Iyasere M. (1984) Responding to Colleagues: Setting Standards in Multiple-Section Courses. Improving College and University Teaching, 32(4), pp. 173-179.

Kobler, J. (1978). Departmental Standards in Freshman Composition. College Composition and Communication, 29(3), pp. 264-266.

Lei S., Bartlett K., Gorney S., & Herschbach T. (2010). Resistance to Reading Compliance among College Students: Instructors' Perspectives. College Student Journal, 44(2), pp. 219-229.

Sappington J., Kinsey K. & Musayac K. (2002). Two studies of reading compliance among college students. Teaching of Psychology, 29(4), pp. 272-274.

Stickles P. & Stickles J. (2010). The Effect of Reading Assignments on Test Question Performance in Calculus. Mathematics and Computer Education, 44(3), pp. 205-212.

Zywno M.S. (2007) Optional Activities as a Way to Improve Student Engagement and Academic Achievement in a Large Engineering Class, International Conference on Engineering Education, Coimbra Portugal, September 2007.

Staying-Power in Business Ethics Education

Catherine Smith

Department of Philosophy, Cornell University
cmm472@cornell.edu

Introduction

There is a large body of research on teaching business ethics and much of it shares a common motivation: ethical improvement in the business world. A glance at the first sentence of the typical article about teaching business ethics makes the centrality of this motivation apparent: "Recent misconduct and highly questionable behavior in both religious and secular environments have fostered considerable distrust [and] cynicism... among the American populace toward the leadership of virtually all social institutions, especially business organizations" (Allen et al., 2005). "Society has lost faith in business integrity as fallout from widespread corporate scandals since the 1990s has affected the lives of literally millions of people" (Baker et al., 2012). "The multiple corporate scandals in recent decades continue to negatively impact the public's trust in business" (Schmidt and Davidson, 2013). In each of these introductions, the authors waste no time raising the specter of business immorality as a primary reason to look more closely at business ethics teaching.

The hope, shared by many researchers, is that better ethics education will make for improved ethics in a new generation of business leaders. But what really is better business ethics education? This question is a complex one in part because ethics itself can be a complicated matter, involving intense disagreements. In my research, I conducted interviews and surveys to learn something about one of the many questions that fits into this complex whole: What do learners themselves think they are getting out of their business ethics education that actually matters, that will stick with them enough to shape their thinking and behavior into the future?

Literature Review

With such a pressing real-world motivation in place, it is hardly surprising to find that there is a wide variety and a large quantity of literature on the teaching of business ethics, particularly on how effective ethics education is at actually shaping student thinking and behavior in a morally desirable way. Results of this research are not conclusive. There is certainly evidence that teaching business ethics can be effective in increasing student awareness of, attention to, and reaction to moral considerations (Nguyen et al., 2008; Weber & Glyptis, 2000). For example, Knotts et al. (2000) found, in a business school that emphasized ethics in its mission statement, that business students actually had superior ethical awareness compared to their non-business counterparts.

There is, however, no shortage of evidence that business ethics classes can fail to make any real difference. Smyth (2009), for example, found that business students were even more tolerant of unethical behavior, in a study that focused on cheating. Allen et al. (2005), found that senior undergraduate students in business, trained in ethics, placed no more value on traits like honesty, generosity, and compassion than freshmen in their program. As the authors explain, this result suggests that four years of participation in a curriculum that included the integration of business ethics had done little to solidify their moral character. In fact, it was found that senior students considered many of these traits less important and reported them as less "activated" by their education than freshmen. Ritter (2006) also found that students did not demonstrate more awareness of ethical issues or include more ethical components in their reasoning after being given ethics training. And although Cagle (2005) found that studying ethically model companies did give students more confidence that businesses are not all ethically bad, no improvement in the students' ethical standards accompanied this change. (These results were replicated in a follow-up study (Cagle, 2008).) Some studies also indicate that undergraduate business students begin their programs with low levels of ethical awareness (Galbraith and Webb, 2010; Lowry, 2003). Aside from this research, there is also the disheartening reality of continued business scandals, with Volkswagen's emission scandal taking over headlines as recently as September of 2015 (Hakim et al., 2015).

These mixed results could in part be due to differences in teaching. Business ethics classes are not always based in the same department, and different business programs choose to integrate ethics requirements into their curriculum in different ways, resulting in little standardization across the field (Petrick et al., 2011; Nguyen, 2008). The inconsistencies between different business ethics classes and programs

are serious enough that Avecedo (2013) considers some of them evidence that there is widespread misunderstanding of ethical behavior across the field. Studies do indicate that different ways of delivering content to students can make a large difference to the effectiveness of teaching ethical behavior (McManus et al., 2012), suggesting that best practices matter a great deal for successful business ethics education. In fact, much recent work on business ethics is devoted to sharing new methods of delivering business ethics content in hopes of using these to make improvements, along with reports on classroom outcomes (Baker & Comer, 2012; Schmidt et al., 2013; Thomson, 2011; Yiu, 2008).

A Possible Gap

In sum, the business ethics education literature can by and large be divided into two categories. On the one hand, there are quantitative research studies investigating the effectiveness of business ethics education. On the other there are mixed method and quantitative studies, which research specific classroom implementations of more precise strategies for teaching business ethics. One thing that seems to be missing is direct conversation with former business ethics students about their own experience of the difference their business ethics education made to them. If the desired outcome of business ethics education is for students to come away from class with changed minds that they will carry forward into their workplaces, it would be beneficial to find out what former students' have retained and applied from their business ethics experience.

Hearing from students what class activities, readings, or interactions seemed meaningful to them could help provide first-person support to claims underlying research which proposes new methods for teaching in the business ethics classroom. These proposals are often supported by plausible but unproven claims of their benefit, such as that ethics education will be more effective if it is made more personal or concrete to the student (Baker & Comer, 2012), or that students will only adopt ethics in a shallow, appearance-oriented way if their class has an implicitly "materialistic" worldview (Giacalone & Thompson, 2006).

Speaking to students in an open-format interview may also help to identify ways that current research on effectiveness may be overlooking effects that these classes have on students. There is no one, agreed-upon way to assess how effective business ethics education has been for strengthening the moral competence of students, and studies test for learning outcomes in business ethics classes in a variety of ways.

Effects on student morality are assessed in different studies by asking students to rate how important they think various ethical qualities are to business success (Allen et al., 2005), by having them respond to case studies (Keith et al., 2009; Ritter, 2006), and by having them rate how important various social issues are to them (Weber and Glyptis, 2000). But it is not clear that these measures are the best way to capture moral growth. To focus on one measure, many studies judge the success of ethics education by assessing how severely students would choose to punish unethical behavior (Smyth, 2009; Cagle, 2008, 2005; Ritter, 2006; Knotts et al., 2000)—but the assumption that severe punishment for unethical behavior is the most effective or ethical way of responding to it is itself an assumption that students might not share with researchers. Hearing from students themselves should allow for them to explain how their moral outlook has been affected by their class without having to correctly intuit beforehand how these effects would manifest themselves in different moral individuals.

In my study, the goal was to gather first-personal information from the currently under-researched category of former business ethics students, in the underutilized form of open-ended interviews.

Methods

I reached out to former business ethics students at Cornell University and in the surrounding area in Ithaca, New York to find individuals willing to be interviewed about their experience with business ethics education. I emailed former students of my own, who had taken a business ethics course with me the previous spring, and I also placed fliers with my information and a description of my study around the buildings of business-oriented fields of study. After my initial hopes that I would get a high response rate from my former students faded, I also added a few other methods to my strategy. I canvassed businesses in downtown Ithaca, reaching out especially to small business owners who may have been trained in business ethics; I sent out a general call to my contacts on social media; and I also designed a survey to be given to the students to whom I was currently teaching business ethics, near the end of the semester.

Participants

In all, I ended up with a small but diverse number of interviewees, and a larger group of survey respondents from my class. My interviewees were diverse in educational and employment status (from college sophomore to business master's student to small business owner), and they also varied in the type of business ethics teaching they had experienced

(from writing-intensive introductory courses in business ethics to business ethics units integrated over several classes to short-term community training sessions). This variation did not trouble me, as part of my goal in the study was to see what kinds of the full variety of methods for teaching business ethics might be effective across different career-trajectories and over long periods of time. The difficulty of finding people willing to be interviewed in person for no fee, however, did limit the quantity of interviews I was able to obtain. It also may have led to my interviewing a selection of individuals who were particularly generous and/or self-confident.

All students present on the day I conducted my in-class survey opted in, making for fourteen completed surveys. My survey participants were, for obvious reasons, more homogenous than interview participants. All were enrolled in a writing-intensive course for first-year students on the topic of business ethics (with the broader topic-listing of "Philosophical Problems"). Every freshman student at Cornell University is required to take at least one (and sometimes two) of these "writing seminars," as they are called, in some topic or other. Students are given the chance to rank options, and efforts are made to assign them to a class that focuses on a topic in which they are interested.

Given these conditions, it is likely that all the individuals surveyed were somewhat interested in philosophy, business ethics, or both (they at least preferred a philosophical business-ethics class to many, if not all, alternative options)—but they were not necessarily so interested that they would select the course as an extra-curricular course. Students also sometimes rank their choices with attention to which courses will fit into their schedule or with expectations about which topics will prove more or less time-consuming. This reality further lowers the concern that all students surveyed were outstandingly interested in philosophical or business-related thinking before beginning the course. Anecdotally, I did once observe a student saying to a friend before class that they really were not the "philosophy type."

Interview Method

My interview format was designed to be relatively open-ended, so that interviewees could have the latitude to discuss what seemed most important to them about their business ethics training. This was particularly important to me because one of my motivations for the study came from the concern that different people might experience changes in their moral awareness in distinct ways, not all of which might be captured in any given survey about moral competence. Each interview was recorded, audio-

only on my work computer. I began by asking participants what they remembered learning from their class. Then, I asked them if anything about the class had seemed particularly meaningful to them, and if so, why. Each participant was also asked if they thought that anything from the class had changed the way they thought or acted (again, with a follow-up question of why). I concluded each interview by giving the participant time to say anything else they might want to say, in case I had not asked them a question they were hoping I would ask, or in case they remembered anything else during our conversation that they wanted to add.

Survey Method

My survey questions were designed to resemble my interview questions and to preserve the open-ended nature of my inquiry (For example: 1) Is there anything we did or discussed in class that you think will stick with you even after class is over? Or did anything we did or discussed seem particularly meaningful or important to you? 2) If yes: Why do you think this thing/these things might stick with you after class? If no: What kinds of things, in your experience, typically do make for more meaningful learning experiences (this need not be in the classroom context)?). Participants were given paper surveys to fill out in handwriting. They were informed that doing so was voluntary and not a part of their assigned work for the class. Students interested in participating were instructed to fill out the surveys as honestly as possible, anonymously, and put them in an envelope at the front of the room. I left the classroom before writing began and waited for 10-15 minutes to make sure that student responses (or choices not to participate) were fully anonymous. When I returned, I collected the envelope and went back to my office.

Coding

To code my data, I first typed up all of the main ideas I heard or read expressed in the recorded interviews and surveys. Then, I went through the results looking for any commonalities. I noticed and tracked six rough categories, some of which were inter-related, as I will discuss in more detail in the next section. My categories were: realness, the prudential advantages of business ethics, ease or simplicity of concepts, other or differing perspectives, people, and tough decisions. On final analysis, the most consistently mentioned categories were other/differing perspectives and realness. Of four interviewees, three brought up differing perspectives in the course of discussing the class and what stayed with or seemed meaningful to them, and three brought up what I

have called realness. These categories were also safely the most consistently mentioned throughout written survey responses.

Results

Realness

Interviewees and survey participants alike voluntarily brought up aspects of the course that they considered to be the most real-world-relevant when asked what stood out to them and changed their thinking in their classes. One interviewee equated boring and "dry" material with material that did not have much "relevance to today's world" and said that such material did not really stay with them after class. Survey respondents noted that they thought they would remember theories that were "the most applicable and useful in [the] business context" or material with "practical applications." One student who speculated that they may forget most of the material from the class explained this by noting that philosophical ethical theories "can't be used to change everyday behavior." Overall, it did seem that many participants found that perceived real-world relevance made material seem more important and encouraged them to pay closer attention to it.

These results made sense with my experiences leading discussion in the classroom, as well. During the spring semester, as I was teaching the survey respondents, there was a noticeable increase in engagement in those parts of the class in which I asked students to use the moral theories we were learning to analyze real-world business events. Students also seemed more eager to participate when I offered analogies that were meant to bring business dilemmas closer to their school-related experiences, which were presumably more real-world-seeming to them at the time.

The Prudential Advantages of Business Ethics

Participants in the study also occasionally mentioned that they were persuaded by their business ethics class or training that being more ethical would lead to more business success, and that this was important to the material sticking with them. This came up more frequently among the interviewees than the survey participants, which may have to do with the fact that interviewees were adults who in most cases were (or had been) more fully involved than survey participants in the business world already (as a small business owner or full-time employee). One small-business owner who was interviewed found the prudential advantages of ethical behavior particularly important to their business ethics training overall, explaining to me that the consequences of

being prosecuted or fined are particularly devastating for small businesses. It may be appropriate to say that the prudential importance of ethics actually functioned for such participants as a personal specification of what it meant for a topic to be "real" or "relevant to the real world" for them.

Ease/Simplicity of Concepts

If, as suggested in my analysis of how the prudential advantages of business ethics mattered to different participants, different surrounding contexts or experiences can change what stands out to learners about their class material, that would help explain another interesting mismatch between my interviewees and my survey participants. The survey responses were the only reason I decided to track mentions of ease or simplicity of concepts as a category. Interviewees did not, in even a single case, cite simplicity or ease as a factor in what they retained from class or what seemed significant or behavior-changing for them.

One possible explanation for the fact that survey participants did mention this factor, which was evidently obsolete to interviewees, is that survey participants were students, surveyed at a time in the semester when all were preoccupied with final exams and papers. Bearing in mind that these individuals were embroiled in the task of trying to think about and accomplish many things at once, it is not surprising that easy and simple material would stand out to them as more likely to stick.

Other/Differing Perspectives

The other category that was mentioned with particularly notable frequently in my results, aside from realness, was the other/differing perspectives category. In some cases, the fact that some perspective was new seemed to be enough to provoke interest and to facilitating learning that participants found especially meaningful. One interviewee stated that the topic that stood out most in their memory of class was a topic the class put in "a different context" or in "a different light" than what the interviewee had previously heard in conversations about it. Survey respondents also wrote that the class had made them look at issues "in a unique way" and made them realize that different "scenarios" "can result in different ways of thinking or strategies." Multiple students cited their exposure to "different perspectives" (others used the phrases "new frame of reference" or "something I had not thought about before") as something that was interesting and would stick with them after class.

Differing perspectives were also mentioned specifically as something that students had learned to keep

in mind when making business decisions. Survey participants made frequent mention of stakeholder theory and the fact that this theory requires one to take into account the many different people who may be affected by a business decision.

Prevalent or differing perspectives were very frequently mentioned, however, in the connection to (spoken, written, or imagined) conversations. Various survey respondents wrote things like: "Learning my classmates' different type of views helped," "Other classmates and [the teacher] made good thought-provoking points and that... led to some changed thinking," and "Having open discussion and challenge in class led me to change my mind; I think if debate were more prevalent/intense I would have changed more."

Conversations also seemed significant to the mentions of perspectives in interviewees. One interviewee noted that they had changed their behavior as a consumer particularly because of the "interplay" between the perspective offered in their business ethics class and in a sociology class they were taking simultaneously. The interviewee found that their exposure to these contrasting ways of approaching the issue actually increased their confidence that they had good reason to change their behavior. Yet another interviewee spontaneously described one of the benefits of their business ethics education in terms of their ability to participate in new kinds of conversations. As they explained: *If I were to talk to you [before the class] about why I believed certain things, especially in the business world, [I would have just said] this is what I believe; I don't really have much behind it. Now I have more structure to my argument on why I think you should live a certain way.* It is worth noting that this interviewee was not one of the participants who had taken my own philosophy-oriented class: this makes their mention of being able to argue for something more striking than it might have been if they had taken a class in which philosophy, and hence argumentation, was front and center.

I also have had some anecdotal experience that fits well with the fact that participants were thinking about their learning in terms of their ability to competently have a conversation about an ethical decision with another person. When I taught business ethics for the first time, I once asked my students, somewhat frustrated by the unperturbed way in which they seemed to be accepting some serious business malpractices we had discussed, whether they really thought such things were not concerning. The response I got from a few students who spoke up, to which everyone else seemed to nod in agreement, was that they all thought it was clear enough that much of what happens in the business world is not ok. Still, they said, in the business world that was just how things were done—no one was going to listen to ethical reasoning. If this is how many people feel going into a business ethics class, it seems that one thing of value they might be looking to gain would be precisely the ability to talk to really talk through ethical issues with others. In the real world, deciding what would be best is only part of the process—after that, there will be conversations, with one's employees, boss, or stockholders.

People

The seeming importance of conversation brings me to another category I tracked in my coding: people. The people category was particularly difficult for me. I struggled with whether or not to remove the category, since mentions of people were also often affixed to other categories, especially differing perspectives. Mentions of different perspectives often came along with student awareness that these perspectives really belonged to other people—whether those people were individuals who could be affected by a business decision, classmates, or a teacher. One interviewee said that one of the things they would remember most clearly from their business ethics class going far into the future was a class poll in which students used clickers to indicate what they would be willing to do in order to secure a promotion. The interviewee expressed shock at how far some of their actual classmates claimed they would go.

In the end, I decided that it still made sense to think of people as its own category, since it seemed to provide insight into how there might be more unity between the responses highlighting realness and responses highlighting perspectives than might initially seem to be the case. The fact that different perspectives were actually discovered in living human beings, some of them in the classroom, was often significant to participants, perhaps because this made those perspectives more obviously real-world. One interviewee noted that it was "powerful" to see that people had such different opinions in class discussion on some of the issues, when they had gone into class expecting everyone to be more or less on the same page.

Tough Decisions

My final category was the tough decisions category. Participants did not bring this up with especially great frequency, but instances spread over both the surveys and the interviews. One interviewee explained that the case studies they remembered from class helped bring out the fact that sometimes there was "no clearly right answer" in ethical decisions, for instance. Survey participants mentioned that the class

had made them think "in-depth" about things they normally would not have thought so hard about and that the class "pushed" them to "think outside the box."

Toughness of cases also seems to have been associated with the fact that participants were made to think of differing perspectives. For instance, it is notable that one survey participant I quoted earlier, who linked their learning to class discussion and commented that even more debate would have been helpful, worded their comment in terms of "challenges" (especially "intense" ones) being edifying. The survey participant who wrote that the class had pushed them also did so right after commenting that the class taught them to consider many different groups of people and how they would be affected by a business decision. One interviewee commented that they had recognized "tough cases" particularly after pausing to put themself in others' places. Analysis of this category leads me to believe that part of the usefulness of differing perspectives for these learners actually came, perhaps counter-intuitively, from the fact that such perspectives complicate matters.

Overall

The responses I gathered with the open-ended questions for this study displayed a great deal of variety. This is evident, for instance, in the other/differing perspectives category. Although many participants noted that awareness of other/differing perspectives was part of what would stay with them after class and what had struck them during class, students pulled this awareness from different aspects of class. Some noted that it was actual interaction with classmates that brought this thought to mind, while others noted the fact that their class included arguments presenting various different sides of an issues debate, and still others came to this realization simply by coming to understand one particular theory of business ethics (stakeholder theory) that required them to think through the needs of many stakeholders before making an assessment on a case study. This variety seems to suggest that it makes good sense to retain diversity in content-delivery methods in these classes. The clear meaningfulness of simplicity to the survey participants and the lack of any mention of this in the interview participants also seems to speak in favor of the importance of thinking about the circumstances surrounding the classroom and the other challenges students are dealing with when designing a course that they will be able to really engage with and internalize.

The importance of realness to many of the participants suggests that it was true of them, as some authors in the literature have suggested, that they found relevance to the real-world and concreteness important. These factors speak in favor of the continued use of case-studies and theory application in the classroom. The importance of perspectives was a more surprising result, suggesting that students will find material more interesting if they think of it as located in a conversation between people (whether their classmates, authors of readings for the class, or people they may be conversing/arguing with in the future). This speaks in favor of facilitating discussions in which students can express opposing opinions to each other. It also suggests that it is a good idea to present not just case studies and theories, but also differing takes or perspectives on these materials.

One might expect that offering different perspectives would overwhelm students or leave them confused, but participants expressed none of these reservations. In fact, those who mentioned the difficulty or toughness that resulted from considering different perspectives found it to be helpful to their learning experience. One interviewee even noted that they remembered a guest lecture from a course they had taken years ago in which they asked the speaker if she had even found herself conflicted and sympathetic to the unethical side of the business ethics issue under discussion. The speaker basically refused to countenance this idea. This data point stands out because one of the things this interviewee remembered about their class was its failure to include the kinds of conversations and conflicts of opinion that were so meaningful to many other participants.

Conclusions

Given that my results for this study leaned heavily on students who were enrolled in a philosophy course, I think there is reason to be somewhat cautious about how generalizable they are. However, it is also the case, as I mentioned, that survey participants were not voluntarily enrolled in a philosophy extra-curricular: they simply selected a topic categorized under "Philosophical Problems" from among a list of options for a required course. In addition, the majority of interviewees did not have business ethics training that was presented in a philosophy context. Therefore, these results remain worth considering for other classrooms.

My study also did nothing to test student's moral competence in an objective way. So, although many participants expressed an eagerness to think about things from other points of view and especially to consider how everyone affected by a decision might view it, I cannot claim to know that they are better at actually carrying this out on the basis of their ed-

ucation or that they are more likely to actually follow through on their resolve in real-life situations. Still, gaining the awareness that moral decisions are complicated and often should be difficult is no small step, and it was encouraging to see these results.

In sum, this research provides at least tentative support for teaching business ethics with special attention to real-world applications, especially those relevant to learners themselves in their daily lives—but also with attention to providing differing perspectives (whether through readings or disagreement-friendly class discussion) that complicate the material at hand. It also implies some suggestions about what we might test for if we want to find out whether students are actually gaining the competencies that they think they have, and that they considered ethically significant, from the course. For instance, instead of testing for the ability to attach the appropriate punishment for wrongdoing in the workplace, we might test for the ability to identify those people who are likely to be affected by a given business decision, figure out how they will likely be affected, and come up with solutions that actively attend to these effects, granting such considerations real weight.

References

Allen, W.R., Bacdayan, P., Kowalski, K.B., and Roy, M.H. (2005). "Examining the impact of ethics training on business student values." Education + Training, 47(3), pp. 170-182.

Avecedo, A. (2013). "But, Is It Ethics? Common Misconceptions in Business Ethics Education." Journal of Education for Business, 88, pp. 63-69.

Baker, S.D., and Comer, D.R. (2012). "'Business Ethics Everywhere': An Experiential Exercise to Develop Students' Ability to Identify and Respond to Ethical Issues in Business." Journal of Management Education, 36(1), pp. 95-125.

Cagle, J.A.B. (2005). "Case Studies of Ethical Companies to Emulate: Effects on Ethical Perceptions of Finance Students." Journal of Financial Education, 31, pp. 41-56.

Cagle, J.A.B. (2008). "Using Vignettes in Introductory Finance Classes: Impact on Ethical Perceptions of Undergraduate Business Students" Journal of Education for Business, pp. 76-83.

Galbraith, D.D., and Webb, F.L. (2010). "Business Ethics: A View From the Classroom." Journal of College Teaching & Learning, 7(4), pp. 39-51.

Giacalone, R.A., and Thompson, K.R. (2006). "Business Ethics and Social Responsibility Education: Shifting the Worldview." Academy of Management Learning and Education, 5(3), pp. 266-277.

Hakim, D., Kessler, A.M., and Ewing, J. (2015). "As Volkswagen Pushed to Be No. 1, Ambitions Fueled a Scandal", New York Times, Sept. 26, 2015.

Keith, N.K., Perrault, H.R., Chin, M., Keith, M. (2009). "The Effect of Gender on the Importance of Business Ethics and Managerial Decisions: A Student Perspective." The Delta Pi Epsilon Journal, 51(3), pp. 125-136.

Knotts, T.L., Lopez, T.B., and Mesak, H.I. (2000). "Ethical Judgments of College Students: An Empirical Analysis." Journal of Education for Business, 75(3), pp. 158-163.

Lowry, D. (2003). "An Investigation of Student Moral Awareness and Associated Factors in Two Cohorts of an Undergraduate Business Degree in a Britsih University: Implications for Business Ethic Curriculum Design." Journal of Business Ethics, 48(1, Special Issue), pp. 7-19.

McManus, L., Subramaniam, N., and James, W. (2012). "A Comparative Study of the Effect of Web-Based Versus In-Class Textbook Ethics Instruction on Accounting Students' Propensity to Whistle-Blow." Journal of Education for Business, 87(6), pp. 333-342.

Nguyen, N.T., Basuray, M.T., Smith, W.P., Kopka, D., McCulloh, D.N. (2008). "Ethics Perceptions: Does Teaching Make a Difference?" Journal of Education for Business, 84(2) pp. 66-75.

Petrick, J.A., Cragg, W., and Sñudo, M. (2011). "Business Ethics in North America: Trends and Challenges." Journal of Business Ethics, 104(Supp.1), pp. 51-62.

Ritter, B.A. (2006). "Can Business Ethics be Trained? A Study of the Ethical Decision-making Process in Business Students." Journal of Business Ethics 68: 153-164.

Schmidt, C.D., Davidson, K.M., and Adkins, C. (2013). "Applying What Works: A Case for Deliberate Psychological Education in Undergraduate Business Ethics." Journal of Education for Business, 88, pp. 127-135.

Smyth, L.S. (2009). "Students' Perceptions of Business Ethics: Using Cheating as a Surrogate for Business Situations." Journal of Education for Business, 84(4), pp. 229-239.

Thomson, G.S. (2011). "Good Conversations: An Enhanced Model to Teach Business Ethics." Journal of International Education Research, 7(1), pp. 73-80.

Weber, J., and Glyptis, S.M. (2000). "Measuring the Impact of a Business Ethics Course and Community Service Experience on Students' Values and Opinions." Teaching Business Ethics, 4(4), pp. 341-358.

Yiu, C.Y. (2008). "A New Model to Help Students Develop Professional Ethics." Journal of Real Estate Practice and Education, 11(1), pp. 41-56.

Anarchy in the Classroom: The Efficacy of Self-Directed Learning for Critical Whiteness Pedagogy

Katherine Thorsteinson

Department of English, Cornell University
klt55@cornell.edu

Introduction

This autoethnographic study traces my experiences designing and teaching two freshman writing seminars on topics of race at Cornell University: "Theories of Blackness from Around the Globe" in Fall of 2016, followed by "Race and (Dis)ability" in Spring of 2017. As a white female graduate instructor who focuses on anti-racist pedagogy and praxis, I have wrestled with problems of power imbalance in the classroom. Firstly, I understand that certain lessons are better explained by those who have experienced racial oppression. Secondly, many of these so-called lessons might be triggering or simply beyond the scope of explanation, and it is not the responsibility of my students to take on the role of "expert" in these cases. Thirdly, the different racial knowledges my students are bringing into the classroom are largely determined by their own experiences and will thus require a very flexible or diverse pedagogical approach.

Moreover, these experiences tend to be quotidian, accumulating on an unconscious level through what we might call a racial habitus (Bonilla-Silva 2006). In particular, given the ubiquity and normativity of white privilege, (white) students may be less cognizant about the cultural specificity of whiteness and lack a critical historical knowledge of racism (Smith 2004; Nelson, Adams, & Salter, 2013; Brooms & Brice 2017). Due to the insistence of white fragility and guilt, white students may also be more comfortable discussing race, risking their own biases, and opening themselves up to criticism with a teacher who is also white (Picower 2009; Brooms & Brice 2017). Students of color, however, may be justifiably wary of white teachers given the long history of racism in education policy and practice. So, in designing my first course, I set out to locate, interrogate, and decenter my whiteness in the classroom. I turned to models of Self-Directed Learning (SDL) for extending agency and responsibility to my students over their own racial education.

Such an approach shaped multiple facets of the course: students personally selected their own readings each week from a list of choices relevant to our topic, organized the speaking order of their own discussions by passing a ball to "take the floor," chose their own essay themes and entirely determined the nature of their own final projects, presented their work to each other in peer review workshops and individual presentations, and collaborated with me to develop assessment rubrics. I also met several times with students both individually and as a class to create the syllabus, share suggested readings, evaluate my own instructional style, assess their level of interest and comfort with the course, and track their progress. Teaching voice was further diversified by several graduate student colleagues I invited to lead our seminars, the online conversations students had with each other on Blackboard, as well as some field trips we made to the university archives, library, and art museum.

I found throughout this process that students were very engaged in course content and excited about our class dynamic, but they frequently felt overwhelmed or confused with the logistics of this radical democratic approach. Upon reflection, I realized that we were inhabiting the double meanings of "anarchy." This was, as bell hooks has put it, "education as the practice of freedom" and yet could too easily devolve into chaos. Ultimately, I found that my approach to Self-Directed Learning helped center student voice, build community, and improve metacognitive learning. However, such strategies also contributed to feelings of insecurity, discomfort, or confusion. This is particularly problematic for critical whiteness pedagogy because these feelings are already likely to arise when white students confront anti-racist material. First-year students are also dealing with the stress of their new environments and competitive programs. Discomfort and confusion can actually be important for the racial (un)learning process, particularly for white students. However, unnecessarily contributing to this already highly emotive context makes this work more difficult.

In my second course I thus devised a number of strategies to harness these emotions for a critical whiteness pedagogy tailored for this particular context. Most importantly, I had to find a democratic model through which I could cultivate what Christina Sharpe (2016, p. 131) has called "an ethics of care" amongst my students and with myself. Without

this foundation of care and trust, I feared that defensiveness, resentment, and impatience would rule the classroom. Shifting democratic models, from individualism to collectivism, meant that students were better able to monitor, organize, and educate each other. Thus, in "Race and (Dis)ability," student groups chose course readings to supplement those that were required, led classroom exercises and lessons every Wednesday, engaged in peer review and collaborated on projects. I continued to reach out personally through email before and immediately after the beginning of term; ask students to complete questionnaires to better ascertain their needs, goals, and experiences; and honestly explain my own stakes in the material. I also set out the goals of our course, my rationale for the syllabus, and offered flexible submission dates as well as opportunities for resubmission. By minimizing the punitive aspects of learning, I felt students may be more likely to risk themselves and make mistakes—crucial orientations for any anti-racist pedagogy.

My research thus explores the following three interrelated questions: (1) How does social identity inform teaching identity? Or, given the inevitable power imbalances in the classroom, what does it mean to be a white woman teaching race courses? (2) Despite the majority of literature extoling Self-Directed Learning, we should be more attentive to the variations of this approach and the contexts in which it is used. For example, extending too much agency to students without the requisite support and structure may lead to intimidation, confusion, or exhaustion. When properly harnessed, these emotions may prompt white students to risk their ideological attachments or take on work normally relegated to racial minorities. However, such emotions may provide little pedagogical benefit to students of color and white students may eventually give up or shut down. What, then, are the utilities and limitations of Self-Directed Learning for critical whiteness pedagogy? (3) What does bell hooks's call for "education as the practice of freedom" mean in today's political and economic climate? It is important to make anti-racist pedagogy compatible whenever possible with our students' other academic and professional obligations—not just to appease our students, but also to account for how their ability to sustain critique and undergo transformation is tied to their other emotional attachments. Much of what students learn from critical whiteness pedagogy may target the operations of higher education, the neoliberal economy they are entering, and the notions of meritocracy structuring our sociopolitical order. Nevertheless, our methods for teaching these lessons should avoid contributing further to

feelings of insecurity, uncertainty, and abandonment.

Literature Review

Critical Whiteness Pedagogy

Throughout the stages of formal education from preschool to graduate school, students find very few opportunities to determine the materials or methods by which they are taught. "Occasionally, students do have some voice and choice," admits Kimberly M. Williams (2015 p. 161): "they sometimes get to choose classes and what they do projects on but, by and large, students have little to no choice over the most essential parts of their education." This problem is exacerbated for many students of color who also find their voices drowned out by what Brooms and Brice (2017) call "white noise." When white students deflect or disengage from discussing material that undermines their privilege, their silence produces a kind of "white noise" that can drown out or disrupt the learning process. Amidst this "ambient sense of privilege," conversations of race and racism are often interpreted as noise, meaningless chatter, or angry ranting. As they explain further, it is difficult for white students to discuss whiteness because of the "denial of their privilege, the myth of meritocracy, and [the] promotion of colorblind ideology" (Brooms & Bryce 2017 p. 152).

The same noisy silences may be found in the research of white educators, even in the literature on anti-racist pedagogy. For example, Picower (2009) notes that most of the research on race in education focuses on students of color as opposed to the predominately white educators teaching them. This absence is particularly problematic because when white teachers are challenged for their hegemonic understandings about race and difference they tend to resist through what Picower (2009) calls "tools of Whiteness" which are often emotional, ideological, or performative in nature. Ladson-Billings (2001, p. 81) similarly argues that white middle-class prospective teachers often "have little to no understanding of their own culture. Notions of Whiteness are taken for granted. They rarely are interrogated." Importantly, this lack of critical self-reflection can lead white educators to underestimate the relevance of teaching from a culturally relevant perspective which can have negative effects for students of color (Sleeter and McLaren 1995; Ayers et al. 1998; Ladson-Billings 2001; Cochran-Smith 2004; Jordan-Irvine 2003; Picower 2009).

Whiteness is under examined in education research and also goes unnoticed by many white teachers, so it is not surprising that critical whiteness pedagogies have rarely been discussed or developed in post-

secondary education. Indeed, Matias and Mackey (2016 p. 33) lament the dearth "of pedagogies specific to critical whiteness studies." "Beyond acknowledging white racial privilege," they continue, "how might educators engage in racially-just teaching when it comes to critical whiteness studies? What do critical whiteness studies embody and how does understanding the theory, and pedagogically implementing it, impact antiracist efforts?" Fewer still have considered these questions in a post-secondary context. Despite the persistence of exclusionary practices, colleges and universities are uniquely diverse institutions—bringing people together from across the country and around the world, increasingly open to different economic classes, political beliefs, sexual orientations, ages, religions, and cultural backgrounds. When designing my first course, "Theories of Blackness from Around the Globe," I thus experimented with Self-Directed Learning (SDL) as one such method to decenter and interrogate whiteness in the post-secondary classroom.

What Is SDL?

Schmidt (2000, p. 243) describes Self-Directed Learning broadly as "the preparedness of a student to engage in learning activities defined by him- or herself, rather than by a teacher." Referring to the relative willingness and competency of students, "preparedness" thus suggests that SDL is an acquired "skill and disposition" (Loyens et al. 2008). Occasionally Loyens et al. (2008, p. 415) reiterate this sense of the term: "SDL entails the ability to assess learning needs, effective planning, and time management, a critical evaluation of the literature resources, as well as a critical evaluation of their own SDL skills." For Candy (1991) on the other hand, SDL is both a process and a goal with the following four dimensions: "personal autonomy, self-management in learning, the independent pursuit of learning, and the learner control of instruction" (Loyens et al. 2008, p. 414). Knowles (1975) similarly defines SDL as a process; one in which students diagnose their learning needs, formulate goals, identify human and material resources, choose and implement appropriate learning strategies, and evaluate learning outcomes.

Despite this apparent disagreement about the nature of Self-Directed Learning, we can see that these definitions actually form two sides of the same coin. I would argue simply that thinking about SDL as a skill and disposition takes the student perspective, whereas thinking about it as a process and goal takes the teacher perspective. That is, teaching through a process of SDL enables students to acquire this disposition. So too, teachers will set SDL as a goal which students will then acquire as a skill. In reality, though, this strict divide between teacher and learner can be more complex. For, in cases of pure "self-teaching" these perspectives can be held by the same person (Shanley 2007). Moreover, as Loyens et al. (2008, p.414) emphasize, "Learning does not take place in isolation but in association with others such as teachers, tutors, and peers." While students bear much of the responsibility and agency for their own learning success, they should not be left stranded without any instructional guidance. In fact, methods of Self-Directed Learning usually require more preparation, monitoring, and feedback from teachers (Schmidt et al. 2007). The kind of work involved with teaching through SDL is less prescriptive and didactic, but more improvisational and dynamic. Thus, teachers must shift their pedagogical methods, skill-sets, and expectations.

How to Foster SDL?

Self-Directed Learning may involve active learning strategies such as preparation (goal setting, pre-assignment outlines), monitoring (schematic note-taking, cue card quizzes, consultation), and assessment (reflections, revision, peer review). But SDL also engages students at the fundamental level of course design and class plans, as exemplified by Problem-Based Learning (PBL). PBL emerged out of medical education during the mid-1960s as a way to help first-year students apply otherwise abstract material from anatomy, physiology, or biochemistry to their future professions as doctors (Barrows and Tamblyn 1980; Schmidt 1983). Instead of waiting until internships to work with real patients, students were asked to respond to realistic medical problems that physicians often encounter. While the term PBL originated in the medical context, similar methods have been status quo in law and business education for some time (Barrows 1996). Still, medical educators have made at least one very important contribution to PBL: while law and business students tend to engage with case studies only after some competency has been achieved, medical students are expected to work through these problems before any significant knowledge of the topic has been acquired. Thus, PBL uses these problems as the very starting point of the learning process (Loyens et al. 2008).

It is this "creative" aspect of PBL that qualifies the method as a form of Self-Directed Learning. Since students are not simply given the relevant information to respond to these case studies in advance, they must find creative means of devising answers through their own research, discussion, and critical thinking (Schmidt 1983; Hmelo-Silver 2004;

Loyens et al. 2008). Moreover, PBL is fundamentally student-centered in that it requires individuals to identify knowledge deficits, generate their own learning issues, critically evaluate literature resources, apply new knowledge, confront real-world situations, as well as engage in both solitary and collaborative reflection (Hmelo and Lin 2000; Loyens et al. 2008). Partially inspired by the Problem-Based Learning approach, K. M. Williams (2015 p. 162) and her colleague K. Hussain attempted to create a "truly democratic university classroom" by enabling students to make the following six decisions: choosing 1) an inquiry group, 2) whether or not they wanted to attend certain weekly meetings, 3) course readings and theories, 4) assessment mechanisms, 5) their preferred method of class instruction, 6) debate groups and topics. Williams (2015) found that adapting to this flexible class structure was challenging at first for teachers and students alike, but eventually students became more confident, engaged, and effective in the learning process. My own ethnographic research is highly indebted to this model developed by Williams and Hussain. Yet, while their course material was centrally concerned with race issues, the logistical and emotional problems arising from student racial diversity did not motivate their case study. Building upon their work, I note the often invisible articulations of whiteness in the classroom and thus consider how our very approaches to democracy are racially inflected.

Different models of SDL have also been developed by museum, library, archive, and gallery curators (Wright 1989; Banz 2008). For example, the "self-guided tour" and "interactive exhibit" have become common place in most museums, encouraging visitors to "forge their own path [and to redevise] the path as new experiences warrant" (Banz 2008, p. 50). This method of SDL has implications for architectural planning and the design of exhibits. The construction of walls and doors or the placement of objects, images, and text will all shape visitors' possible interpretations. Others have developed methods of Self-Directed Learning through technology. Bryan (2015), for example, explains how collaborative learning can be fostered and new knowledges can be formed through "virtual clouds"—social networking sites like Facebook, LinkedIn, and Twitter or production tools like YouTube, Flickr, and SlideShare. "The collaboration in the digital world does not take away the benefits of self-direction," Bryan (2015, p. 43) maintains; "rather it heightens the need for individuals to succeed together." Yet she is careful to point out that technology is not inherently liberating, communal, or constructive. Nor are its users neces-

sarily proficient in SDL on their own. Rather, the possibilities and pitfalls of technology are a call to action. Given that information of all kinds and qualities is increasingly available to people with a range of educational experiences, it is more important than ever to make Self-Directed Learning our priority in the classroom.

Besides Problem-Based Learning, open syllabi, self-guided tours, interactive exhibits, and virtual clouds, others have demonstrated how SDL can be incorporated into the classroom through instructor and peer feedback. Choi and Anderson (2016) designed what they call the SelFeed system (Self-Directed Learning with Feedback), a weekly writing assignment in which students summarize lecture content, develop their own questions in response, and then provide logical answers. Students were also encouraged to research relevant primary literature and to brainstorm in groups. Answers were evaluated according to logical reasoning as opposed to accuracy, and grades were relatively low-stakes. Once these SelFeeds were submitted, the instructor gave feedback, corrections, and two or three follow-up questions. Students chose one of these to answer in a short paragraph. Based on surveys and final grades, the SelFeed method indicates an increased sense of satisfaction with student-teacher relations and significant academic improvement (Choi & Anderson 2016). These five examples represent only a handful of the methods educators have developed for SDL, but I have surveyed these in particular because they approximate the approaches I have taken in my teaching.

Why Foster SDL?

Proponents of SDL claim that it enables "deep level processing" and "lifelong learning" (Candy 1991; Boyer, Edmondson, Artis, & Fleming 2014). Deep-level processing involves skills such as elaboration, pattern recognition, and identification of core principles, whereas surface-level strategies tend to consist of rehearsal and memorization. Thus, deep-level processing requires the higher levels of Bloom's Taxonomy—analyzing, evaluating, and creating (Entwistle & Peterson, 2004; Loyens et al., 2008). Loyens et al. (2008) point out that when classroom activity is focused on assessment and teacher demands students can assume their success depends on following or even imitating instruction. As they summarize, "Learning environments that foster SDL... are believed to promote deep-level processing because learners have the freedom to choose what they learn and how they learn it (Loyens et al. 2008, p. 415).

Quantitative studies have supported these hy-

potheses, demonstrating the benefits of self-regulation—one component of SDL—to student learning at multiple levels of study. For example, Mason (2004, p. 283) observed "two rigorous strategic approaches to reading comprehension for thirty-two 5th-grade students who struggle[d] with reading." One approach encouraged metacognition before, during, and after reading while the other more conventional approach asked students to read silently and then answer specific questions posed by the teacher. While no significant differences in self-efficacy, motivation, or written comprehension were found between groups, Mason found drastic improvements in oral reading comprehension for those students who engaged metacognition. Moreover, SRL is a good predictor of future academic performance (Minnaert & Janssen 1999; Loyens et al. 2008) and others have discovered its benefits for the workplace-learning setting.

Several studies have demonstrated that SDL is positively correlated with high-demand skills and qualities like motivation, creativity, inquisitiveness, and confidence (Edmondson, Boyer, & Artis 2012). Boyer et al. (2013) explore the relationships between Self-Directed Learning and five key nomologically related constructs for effective workplace learning: control over events and outcomes, motivation, performance, teamwork, and confidence. Through a meta-analytic review of SDL research from various academic disciplines across five countries during the past thirty years, they discovered that all five of these variable relationships evidenced significant positive correlations (Boyer et al. 2013). The explanation for these results is simple: when college students learn how to be proactive and flexible through Self-Directed Learning they are better able to anticipate organizational needs and adapt their own behavior accordingly (Tobin, 2000; Cron, Marshall, Singh, Spiro, & Sujan, 2005; Artis & Harris, 2007).

Self-Directed Learning does not only positively affect academic and professional performance, it can also be an effective strategy for teaching against-the-grain. Indeed, gaps or biases in existing literature can themselves function as "teaching moments" if problem solving and critical thinking are intended learning outcomes. For example, in a course about LGBTQ related health issues, Kirkpatrick et al. (2015) used SDL to teach beyond the textbook. By conducting their own research nursing students were forced to confront the latent homophobia in their field and found creative means of overcoming the absence of LGBTQ educational materials in healthcare organizations. For not unrelated reasons, Davis, Cook, and Ostenson (2015) discuss one of the most significant reasons for their decision to engage Self-Directed Learning: "trust and respect." While most widely used in post-secondary and adult education (Loyens et al. 2008), Davis et al. (2015) engaged a group of 105 ninth graders in SDL. They selected those students who fell below the benchmark in Reading and English the previous year, comprising the bottom twenty-five percent of scores. As they explain, well-meaning educators merely contributed to the years of stigmatization experienced by these so-called "remedial", "deficient," "struggling" students. It is counter-productive and arrogant to think that the "magical instruction" of individual teachers can effectively "fix" anyone (Davis et al. 2015, p. 42). In this case, Self-Directed Learning consisted of a real attempt to extend agency and responsibility to students in a way that directly contended with classroom dynamics of power.

These last two examples are particularly relevant to my own research. Instead of lecturing on certain race topics, for example, Matias and Mackey (2016, p.41) report using SDL to challenge white teacher candidates to "'look it up [themselves]' because this is the same process exacted on people of color, women, and any other marginalized person [who] experiences Euro- and male- centric curricula." In my own courses, SDL prompted students to fill the gaps in existing literature, asked white students to take on the burden of research usually carried by students of color, and also opened up more opportunities for students of color to determine their own racial education. Moreover, because SDL requires a great deal of respect and trust, white teachers can avoid the problematic emotions of pity and shame that often motivate them towards anti-racist pedagogy in the first place (Matias & Mackey 2016).

Methodology

I primarily employ an autoethnographic approach in this teaching as research study, but I also consider the work, questionnaires, and evaluations I received from students in both courses. I documented my teaching methods and classroom experiences in three primary ways: personal journal, peer consultation, and public website. Every week I made brief notes about the effectiveness of past exercises, my ideas for future classes, and student progress. These observations are at once the most candid and also the most subjective. To gain some more objectivity, I asked faculty to provide me feedback on the classes they witnessed. I also worked in peer collaborations and training groups through which I had the opportunity to air some of these classroom experiences and ideas. Such consultations have been invaluable for

gaining perspective from other educators. Finally, my website has been a place where I have compiled some of these documents into a narrative about my pedagogical trajectory as a whole. This online portfolio has prompted me to communicate my teaching decisions to an audience both in and out of the education field. Reflecting on my personal experiences in the classroom through these three sources has helped me understand how social identity informs teaching identity and, in particular, how my whiteness shapes the way students can engage with race. As Brooms and Brice (2017, pp. 147-148) emphasize, in "addition to our pedagogical approach and teaching experiences. . . our racialized and gendered identities impact how students receive the content—and resist it as well. Thus, the autoethnographic approach is suited quite well for [this] inquiry." Because of its reflexivity, autoethnography is also appropriate for a teaching as research study. That is, autoethnography provides a "way of knowing" through one's own identities and experiences (Dyson 2007; Brooms & Brice 2017).

The student work I draw on in this study has all been dis-identified and given consensually. I reflect primarily on the overall trends I observe in the quality of student work and levels of student engagement with particular assignments. Due to time constraints and the nature of changing research questions, I did not collect survey information from my first course "Theories of Blackness from Around the Globe." At the time of this course, my questions focused on the problem of whiteness in the classroom but did not yet engage with the pitfalls of Self-Directed Learning. It was only after my own self-reflection and peer consultation that I considered how SDL may not only be an asset to critical whiteness pedagogy, but also a potential hindrance. My research questions and theoretical models grew directly out of the teaching itself, and has thus been an exercise in "grounded theory method" (Glaser & Strauss 1967; Picower 2009). For this course, then, I rely on official student evaluations and my own reflections. The survey data from my second class, "Race and (Dis)ability," was collected a month after the course had finished and grades had been finalized. Although this deferral may have skewed some of the results due to forgetfulness or distraction, I felt it would also prompt students to be more honest and possibly less conflicted in their answers. To account for this gap in time, I also consult the official evaluations students provided on the last day of semester. These responses will offer a better snapshot for their feelings about the course at that time.

My own scholarly interests and political investments have led me to take a variety of critical race courses and diversity training seminars, and I have contributed my own research to the fields of African American folklore, diaspora studies, queer studies, and education. Equipped with these experiences and theoretical frameworks, I knew that my identities in the classroom would matter. I was determined to make anti-racist pedagogy about more than the reading lists my students would consume or the essay topics they would regurgitate. For this reason, I was excited and nervous to create my first ever course from the ground up in fall of 2016. My nervousness arose from a number of sources: Was I equipped to teach about theories of Blackness as a white woman? How would I deal with stubborn white students or outright racism? What if my whiteness "got in the way" of the anti-racist work I wanted us to do together? Attempting to decenter my whiteness and my power as teacher, I prompted students to make the following six decisions: choosing 1) weekly readings from a suggested list, 2) how to moderate class and online discussion, 3) which theoretical models to focus on and how to integrate them, 4) essay topics and final projects, 5) assessment methods, 6) classroom activities and fieldtrip destinations.

Clearly the tradition of Black studies is varied, rich, nuanced, and complex. To say a half semester course that considers this body of thought from a global perspective is ambitious would be an incredible understatement. The task is impossible, but this inevitable failure was also built into the flexibility of the course structure. For, students would choose their preferred readings, not to find the "best" theory or gain the "fullest" understanding of what Blackness is, but rather to frustrate any common sense notions of race and to theorize for themselves what Blackness could indeed be(come). As I explained in the description,

> This course is not at all about mastery. Indeed, we should be wary of all the historical baggage that word carries. The course readings and assignments are designed so as to extend as much agency and responsibility to you as possible. But they have also been designed with a certain acceptance of incompleteness and failure. We will all inevitably leave this semester with a mere sketch of some ideas, questions, directions— each one of these unique from, yet inspired by, others in this class.

All of the choices I made in "Race and (Dis)ability" regarding course development, classroom instruction, or assessment were made in critical response to the successes and failures of this previous one. Firstly, I

reduced the number of required and suggested readings dramatically. Instead of offering buffet style reading options every week, I solicited suggestions from students before the semester and periodically asked students which readings they would rather skip or spend more time with as a group. Students also had the opportunity to assign short readings to the class for their group presentations each week. These groups consisted of three or four students who developed brief lectures on topics of their choice and led classroom exercises that engaged the readings they assigned. Furthermore, I tended to provide more prompts for online and class discussion when students appeared confused or unfocused. I reduced the number of assignments in this second course, allowing students more time for revision and peer review. In most other ways my approaches to the two courses remained very similar, but these changes indicate how I shifted from an individual to a collective democratic model. Both classes consisted of sixteen students in their first year of undergraduate study at an Ivy League university who were enrolled in the fall and spring semesters respectively. These freshman writing seminars are required for all students in their first year of study, though they do have some choice over which specific courses they take. A number of different disciplines in the humanities, sciences, and social sciences offer versions of these seminars, and each course tends to be designed by the instructor in accordance with general guidelines. The demographics of both courses I taught differ in important ways. "Theories of Blackness from Around the Globe" involved all women, roughly 75% were students of color and most of whom were Black identifying. Still, the classroom was very diverse for the variety of languages these students spoke, the range of countries with which they maintained ties, and the academic disciplines they planned to enter. "Race and (Dis)ability" was comprised of roughly 65% women and 65% students of color who identified by a variety of different racial classifications. Many of these students also spoke multiple languages, several maintained ties with other countries, and most aspired to different professions. To put these numbers in perspective, consider that the overall undergraduate population is roughly 40% white, 17% Asian, 11% Latin@, 6% Black or African American, 4% multiracial, and 0.3% indigenous.

Results

Observations

While such indicators are difficult to measure, I noticed a pronounced shift in energy and emotion between both semesters. By the end of "Theories of Blackness from Around the Globe," many students appeared less engaged and excited about course material. Some dominated the classroom discussion while others mentally "checked-out." In my journal reflections for "Race and (Dis)ability," however, I noted that students completed their readings more often and thus generally demonstrated a deeper understanding of course material in our class discussions. These students were also more willing to engage in group projects, share the floor equally in class discussions, and converse in online and out-of-class environments. Attendance improved and work was submitted more punctually in the second semester. Students in "Race and (Dis)ability" were also more likely to contact me for meetings or ask for consultation via email. Moreover, the overall class average for final grades increased from a B+ to an A- between semesters. Based on my journal reflections and these final course grades, I understand that students in "Race and (Dis)ability" felt more safe and supported in their work.

Interviews

I was lucky enough to retain two of my students across both semesters—a rarity since students usually choose these writing seminars based on whether they can accommodate the scheduling of other degree requirements. I take this as a sign that at least these two students appreciated the course content and my teaching methods enough to prioritize continued study with me. The fact that they explicitly opted for my class in the second semester may pose some confirmation bias. Nevertheless, I conducted an interview with them to gather more detailed and in-depth qualitative data to support the survey results and student evaluations I received. These students agreed with my general observations of the first semester I outline above. As they noted, "Theories of Blackness from Around the Globe" was largely successful when it came to extending student agency and decentering my whiteness in the classroom. Nevertheless, they encountered the following drawbacks: the order of in-class discussion could be unclear and some students dominated the floor, online conversations could be unfocused and many students neglected to take part, the assignment of readings could be confusing or overwhelming. In their account, the collective democratic approach of "Race and (Dis)ability" mostly corrected these problems. They also emphasized that having fewer readings in the second semester enabled them to engage more deeply with course material. Significantly, though, even after considering some of the drawbacks in the first semester, these students supported the use of

Self-Directed Learning. As one student explained, "In both classes I loved the freedom that we were given in terms of assignments and day to day class time."

Survey and Evaluations

When asked "How 'in control' of your own education did you feel?" students from "Race and (Dis)ability" answered 55% "very," 45% "somewhat," 0% "average," and 0% "less or not at all." Comparing this concrete data from "Race and (Dis)ability" to my observations in "Theories of Blackness from Around the Globe," I suspect the near equal tallies for "very" and "somewhat" demonstrate a slight decrease in the general feelings of agency between both semesters. I believe this slightly lesser sense of student agency reflects my shift from individual to collective democratic models. When asked to check any number of course facets that contributed to this sense of agency, students responded as follows: 100% "choosing essay topics and final projects," 91% "leading lessons in groups," 82% "verbal/email interactions with instructor," 64% "instructor's method of moderating class discussions," 55% "peer review," and 36% "solicitation of syllabus ideas." This data demonstrates that students overwhelmingly felt greater control over their own learning by the end of the course, and this was largely the result of making course material more flexible through open essay topics and student led classes. Also central to this sense of agency was the personal and communal support students found in each other and with me. The balance between independence and guidance was established through our commitment to regular meetings, classroom attendance, and online communications. It is important to note that no students reported feelings of insecurity, confusion, or frustration with the course structure or methods of instruction except three (18% of data pool) who noted some initial uncertainty about the openness of essay topics. Since this feature of the course increased a sense of agency for 100% of respondents, it appears to me a necessary and beneficial compromise.

Significantly, there was complete affirmation to the question "Do you think this emphasis on student agency helped you better reflect upon and understand issues of race and disability?" As one student explained, "I left this class knowing so much more about intersectionality and the everyday ways I interact with race. Of all the classes I took as a freshman this is the only one I think about almost everyday... The independent thinking I used while writing my essays, especially working through the concepts to figure out how they apply to my sources and my life,

definitely helped me form my own thoughts about the concepts we learned." Other students echoed this point about how they were better able to reflect on course content in terms of their own lives: "learning about issues of race and disability involves a great deal of drawing upon personal experience, so allowing us to be quite self-directed was beneficial." It "allowed me to learn more about the intersection of race and disability," said another student, "as well as be able to apply that to my own race." In particular, students noted that this collective approach to Self-Directed Learning enabled them to better reflect on their own privileges: "I was able to learn about my own privileges," one student explained, "when it comes to how race and disability intersect in my own life."

This collective method of SDL also enabled me to highlight and decenter my own privileges as a teacher. As one student noted, "I think that my teacher focused on race and disability from a very unbiased approach. I think she allowed herself to understand her students of color in such an amazing way." Students also note how this collective approach to SDL contributed to what Christina Sharpe calls "an ethics of care" in our classroom: "I learned lots from other students in the class because our instructor always urged people to voice their opinions," said one student. Another agreed: "I strongly believe that the culture developed in our classroom fostered deeper conversations and examinations of critical topics. Speaking from the 'I perspective,' it felt as if there were no restrictions preventing me from speaking my mind." While this lasses-faire attitude might be a tad misleading (the community we created in the classroom encouraged students to gently "check" each other), it does reflect my attempts to step back and avoid punitive interactions. Finally, students also described how greater freedom over their learning was vital to learning the practice of critique itself: "Student Agency helped us do more of the interpreting as opposed to many other traditional English classes where teachers often overlook student agency and lead discussions with their own interpretations." Thus, the collective approach to SDL is not only beneficial to engage students in course material but is essential to the practice of critique itself.

Discussion

While existing literature focuses on the educational benefits of Self-Directed Learning, little research has been conducted into its value for critical whiteness pedagogy. My own experiences suggest that such practices are highly contextual and that SDL may be an appropriate anti-racist tool depending on the specific methods used. However, in small freshman sem-

inars dealing with ethically weighty and emotionally difficult content, certain forms of SDL can actually impede the learning process. I believe my research contributes to these fields by distinguishing kinds of Self-Directed Learning (namely, collective and individual), outlining how SDL can be used for critical whiteness pedagogy to different ends depending on student identity or experience, and thus offering a number of different strategies for harnessing student emotions. In any case, it is important to be upfront with students about the rationale for such decisions and to discuss their levels of comfort throughout the learning process. Some compromises will inevitably need to be made. For example, the 18% of students who felt some unease with even collective forms of SDL should be balanced with the 100% of students who found educational benefit in these methods.

Indeed, for anarchist educator and theorist Paulo Freire, doubt and disagreement are the foundations of freedom (39). While some students may become uncomfortable when they are developing this critical consciousness ("conscientização"), he believes "Such an individual is actually taking refuge in an attempt to achieve security, which he or she prefers to the risks of liberty" (36). "Men and women rarely admit their fear of freedom openly," he continues, and "they confuse freedom with the maintenance of the status quo; so that if conscientização threatens to place that status quo in question, it thereby seems to constitute a threat to freedom itself" (36; original emphasis). Significantly, Freire advocates a method of Self-Directed Learning to achieve critical consciousness: "a dialogical and problem-posing education" (40). Nevertheless, I think it is naïve and unfair to believe that students' emotional responses to these methods are irrelevant to their learning success. To employ more radical forms of SDL despite student discomfort would in fact undermine the whole democratic enterprise. Ultimately then, closer attention should be given to the different experiences and identities our students are bringing into the classroom. Such differences will inevitably impact the ways they engage with you and the methods of SDL you employ.

References

Artis, A. B., & Harris, E. G. (2007). Self-directed learning and sales force performance: An integrated framework. Journal of Personal Selling and Sales Management, 27(1), 9-24.

Ayers, W., Hunt, J.A. and Quinn, T. (1998). Teaching for social justice: A democracy and education reader, New York, NY: W.W. Norton.

Banz, R. (2008). Self-Directed Learning: Implications for Museums. The Journal of Museum Education, 33(1), 43-54.

Barrows, H. S. (1986). A taxonomy of problem-based learning methods. Medical Education, 20, 481-486.

Barrows, H. S. (1996). Problem-based learning in medicine and beyond: A brief overview. In W. H. Gijselaers (Ed.), New directions for teaching and learning (vol. Vol. 68, (pp. 3-11)). San Francisco: Jossey-Bass.

Barrows, H. S., & Tamblyn, R. (1980). Problem-based learning: An approach to medical education. New York: Springer.

Boekaerts, M. (1999). Self-regulated learning: Where we are today. International Journal of Educational Research, 31, 445-457.

Bonilla-Silva, E. (2006). Racism without racists: Color-blind racism & racial inequality in contemporary America. Lanham, MD: Rowman & Littlefield.

Boyer, S. L., Edmondson, D. R., Artis, A. B., & Fleming, D. (2014). Self-directed learning: A tool for lifelong learning. Journal of Marketing Education, 36(1), 20-32.

Brooms, D. R., & Brice, D. A. (2017). Bring the noise: Black men teaching (race and) white privilege. Race and Justice, 7(2), 144-159.

Bryan, V. C. (2015). Self-directed learning and technology. Education Digest, 80(6), 42.

Candy, P. C. (1991). Self-direction for lifelong learning. San Francisco: Jossey-Bass.

Choi, Y., & Anderson, W. J. (2016). Self-directed learning with feedback. Journal of College Science Teaching, 46(1), 32.

Cron, W. L., Marshall, G. W., Singh, J., Spiro, R. L., & Sujan, H. (2005). Salesperson selection, training, and development: Trends, implications, and research opportunities. Journal of Personal Selling and Sales Management, 25, 123-136.

Davis, J., Cook, R., & Ostenson, J. (2015). Empowering stereotyped readers through self-directed learning. English Journal, 105(1), 42-47.

Dyson, M. (2007). My story in a profession of stories: Auto ethnography—An empowering methodology for educators. Australian Journal of Teacher Education, 32, 26–48.

Edmondson, D. R., Boyer, S. L., & Artis, A. B. (2012). Self-directed learning: A meta-analytic review of adult learning constructs. International Journal of Education Research (IJER), 7(1), 40.

Entwistle, N. J., & Peterson, E. R. (2004). Conceptions of learning and knowledge in higher education: Relationships with study behaviour and influences of learning environments. International Journal of Educational Research, 41(6), 407-428.

Freire, P. (2000). Pedagogy of the oppressed. (M.B. Ramos, Trans.) 30th anniversary ed. New York: Continuum.

Greveson, G. C., & Spencer, J. A. (2005). Self-directed learning—The importance of concepts and contexts. Medical Education, 39, 348-349.

Hmelo, C. E., & Lin, X. (2000). Becoming self-directed learners: Strategy development in problem-based learning. In D. Evensen, & C. E. Hmelo (Eds.), Problem-based learning: A research perspective on learning interactions (pp. 227-250). Mahwah: Erlbaum.

Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? Educational Psychology Review, 16, 235-266.

hooks, b. (1994). Teaching to transgress : education as the practice of freedom. New York, NY: Routledge.

Jordan-Irvine, J.J. (2003). Educating teachers for diversity: Seeing with a cultural eye, New York, NY: Teachers College Press.

Kirkpatrick, M. K., Esterhuizen, P., Jesse, E., & Brown, S. T. (2015). Improving self-directed Learning/Intercultural competencies: Breaking the silence. Nurse Educator, 40(1), 46-50.

Knowles, M. S. (1975). Self-directed learning: A guide for learners and teachers. New York, NY: Association Press.

Ladson-Billings, G. (2001). Crossing over to Canaan: The journey of new teachers in diverse classrooms. San Francisco, CA: Jossey-Bass.

Loyens, S., Magda, J., & Rikers, R. (2008). Self-Directed Learning in Problem-Based Learning and its Relationships with Self-Regulated Learning. Educational Psychology Review, 20(4), 411-427.

Mason, L. H. (2004). Explicit self-regulated strategy development versus reciprocal questioning: Effects on expository reading comprehension among struggling readers. Journal of Educational Psychology, 96(2), 283-296.

Matias, C. E., & Mackey, J. (2016). Breakin' down whiteness in antiracist teaching: Introducing critical whiteness pedagogy. The Urban Review, 48(1), 32-50.

Minnaert, A., & Janssen, P. (1999). The additive effect of regulatory activities on top of intelligence in relation to academic performance in higher education. Learning and Instruction, 9(1), 77-91.

Nelson, J. C., Adams G., & Saltor P. S. (2013). The Marley hypothesis: Denial of racism reflects ignorance of history. Psychological Science, 24, 213–218.

Paris, S. G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. Educational Psychologist, 36, 89-101.

Picower, B. (2009). The unexamined whiteness of teaching: How white teachers maintain and enact dominant racial ideologies. Race Ethnicity and Education, 12(2), 197-215.

Schmidt, H. G. (1983). Problem-based learning: Rationale and description. Medical Education, 17, 11-16.

Schmidt, H. G. (2000). Assumptions underlying self-directed learning may be false. Medical Education, 34, 243-245.

Schmidt, H. G., Loyens, S. M. M., Van Gog, T., & Paas, F. (2007). Problem-based learning is compatible with human cognitive architecture: Commentary on Kirschner, Sweller, and Clark. Educational Psychologist, 42, 91-97.

Shanley, P. F. (2007). Leaving the "empty glass" of problem-based learning behind: New assumptions and a revised model for case study in preclinical medical education. Academic Medicine, 82, 479-485.

Sharpe, C. (2016). In the Wake: On Blackness and Being, Durham, NC: Duke University Press.

Sleeter, C.E. & McLaren, P. (1995). Multicultural education, critical pedagogy, and the politics of difference, Albany, NY: State University of New York Press.

Smith, W. A. (2004). Black faculty coping with racial battle fatigue: The campus racial climate in a post-civil rights era. In Cleveland D. (Ed.),

A long way to go: Conversation about race by African American faculty and graduate students (pp. 171–190). New York, NY: Peter Lang.

Tobin, D. R. (2000). All learning is self-directed: How organizations can support and encourage independent learning. Alexandria, VA: American Society for Training and Development.

Williams, K. M. (2015). Doing Research to Improve Teaching and Learning: A Guide for College and University Faculty. New York, NY: Routledge.

Winne, P. H. (1995). Self-regulation is ubiquitous but its forms vary with knowledge. Educational Psychologist, 30, 217-221.

Wright, P. (1989). The Quality of Visitors' Experiences in Art Museums. In Peter Vergo (Ed.), The New Museology (119-148). London: Reaktion Books.

Zimmerman, B. J. (1989). A social cognitive view of self-regulated academic learning. Journal of Educational Psychology, 81, 329-339.

Student Engagement with Course Content and Peers in Synchronous Online Discussions

Allison M. Truhlar[1,2,*], Kimberly M. Williams[2], and M. Todd Walter[1]

[1] Department of Biological and Environmental Engineering, Cornell University
[2] Center for Teaching Innovation, Cornell University
[*] Corresponding Author: amt94@cornell.edu

Abstract

As higher education institutions in United States introduce online courses to their curricula and offer these courses to growing audiences, there is an increasing desire to understand how best to engage students with both course content and their peers. In one online introductory-level sustainability course, synchronous chats have been the primary means by which students interact with one another. This study examines the effects of assigning chat roles and facilitating self and group reflection on student-content and student to student interaction outcomes in synchronous chats. We also considered what occurred within group reflections to inform how they are structured in the future. We found that assigning roles increased the proportion of critical student-content interactions and significantly increased critical student-student interactions; $t(9.42) = 0.94$, $p = 0.19$ and $t(12.85) = 3.10$, $p = 0.004$, respectively. Self- reflections had no effect on either interaction type. Groups completing group reflections had a significantly greater proportion of critical student-content interactions in the third chat and critical student-student interactions in the fourth chat than groups that did not complete the group reflections; $t(7.99) = 2.07$, $p = 0.04$ and $t(6.73) = 1.94$, $p = 0.05$, respectively. This suggests group reflections help students to maintain critical thinking throughout the course of the semester. Based on these results, combined with the qualitative analysis of the group reflections, we plan to keep roles going forward and eliminate the self-reflections. Furthermore, to increase the effectiveness of the group reflections, we recommend changes to increase student agency and ability to convert their ideas into change during subsequent chats.

Introduction

The introduction of the World Wide Web in 1991 allowed for the growth of online education in the United States (Sun & Chen, 2016). Twenty years after this milestone, 89% and 60% of public and private four-year colleges in the US, respectively, offered either fully online or blended online courses (Parker, Lenhart, & Moore, 2011). More recently, "massive open online courses" (MOOCs) were introduced to online education and heralded as a way to make higher education accessible to anyone with a computer. The percentage of higher education institutions offering MOOCs has steadily increased from 2.6% in 2012 to 11.3% in 2015 (Allen & Seaman, 2016). This paper explores strategies for promoting student engagement and learning in synchronous online discussions, particularly in the case where the class is too large for the primary instructor to facilitate every discussion group. At a time when the higher education landscape is increasingly shifting towards online learning, and in many cases, online learning for the "masses," this investigation is a timely addition to the conversation surrounding student engagement and achievement of learning outcomes in online education.

Interaction in Successful Online Learning

The US Distance Learning Association indicates that interaction is an integral component of quality distance education, including online education (Holden & Westfall, 2007). In any educational context, interaction falls into three categories: student-instructor, student-student, or student-content (Moore, 1989). A meta-analysis of 74 studies, found that implementing at least one type of interaction intervention in distance-education courses increased student understanding of course content (Bernard et al., 2009). Among the three interaction types, student-student and student-content interactions had significantly greater effect sizes on student learning.

In online courses, the majority of student-student interactions take place in asynchronous (e.g. discussion forums students can access on their own time) or synchronous (e.g. live chat rooms) discussions carried out over a technological platform. Aside from individual assignments, discussions are where many student-content interactions occur. Therefore, there is great interest in understanding how to best encourage and support interactions in these discussion formats. Over the past two decades many interventions

have been studied for their ability to improve outcomes of student interactions in online discussions. Examples include assigning roles to students (Wise, Saghafian, & Padmanabhan, 2012) and asking students to reflect on their contributions to the discussion (Kayler & Weller, 2007).

Assigning discussion roles to students has been suggested to have many benefits including increasing student participation (Tagg, 1994), interaction (Hara et al., 2000), and knowledge integration and construction (Schellens, 2005; Strijbos et al., 2007). However, the research on whether these outcomes are actually achieved is inconclusive. Wise and colleagues (2012) suggest this is because research has focused on roles, but not on the actual functions the roles perform in conversation. Therefore, depending on how the instructor or researcher specifies the role to students, roles with the same name could be asked to fulfill different functions or vice-versa. Wise and colleagues (2012) identified six functions common to the roles assigned in 12 studies. Of these six, three were suggested to be most helpful in discussions, namely: (1) give direction, (2) summarize, and (3) provide critical response. The "give direction" function is typically assigned to the student expected to start the conversation and can be expanded to include keeping the conversation on topic for its duration (Hara et al., 2000; Persell, 2004; Tagg, 1994; Wise et al., 2012; Zhu, 1998). The "summarize" function asks students to synthesize what has been stated in the discussion up to a certain point, usually set at the end of the discussion period. Posts made by students assigned this function are consistently associated with higher levels of thinking due to the synthesis required by the function description itself (De Wever et al., 2007; Schellens, 2005). Summary posts have also been shown to stimulate higher-order responses from other students (Wise & Chiu, 2011). The "critical response" function is defined as "the right to be critical of ideas in the discussion" (Wise et al., 2012). A lack of critical comments on other students' thinking is common to online discussions (De Wever et al., 2008; Schellens, 2005; Schellens et al., 2007). Assigning the critical response function gives students permission to challenge their peers, which they might otherwise avoid (Wise et al., 2012).

Both individual and group reflections are commonly used pedagogical tools to promote better group work (Gibbs, 1994). However, there are few instances of these techniques being applied to online discussions, even though this specific type of group work has the unique quality of producing a written transcript. Therefore, instructors have the ability to hold students accountable of what actu-

ally transpired in the discussion. Multiple researchers have suggested that the reflective reading of transcripts as an intervention to improve students' online discussion strategies is an area that needs more investigation (De Wever et al., 2008; Murphy & Jerome, 2005; Walker, 2004). Qualitative work has found that structured, written self-reflection on contributions to an online discussion can allow students to identify their strengths and areas for improvement (Kayler & Weller, 2007; Murphy & Jerome, 2005). Furthermore, Kayler and Weller (2007) found that students who favored written reflection in their learning style were particularly empowered by this activity, and used it to determine how they could continue to improve their own contributions even when they were not happy with the group as a whole. Students who did not favor written reflection as a learning tool tended to blame bad group discussion experiences on perceived deficiencies in the participation of their group-mates. Therefore, one of the outcomes of this study was the modification of the self-reflection activity to accommodate diverse learning styles with the inclusion of an in-person group-reflection. This work seeks to build on the work by Kayler and Weller (2007), and determine the improvements in chat quality, if any, gained from the addition of group reflections.

Critical Interactions

As noted by Ertmer and colleagues (2011), measuring interactions in online discussions comes with challenges. Specifically, not every post is meaningful and relevant to course content (Ertmer, Sadaf, & Ertmer, 2011). Likewise, many student-to-student posts might be social interactions and not active discussions of course content. Therefore, purely quantitative measures of interaction (e.g. post counts) have not been as useful as those mixed with qualitative measures (e.g. post quality), often defined in terms of critical thinking (Ertmer et al., 2011; Walker, 2004; Wise & Chiu, 2011). For the remainder of this manuscript, "critical interaction" will be considered, as opposed to all student-content or student-student interaction.

Bloom's original taxonomy (Bloom et al., 1956) is often applied to measuring critical thinking about course content. The higher levels of Bloom's original taxonomy (i.e., analysis, synthesis, and evaluation) correspond to the cognitive processes associated with critical thinking (Pear et al., 2001). For student-student interactions, critical thinking occurs when participants hold ideas "publicly accountable" and work to reason out their understanding of those ideas (Mercer, Wegerif, & Dawes, 1999). This is op-
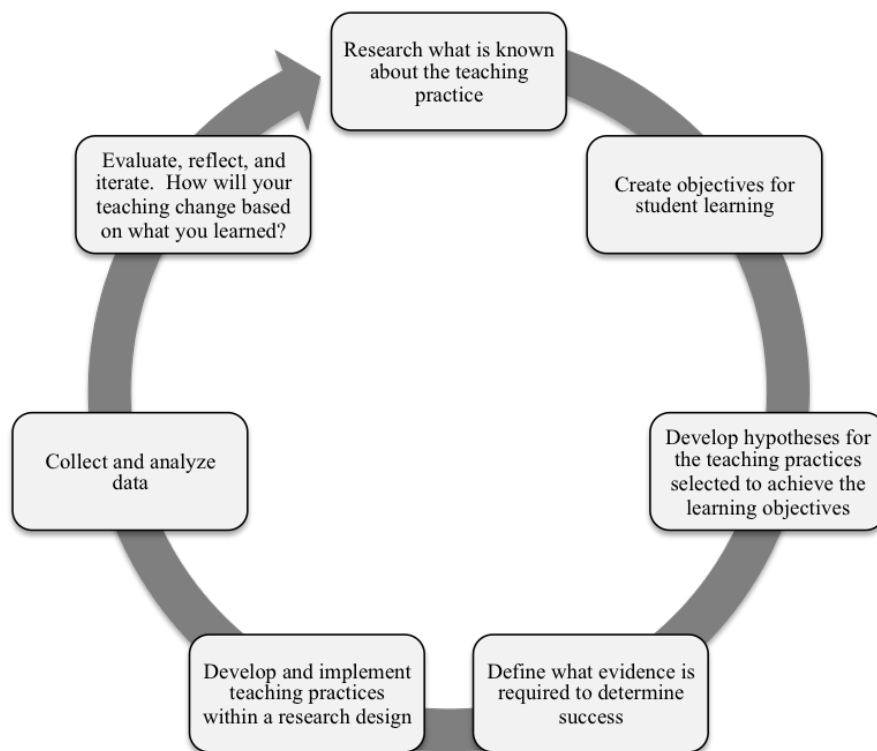
Figure 1. Teaching-as-Research process (Kwako et al., 2005; Williams, 2015).

posed to a discussion where participants reflexively agree or disagree, and thereby only superficially engage with the ideas put forward. Indicators of critical thinking in a conversation are questions that elicit further explanation of a provided fact or opinion, questions that require a defense of a line of argument, and statements containing a counter or alternative argument (Walker, 2004). These question types are labeled by Walker (2004) as "probe", "challenge", and "counter" questions, respectively.

Context

Our research is situated within the framework of "Teaching-as-Research," a formalized process by which an instructor applies research methods to evaluate their own teaching practice in order to improve learning outcomes for their students (Kwako, Courter, & Wright, 2005). The steps of this process are outlined in Figure 1.

Our study pertains to an introductory-level sustainable development course that has been offered in an entirely online format since its inception in 1998. During this time period, synchronous chats, generally held four times over the semester, have been the primary means by which students engage with one another and develop a learning community. In spring 2015, enrollment in the course exceeded three hundred students, requiring the use of 40 undergraduate teaching assistants (TAs) to facilitate chats (one group per TA) and complete other tasks such as grad-

ing. In fall 2016, enrollment was capped at 200 students and 16 undergraduate TAs were recruited. In both semesters, the first author of this paper served as the head graduate TA and also as the primary coordinator for the undergraduate TAs.

In spring 2015, the first author closely read and qualitatively analyzed a selection of chat transcripts to understand how students were engaging in the chats. Three main themes emerged from this analysis, two of which are directly relevant to this study: (1) students were not critically engaging with the course information, as evidenced by a majority of postings being made at low levels of thinking on Bloom's taxonomy, and (2) students were not critically engaging with information provided by their peers, as evidenced by a lack of challenges to new information. The first author had an intuitive sense from reading through chat transcripts throughout the semester that students were not fully engaged. For example, as soon as new questions were posted in a chat, it was common for a rapid series of long responses to follow. The questions were provided ahead of time to allow students to prepare, and the first author suspected that students were simply copying and pasting pre-written responses to earn participation points for being present during the chat. This engagement analysis anecdotally confirmed through conversations with past students.

As a result of these findings, for the last chat in the spring 2015 semester we changed the second
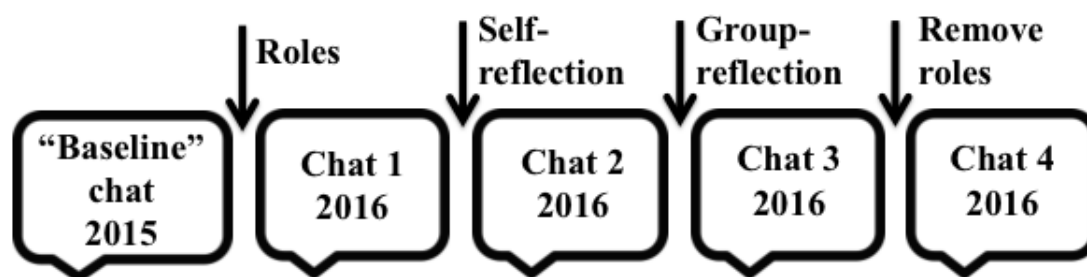
Figure 2. Timeline of interventions used in the online chats.

half (about 30 minutes) of the chat format to center around a debate, as opposed to the previous series of open-ended questions. The idea behind this change was that students would be forced to respond spontaneously to each other, as they could not completely anticipate the arguments of the opposing team. This format change still did not resolve the problem of low student engagement and led us to search the literature for other intervention options. As the literature reviewed here indicates, the majority of online interaction research relates to asynchronous, as opposed to synchronous, discussions. However, we decided to implement one intervention that has been routinely shown to increase both student-content and student-student interactions in asynchronous discussions: assigning discussion roles. In addition, we decided to test the effect of self- and group-reflections on student-content and student-student interaction, as this is an area that has been identified in the literature for further research.

Purpose

This study examines the effects of assigning chat roles and facilitating self and group reflection on student-content and student-student interaction outcomes for an online synchronous chat. The main hypotheses addressed in this study are: (1) assigning chat roles will increase the proportion of chat posts that contain critical student-content and student-student interactions, (2) self-reflections will further increase the proportion of chat posts containing critical student-content and student-student interactions compared to roles alone, (3) when roles are assigned, groups completing group-reflections will have a greater proportion of chat posts containing critical student-content and student-student interactions compared to groups completing only self-reflections, and (4) when no roles are assigned, groups completing group-reflections will have a greater proportion of critical student-content and student-student interactions than groups without group-reflections. In addition to these four hypotheses, and in an effort to

inform the development of future group-reflection assignments, we ask the question: what happens during an online, asynchronous, group-reflection process?

Methods

Interventions

The changes made to the course for the fall 2016 semester were guided by the emergent themes from the earlier qualitative analysis and the existing literature. A timeline of the interventions is shown in Figure 2. Students were assigned to a discussion group from the start of the course and conducted all their chats within this group. All chats were based around a one-hour version of the informal debate structure introduced in spring 2015. As the first intervention, before the first chat each student was assigned one of three specific roles to focus on during the debate. In brief, the three roles were the Traffic-Director, who was responsible for keeping the discussion on topic, the Questioner, who was responsible for asking critical questions of their peers, and the Synthesizer, who was responsible for identifying themes among posts and highlighting topics that still need to be addressed. Students were assigned roles for Chats 1, 2, and 3, and switched them each time.

Following the first chat, students were asked to complete a self-reflection on how they fulfilled their assigned role, drawing on examples of both things they did well and things they could improve on from the chat transcript. To ensure their familiarity with all three roles, they were also asked to reflect on examples from the two roles they were not assigned.

The second chat followed the same procedure as the first, but after the chat, eight of the sixteen groups were assigned to complete both a written self-reflection and an asynchronous group reflection on what went well in the chat and what could be improved. The other eight groups were assigned to complete a written self-reflection only, and to complete an asynchronous group discussion activity that built off the topic discussed in the synchronous chats.

Table 1. Chat interventions, with corresponding hypotheses and t-test results.

Intervention	Hypothesis	Dependent variable	t-value	df	p-value
Chat 1 – assign roles	Assigning chat roles increases critical student-content and student-student interactions. Ho: Chat 1, 2016 = Chat 4, 2015 Ha: Chat 1, 2016 > Chat 4, 2015	Proportion of higher level thinking posts	0.94	9.42	0.19
		Proportion of critical questions	3.10	12.85	**0.004**
Chat 2 – self-reflection	Self-reflections further increase critical student-content and student-student interactions. (Paired t-test) Ho: Chat 2, 2016 = Chat 1, 2016 Ha: Chat 2, 2016 > Chat 1, 2016	Proportion of higher level thinking posts	0.32	9	0.38
		Proportion of critical questions	-1.14	9	0.85
Chat 3 – group-reflection	Group-reflections further increase critical student-content and student-student interactions, as compared to self-reflections alone. Ho: Chat 3, 2016, group reflections = Chat 3, 2016, no group reflections Ha: Chat 3, 2016, group reflections > Chat 3, 2016, no group reflections	Proportion of higher level thinking posts	2.07	7.99	**0.04**
		Proportion of critical questions	-1.03	7.98	0.83
Chat 4 – remove roles	Groups completing group-reflections have greater proportion of critical student-content and student-student interactions than groups without group-reflections, when no roles are assigned. Ho: Chat 4, 2016, group reflections = Chat 4, 2016, no group reflections Ha: Chat 4, 2016, group reflections > Chat 4, 2016, no group reflections	Proportion of higher level thinking posts	0.65	5.96	0.27*
		Proportion of critical questions	1.94	6.73	**0.05**

*Note: If both the yes group and no group "high" outliers are removed, the p value becomes **0.04**; t (3.53) = 2.40.

Table 2. Bloom's original taxonomy, following Bradley et al. (2007).

Level	Example
Opinion	"I think..." with no "because"
Knowledge	Cite reading
Understanding	Paraphrase reading
Application	Relate to own experience
Analysis	Compare information
Synthesis	Draw connections
Evaluation	Justify position

The third chat followed the same procedure as the second chat. The fourth chat followed the same procedure as the second and third chats, except students were not assigned chat roles. Table 1 summarizes the interventions used in each chat and how they relate to our four hypotheses.

Measuring Student-Content and Student-Student Interactions

The first forty posts in five chat transcripts from the last chat in 2015 and ten chat transcripts from each 2016 chat were coded for critical student-content interactions, using the higher levels of thinking from Bloom's original taxonomy (e.g., analysis, synthesis, evaluation) as a proxy. The level of thinking for each post was coded based on a modified version of Bloom's original taxonomy (Table 2; Bradley, Thom, Hayes, & Hay, 2008). If a post included multiple statements, it was coded according to the highest level of thought it contained. The transcripts were also coded for critical student-student interactions using the three types of questions (probe, challenge, or counter) suggested by Walker (2004) as a proxy. One-sided t-tests were used to test each hypothesis; see Table 1 for a full description of the hypotheses and corresponding t-tests used. All statistical analyses were performed with R version 3.2.3 (R Core Team, 2015).

Qualitative Analysis

In order to understand what was happening in the group-reflections, we closely read transcripts from three of the five groups participating in group-reflections. These groups, arbitrarily named A, B, and C, were selected because they had the same students participating in each chat. The other two of the five groups completing group-reflections had minor changes in student chat attendance over the course

of the semester due to scheduling conflicts, although all the students in these groups did complete a group reflection. General themes were identified following the initial reading. We then re-read the transcripts, looking for evidence of each theme in individual posts and labeling them accordingly. The number of times a theme was mentioned was tabulated.

Results

Quantitative Results

Hypothesis 1.

Compared to a chat with no roles assigned, chats with assigned roles had a greater proportion of critical student-content and student-student interactions; $t(9.42) = 0.94$, $p = 0.19$ and $t(12.85) = 3.10$, $p = 0.004$, respectively (Table 1, Figure 3). However, only the increase in the proportion of critical student-student interactions was significant.

Hypothesis 2.

The addition of pre-chat self-reflections to the chat format did not increase either the proportion of critical student-content interactions, or the proportion of critical student-student interactions; $t(9) = 0.32$, $p = 0.38$ and $t(9) = -1.14$, $p = 0.85$, respectively (Table 1, Figure 4).

Hypothesis 3.

The proportion of critical student-content interactions was significantly greater for groups completing group-reflections, compared to groups that did not complete group-reflections; $t(7.99) = 2.07$, $p = 0.04$ (Table 1, Figure 5). There was no difference in the proportion of critical student-student interactions between groups in the two treatments; $t(7.98) = -1.03$, $p = 0.83$ (Table 1, Figure 5).
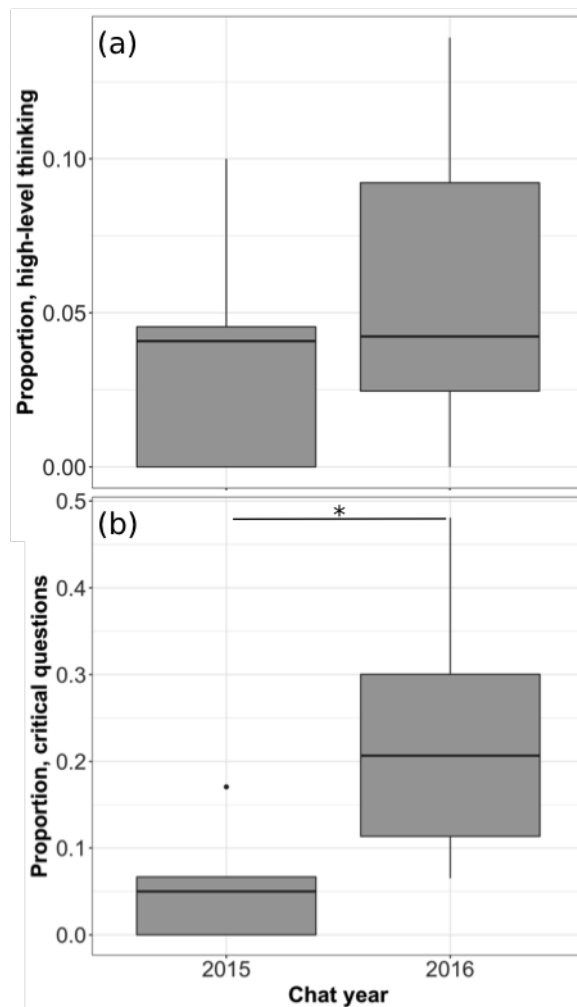
Figure 3. Critical (a) thinking and (b) questions before and after roles.

Hypothesis 4.

When roles were removed for the fourth chat, the groups that completed a group-reflection had a greater proportion of critical student-content inter-actions than groups that did not complete a group-reflection; $t(5.96) = 0.65$, $p = 0.27$ (Table 1, Figure 5). If "high" outliers from both the group with reflections and the group without reflections are re-moved, this difference becomes significant; $t(3.53) = 2.40$, $p = 0.04$. High outliers were defined as val-ues greater than the 75th percentile plus 1.5 times the interquartile range, or the range between the 25th and 75th percentile (Tukey, 1977). In general, the groups completing group reflections maintained higher-level thinking in the chats over the course of the semester (Figure 5). Furthermore, groups completing group-reflections had a greater proportion of critical student-student interactions in the fourth chat than groups that did not complete the group-reflection; $t(6.73) = 1.94$, $p = 0.05$ (Table 1).

Qualitative Results

Three major themes arose around the group reflection process. First, among the groups that completed group reflections, each group differed in the amount of perceived agency (as defined and discussed below) they expressed about their chat experience. Second, each group varied in their ability to suggest concrete changes to the chat. Third, as a whole, groups tended to favor suggesting logistical changes to the chats over learning-outcome related changes. All the above themes held true both with and without role assignment.

Agency.

One group, "Group A," demonstrated more agency in both the first and second group-reflection com-pared to the other groups (Figure 6). Here, we define agency as "feeling enabled and empowered to act" and solve problems (Kayler & Weller, 2007). When members of Group A discussed aspects of the chat they thought were least effective, they tended to em-phasize what they could do, or had already done, to
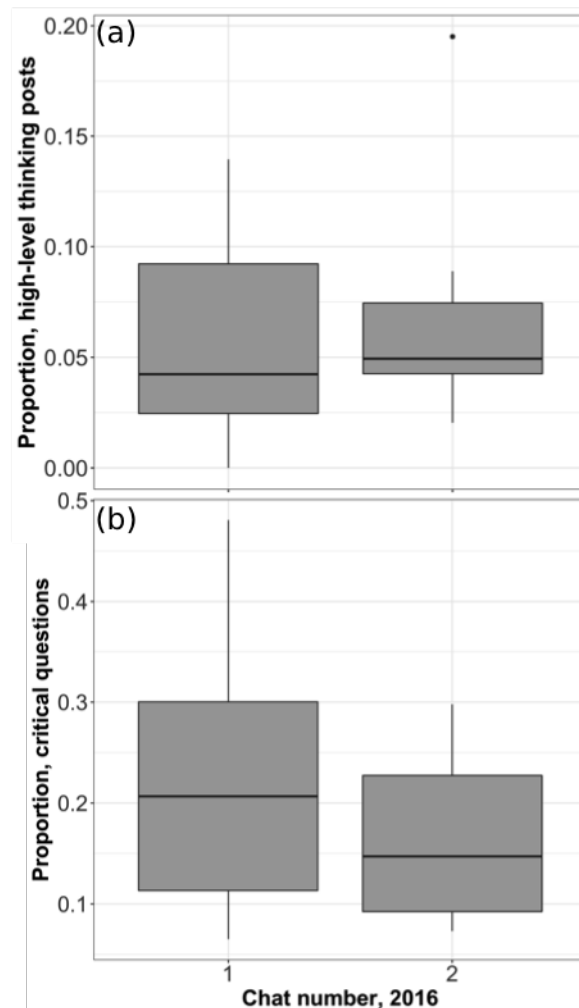
Figure 4. Critical (a) thinking and (b) questions before and after self-reflections.

affect change (Example 1). Many also put forth a narrative that they were learning and improving with each chat (Example 1). Contrary to these patterns, members from the other two groups expressed feeling that the ability to change what was happening in the chat was out of their hands, placing responsibility on inadequacies in the chat platform, other students, or the teaching assistant facilitating the discussion (Example 2).

Example 1: *"The most effective thing about our group's chat was the organization and smooth manner it was carried out [in] because we had discussed it beforehand. After having two chats, we knew how out of hand it can get. So we went subject by subject by first tackling one point, then another, and then the last one."* Example 2: *"I think that the digital format of the chat is responsible for the constraint on timely responses. . ."*

Concrete versus vague suggestions for improvement.

Members of Group A consistently proposed concrete actions they could take to improve ineffective

aspects of the chat (Figure 7; Example 3). In the other two groups, members often did not propose actions to address ineffective aspects of the chat. When they did, these suggestions were vague, as indicated by the frequent use of words like "hopefully" and "maybe" (Example 4).

Example 3: *"I agree with you that the fast-paced nature of the debate is probably the main issue. We should definitely utilize the Raise Hand function."*

Example 4: *"Hopefully next time, all questions will be answered or at least expanded on."*

Mixed focus on logistics and learning outcomes.

The chat aspect that was identified most frequently by each group as needing improvement varied, but overall tended to favor logistical aspects over actions directly related to achieving the target pedagogical outcomes (Table 3).

Discussion

Assigned roles and group-reflections emerged as the most useful interventions for achieving our pedagogical goals of critical, student-content, and,
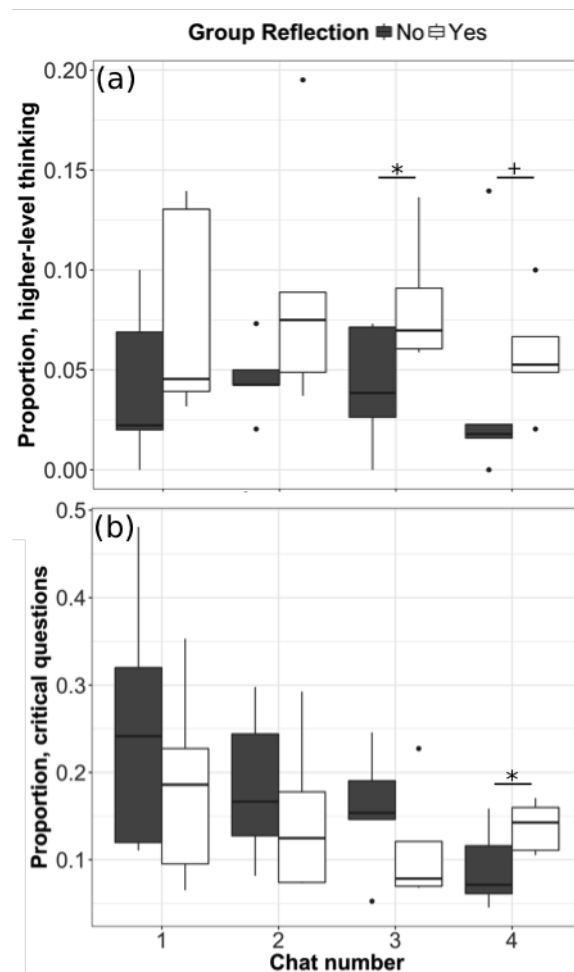
195

Figure 5. Critical (a) thinking and (b) questions with and without group-reflections.

student-student interactions in the synchronous online chats. Specifically, roles significantly increased student-student interaction, while the group-reflections served to maintain higher-level thought throughout the semester, even when roles were removed for the final chat. Group-reflections also resulted in significantly more student-student interactions when roles were removed. The third intervention we tested which was self-reflections, did not result in gains in either higher-level thought or student interaction.

Roles

Our results are in line with previous work that demonstrates roles increase student-student interaction in online chats (Hara et al., 2000). Of the posts directed to another student in Chat 1, 2016, and the 2015 chat, an average of 59% and 24%, respectively, were content-based questions. This increase in questioning suggests much of the student interaction in the 2016 chat was driven by the Questioner role. There is qualitative evidence that other roles, specifically the Traffic Director role, increased

student-to-student posts as well. In Examples 5.1 and 5.2, Teammates A and C are the Traffic Directors. In these exchanges, the Traffic Director posts did not include information relevant to the chat content but were rather intended to organize group members to respond to questions. No organization by students in preparation to answer questions was seen in the 2015 chat, which lacked roles.

Example 5.1: *Teammate A (Traffic Director): Any thoughts on this Teammate B?*

Example 5.2: *Teammate C (Traffic Director): I can address Student X's question. Teammate C (Traffic Director): And then Teammate D idk [I don't know] if you want to address Student Y's?*

Roles did not increase the proportion of posts made at higher levels of thinking. Two roles, the Synthesizer and the Questioner, were specifically implemented due to previous work demonstrating that the functions contained within these roles cause an increase in both critical student-content and student-student interactions (De Wever et al., 2007; Schellens, 2005; Wise & Chiu, 2011; Wise et al., 2012). A couple of observations can explain the failure of these
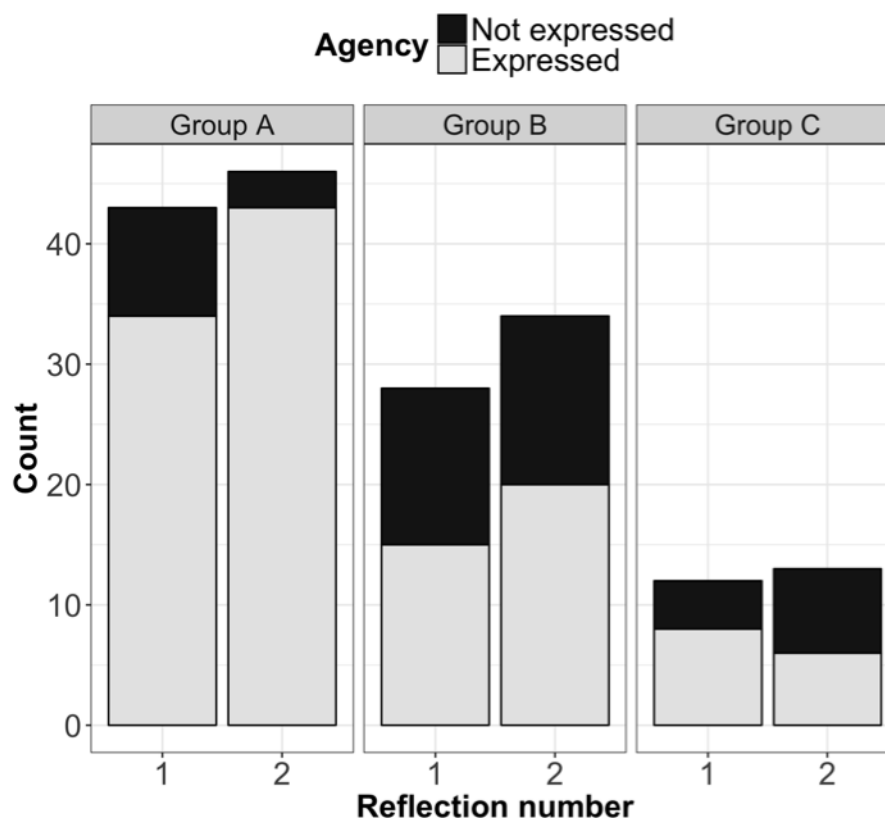
Figure 6. Expression of agency in group-reflection posts.

roles to increase higher-level thought in our case. First, we found students in the Synthesizer role emphasized the "summarize" function found in the role description and would provide short restatements of the main points made throughout the debate, rather than connecting points and identifying areas yet to be explored (Table 1). These summaries were explicitly mentioned as being useful in the group-reflections, due to the fast nature of the synchronous chat (Example 6). Second, of the questions asked across all groups in Chat 1, 83 were posed at lower levels of thinking ("know" or "understand" on Bloom's original taxonomy), while 45 were posed at higher levels of thinking. Notably, no questions were coded as "counter" questions during Chat 1.

Example 6: *". . . it is nice to have the synthesizers to summarize what is going on so that if someone does have a really long response that it can be summarized and easily understood."*

Combined, our observations regarding the Synthesizer and Questioner role implementation suggest that students will pick and choose the functions they find most advantageous or expedient within a role definition. In the case of the Synthesizer role, it seems that students found summaries to better serve their immediate needs due to the fast-paced nature of the synchronous chat, where keeping up with information is the primary challenge and the synthesis

of information becomes a secondary concern. For the Questioner role, although students were given three categories of questions to ask, they neglected to ask any counter questions, which were intended to fulfill the critique function described by Wise and colleagues (2012). Therefore, in both of these roles students effectively selected the functions of the role that favor lower-level thinking.

Self-Reflections

After the addition of self-reflections before the second chat, neither the proportion of critical student-content nor student-student interactions increased. In reading the reflections, we did find that students, as instructed by the self-reflection prompt, considered whether each role was properly executed in Chat 1 and how the roles could be better fulfilled in Chat 2. This is similar to the qualitative results from previous work that found self-reflections caused students to "identify patterns of participation which could enhance the larger learning community" (Kayler & Weller, 2007; Murphy & Jerome, 2005). However, neither of the studies report whether the ability to identify these patterns led to subsequent changes in chat quality. In contrast to these two qualitative studies, one quantitative study found that students assigned to complete self-reflections for online asynchronous discussions did not demonstrate an increase
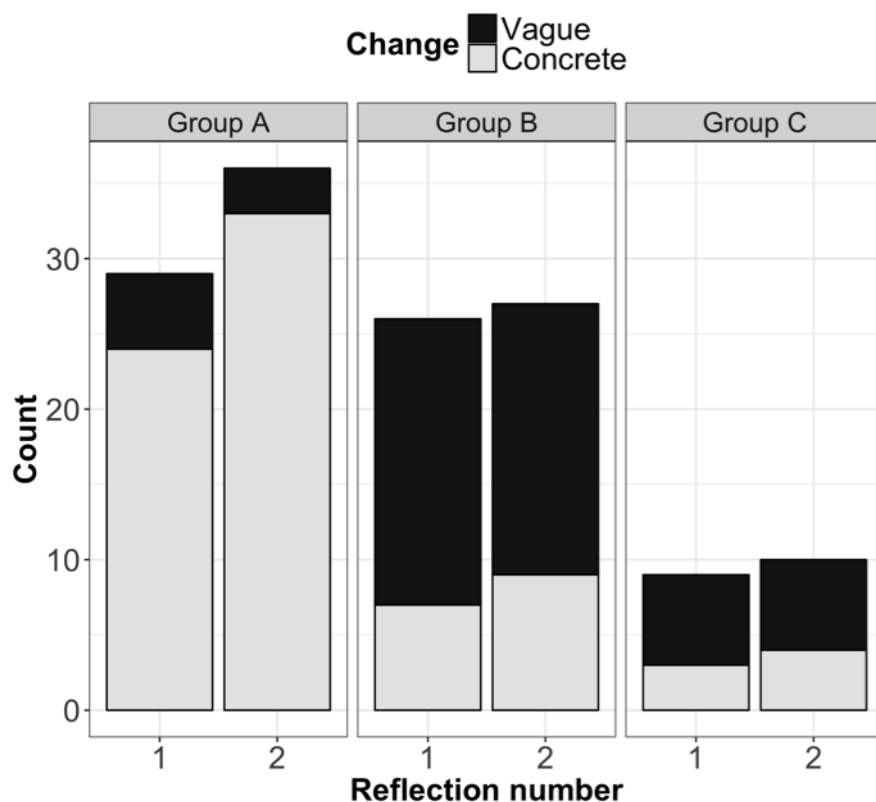
Figure 7. Types of changes suggested in the group reflections.

in knowledge construction compared to students who did not complete self-reflections (De Wever et al., 2008). Our results support these quantitative findings and suggest that self-reflections alone, even if thoughtfully completed by students as indicated by our qualitative results, are not sufficient to cause students to increase their critical thinking or interaction with other students in online chats.

An important similarity between our work and that of De Wever and colleagues (2008) is that undergraduates were enrolled in the course under consideration. Undergraduates may not be able to accurately self-asses the extent of their knowledge construction (De Wever et al., 2008), and in this case, the extent of their critical thinking. The qualitative studies supporting the use of self-reflections both considered graduate-level courses (Kayler & Weller, 2007; Murphy & Jerome, 2005). We therefore echo earlier calls for further research on the effects of providing training for undergraduates prior to completing self-assessments (De Wever et al., 2008). We also suggest that future work could look into whether changing the focus of the self-reflection to emphasize the desired pedagogical outcomes, or adding a step to make the self-reflection salient immediately before the subsequent chat, could result in a stronger quantifiable effect from the self-reflection intervention.

Group Reflections

For both Chat 3 and Chat 4, groups that completed group-reflections wrote a greater proportion of higher-level thinking posts than groups that did not complete group-reflections. This difference was significant for Chat 3 and significant for Chat 4 when all "high" outliers were removed. This suggests that providing groups an opportunity to discuss what happened in a chat and collectively decide on steps to improve the chat quality can help maintain higher-level thinking throughout the semester and can allow students to sustain thoughtful discussion outside the framework of assigned roles. Furthermore, these results offer evidence that the self-reflection intervention alone, which was completed by all groups for Chats 2 through 4, were not enough to sustain higher-level thinking in chats throughout the semester.

In our qualitative analysis, one group, "Group A," possessed unique qualities amongst the three groups analyzed. This group demonstrated more agency than the other two groups and more frequently suggested concrete changes to address inadequacies they saw in the chats. We suggest these two results are logically related; if Group A felt empowered to solve their own problems, it follows that they would also be more likely to suggest concrete actions to take in solving these problems. Previous work has shown that students who feel agency in online chats gen-

Table 3. "Ineffective" chat aspects identified in the group-reflections.

Group	Reflection Number	General category	Specific description	Number of distinct posts
A	1	Logistics	Multiple conversations occurring simultaneously	46
	2	Logistics	Schedule of the debate, specifically the order and timing of when main points should be presented and discussed	25
B	1	Logistics	Schedule of the debate, specifically, the "for" team received more time to present their arguments	19
	2	Outcomes/Logistics	Questions were left unanswered, but it is unclear whether concern about this was driven more by lack of information from peers (outcomes) or by multiple conversations happening at once (logistics)	19
C	1	Outcomes	The debate lost focus due to students getting too bogged down in "specifics" – for example, specific definitions	19
	2	Logistics	Multiple conversations occurring simultaneously	9

erally have a more positive chat experience (Kayler & Weller, 2007). This work suggests that feeling agency over the chat may also translate into students being able to suggest practical actions for improving ineffective aspects of the chat. Therefore, increasing the feeling of agency among group members seems to be a desirable outcome for chat groups. Kayler and Weller (2007) observed a correlation between the preference of face-to-face discussion as a learning style and expressing less agency in self-reflections. Thus, they added a group reflection component that required periodic in-person group meetings as a means of accommodating diverse learning styles, and hopefully promoting a feeling of personal agency in all their students. However, this does not address the question of how to promote personal and group agency in an all-online setting, with the further complication of discussion groups that are too large to be practical for video conferencing. With the increased prevalence of MOOCs, we see this as an important question for future research.

Even though Group A was successful in identifying aspects of the chat, under their control, that were ineffective, and subsequently proposed concrete actions to affect change, the group had mixed success in implementing these actions in the following chat. The group was able to implement the suggestion that they focus on one topic at a time and discussed this in the subsequent reflection as a successful achievement and improvement in chat quality. However,

there was a notable instance where some individuals from the group attempted to try the hand-raise feature included in the chat software, which was agreed upon by many during the group-reflection, but the idea was rejected when proposed during the chat. While students were instructed to come to a consensus over what they would change in the subsequent chat, the discussion forum was not set up in a way to facilitate reaching this consensus in a way that included everyone. In future group-reflection activities, it seems essential that the group have a way of determining buy-in from all group-members. Recapping the group-reflection immediately prior to the subsequent chat may also help to make the reasoning for the change more salient.

The final theme that emerged was that the deficiencies identified in the chats were a mix of logistical and learning-outcome related aspects, favoring logistics over outcomes. In all three groups, a logistical aspect of the chat was identified as the primary ineffective aspect in at least one of the reflections. Specifically, multiple conversations occurring at once and an unequal distribution of time to different portions of the chat were mentioned. There are a couple possible explanations for the focus on these logistical issues. First, it seems evident that the chat format, specifically, the "rush" created by the synchronous posting, was not satisfactory and was seen by students as inhibiting their ability to have an effective chat. Second, the group-reflection prompt did not ex-

plicitly instruct students to focus on the pedagogical outcomes of higher-order thinking and questioning, therefore, it seems reasonable that students would not choose to focus on these aspects. Going forward, changes need to be made both to the chat format and the group-reflection prompt to ensure students are able to focus on the desired learning outcomes over logistics.

Implications for Practice

Based on the results of this study, we plan to make the following changes in the next semester of this course: 1) Focus roles on one specific function rather than wrapping multiple functions into a role, to enhance fulfillment of target functions such as "synthesize" and "critique", 2) Remove self-reflections from the course assignments, 3) Change chat format from synchronous to asynchronous to reduce rushing and multiple conversations, and 4) Change the group-reflection activity to ask students to (i) focus on the desired learning outcomes, (ii) reach a consensus and get buy-in from whole group, and (ii) require students to revisit the group-reflection prior to starting the next chat, in order to make the motivation behind the agreed-upon change more salient.

While our study has the limitation of being a case-study of a single course, we feel some aspects of our results can be applied more generally. We add to the body of literature demonstrating that roles help engage students in chats, especially with their peers, however, our research highlights concerns that are specific to synchronous chats when designing role descriptions. Specifically, the fast-paced nature of the chat might cause students to need the logistical functions in the roles simply to have a manageable discussion. Our research also explores the relatively undocumented process of group-reflections in online settings and provides several guidelines for improving the group-reflection process. We feel group-reflection assignments could allow for increased student autonomy in running online discussion and agency in achieving the pedagogical goals for the course.

References

Allen, I.E., & Seaman, J. (2016). Online report card: Tracking online education in the United States. Babson Survey Group. Retrieved from https://onlinelearningconsortium.org/read/online-report-card-tracking-online-education-united-states-2015/

Bernard, R.M., Abrami, P.C., Borokhovski, E., Wade, C.A., Tamim, R.M., Surkes, M.A., & Bethel, E.C. (2009). A Meta-Analysis of Three Types of Interaction Treatments in Distance Education. Review of Educational Research, 79(3), 1243–1289. https://doi.org/10.3102/0034654309333844

Bloom, B.S., Engelhart, M.D., Furst, E.J., Hill, W.H., & Krathwohl, D.R. (1956). Taxonomy of educational objectives, handbook I: The cognitive domain. New York: David McKay Co Inc.

Bradley, M.E., Thom, L.R., Hayes, J., & Hay, C. (2008). Ask and you will receive: how question type influences quantity and quality of online discussions. British Journal of Educational Technology, 39(5), 888–900. https://doi.org/10.1111/j.1467-8535.2007.00804.x

De Wever, B., Schellens, T., Van Keer, H., & Valcke, M. (2008). Structuring asynchronous discussion groups by introducing roles: Do students act in line with assigned roles? Small Group Research, 39(6), 770–794. https://doi.org/10.1177/1046496408323227

De Wever, B., Van Keer, H., Schellens, T., & Valcke, M. (2007). Applying multilevel modelling to content analysis data: Methodological issues in the study of role assignment in asynchronous discussion groups. Learning and Instruction, 17(4), 436–447. https://doi.org/10.1016/j.learninstruc.2007.04.001

Ertmer, P.A., Sadaf, A., & Ertmer, D.J. (2011). Student-content interactions in online courses: The role of question prompts in facilitating higher-level engagement with course content. Journal of Computing in Higher Education, 23(2-3), 157–186. https://doi.org/10.1007/s12528-011-9047-6

Gibbs, G. (1994). Learning in teams. Oxford: Oxford Centre for Staff Development.

Hara, N., Bonk, C.J., & Angeli, C. (2000). Content analysis of online discussion in an applied educational psychology course. Instructional Science, 28(2), 115–152.

Holden, J.T., & Westfall, P.J.L. (2007). An instructional media selection guide for distance learning. Retrieved from http://eric.ed.gov/?id=ED501248

Kayler, M., & Weller, K. (2007). Pedagogy, self-assessment, and online discussion groups. Educational Technology & Society, 10(1), 136–147.

Kwako, J., Courter, S., & Wright, J. (2005). College classroom course: A guidebook. Center for the Integration of Research, Teaching, and Learning. Retrieved from https://www.cirtl.net/resources/4

Mercer, N., Wegerif, R., & Dawes, L. (1999). Children's talk and the development of reasoning in the classroom. British Educational Research Journal, 25(1), 95–111. https://doi.org/10.1080/0141192990250107

Moore, M.G. (1989). Editorial: Three types of interaction. American Journal of Distance Education, 3(2), 1–7. https://doi.org/10.1080/08923648909526659

Murphy, E., & Jerome, T. (2005). Assessing students' contributions to online asynchronous discussions in university-level courses. Electronic Journal of Instructional Science and Technology, 8.

Parker, K., Lenhart, A., & Moore, K. (2011). The digital revolution and higher education: College presidents, public differ on value of online learning. Pew Internet & American Life Project. Retrieved from http://eric.ed.gov/?id=ED524306

Pear, J.J., Crone-Todd, D.E., Wirth, K.M., & Simister, H.D. (2001). Assessment of thinking levels in students' answers. Academic Exchange Quarterly, 5(4), 94–99.

Persell, C.H. (2004). Using focused web-based discussions to enhance student engagement and deep understanding. Teaching Sociology, 32(1), 61–78.

R Core Team. (2015). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org.

Schellens, T. (2005). The impact of role assignment on knowledge construction in asynchronous discussion groups: A multilevel analysis. Small Group Research, 36(6), 704–745. https://doi.org/10.1177/1046496405281771

Schellens, T., Van Keer, H., De Wever, B., & Valcke, M. (2007). Scripting by assigning roles: Does it improve knowledge construction in asynchronous discussion groups? International Journal of Computer-Supported Collaborative Learning, 2(2-3), 225–246. https://doi.org/10.1007/s11412-007-9016-2

Strijbos, J.W., Martens, R.L., Jochems, W.M.G., & Broers, N.J. (2007). The effect of functional roles on perceived group efficiency during computer-supported collaborative learning: a matter of triangulation. Computers in Human Behavior, 23(1), 353–380. https://doi.org/10.1016/j.chb.2004.10.016

Sun, A., & Chen, X. (2016). Online education and its effective practice: A research review. Journal of Information Technology Education: Research, 15, 157–190.

Tagg, A.C. (1994). Leadership from within: Student moderation of computer conferences. American Journal of Distance Education, 8(3), 40–50. https://doi.org/10.1080/08923649409526865

Tukey, J.W. (1977). Exploratory data analysis.

Walker, S.A. (2004). Socratic strategies and devil's advocacy in synchronous CMC debate. Journal of Computer Assisted Learning, 20(3), 172–182.

Williams, K.M. (2015). Doing research to improve teaching and learning. Taylor & Francis.

Wise, A.F., & Chiu, M.M. (2011). Analyzing temporal patterns of knowledge construction in a role-based online discussion. International Journal of Computer-Supported Collaborative Learning, 6(3), 445–470. https://doi.org/10.1007/s11412-011-9120-1

Wise, A.F., Saghafian, M., & Padmanabhan, P. (2012). Towards more precise design guidance: specifying and testing the functions of assigned student roles in online discussions. Educational Technology Research and Development, 60(1), 55–82. https://doi.org/10.1007/s11423-011-9212-7

Zhu, E. (1998). Learning and mentoring: Electronic discussion in a distance-learning course. In Electronic Collaborators: Learner-centered Technologies for Literacy, Apprenticeship, and Discourse (pp. 233–257). New York, NY: Routledge.

Student Attitudes and Experiences with Self-Selected and Assigned Groups

Laura L. Tupper[1,*] and Kimberly Williams[2]

[1] Department of Mathematics & Statistics, Williams College
[2] Center for Teaching Innovation, Cornell University
[*] Corresponding Author: lllt1@williams.edu

Abstract

This study explores students' views on working in groups to complete course assignments via an anonymous survey administered to advanced statistics students at a large university. Responses to multiple-choice and open-ended questions indicate several trends. While students expressed strong distaste for working with weaker or less hardworking teammates, they also felt that the group structure offered them security when they themselves did not have the time or skill to complete an assignment. There was a general preference for self-selected groups over those assigned externally — but with greater favor toward externally-assigned groups when students are struggling, isolated, or non-native English speakers. Finally, students displayed a narrow view of the goals of group projects, focusing on the quality of the submission and grasp of the immediate course material. A greater emphasis on the broader skills developed through group work may be necessary to encourage students' investment in the group work process.

Introduction

Collaboration is part of working life, and many instructors use group work inside the classroom as well. The Guidelines for Assessment and Instruction in Statistics Education (GAISE) report, emphasizes the importance of the "ability to communicate" on statistical subjects and suggests team-based learning strategies for both large and small classroom environments (Carver 2016). Yet both instructors and students hold widely varying opinions about the value of group work, when it is most useful, and how it should be structured. In this study we sought to explore students' attitudes toward group work in the context of advanced statistics courses.

We surveyed students enrolled in statistics courses aimed at graduate and upper-level undergraduate students about their experiences with group work in the past and their opinions on group work and group formation under a variety of circumstances. A sizable proportion of the enrollment in these courses consists of international students (primarily, but not exclusively, from East Asia), so that we were able to gather responses both from native speakers of the classroom language (English) and from students for whom this was a second language.

Our investigation focused on four main questions:
- Why do students prefer solo or group work?
- Do students' opinions on group formation—choosing their own groupmates or having them assigned—change under different circumstances?
- What goals do students have for group projects?
- Do these attitudes differ depending on whether students are native speakers of the classroom language?

Previous Work

Many researchers have investigated the effectiveness of group work in the classroom, as well as students' response to it. For example, Hamlyn-Harris et al. (2006) surveyed students about their satisfaction with a team project experience, investigating the relationship with such factors as previous experience and team gender balance. Koles et al.(2010) found that students performed better when tested on material learned through group activities—and, notably, that the effect was strongest for the weakest students.

In a statistics context, there is a considerable history of work on group-oriented learning. Garfield (1993) discussed a variety of ways to incorporate small-group activities in the statistics classroom, while Björnsdóttir et al. (2015) reviewed two different practices for cooperative assessment. Keeler and Steinhorst (1995) offered evidence that students' performance in an introductory statistics course—as measured by final grades and student retention rates—improved when group activities and homework groups were introduced, as compared to a standard lecture format. Giraud (1997) went further, comparing two sections of a statistics course taught in the same term by the same instructor; test scores in the section using in-class group assignments exceeded those in the lecture-based section.

If the instructor is to introduce group-based activities or assignments into a course the natural question becomes how these groups are to be formed. For the most part, group formation methods can be seen as lying along a spectrum reflecting how much control students have in the process. At one extreme is the self-selected group, where students are permitted to choose groupmates among all other students in the course. At the other is the externally-assigned group, where an instructor specifies the members of each group. The assignment may deliberately take into account students' skills, backgrounds, and demographics, as well as more mundane information: Tonn and Milledge (2002) described an attempt to assign groups for maximum diversity tempered by the need to match students with compatible schedules. Alternatively, the instructor may take no subjective role in the process. For example, Magel (1998) described the implementation of a random group assignment method, in which students are assigned to new random groups for each activity. St. Clair and Chihara (2012), who recommended consistent groups throughout the course, used random assignment but adjusted the results to ensure demographic balance across groups. Hansen and Hansen (2007) suggested an approach known as "speed teaming," where students report classmates with whom they would or would not be willing to collaborate, but the instructor makes the final assignments.

Both approaches have their proponents, and there is no clear consensus on which is preferable. Hilton and Phillips (2010) reported findings from students' journals on the team experience kept during a large-scale group project, contrasting the reports from self-selected and externally-assigned groups. They found a higher level of initial discomfort in assigned groups, but they also noted that students often came to value the diversity of such groups as the semester progressed. Oakley et al. (2004), however, noted the concern of "isolating at-risk minority students on teams." Bacon et al. (1999) tooka reversed approach, by surveying students about their best and worst experiences with group work and analyzing these responses for associations with different group structures and circumstances; they reported an association between self-selection and "best team experiences." This may be an indication of confounding between emotional and intellectual experience, a theme we observe in the contrast between "least favorite" and "most common" group work problems in our own respondents.

Other research has investigated the classroom experience of language learners and non-native speakers of the language of instruction. Of particular inter-est is the work of Lesser et al. (2013), which analyzed language learners' experience with handling the peculiar vocabulary of statistics in an introductory course: language learners were found to have greater trouble navigating the register shift between statistical and everyday language, and also struggled with examples drawn from non-shared cultural practices. Sisto (2009) described a statistics course in an environment where most students are non-native English speakers, and come from a wide range of nationalities. Sisto uses a modified self-selection approach for undergraduate courses (students are not permitted to work with the same classmate more than once) and instructor assignment for MBA courses, specifically stating a goal of "no more than two common language speakers per group" to ensure "working across cultures and languages."

Methodology

Respondents

The students surveyed were or had been enrolled in at least one of three statistics courses: "Linear Models with Matrices" (fall 2015), "Multivariate Analysis" (spring 2016), or "Applied Design" (spring 2016). Students who were enrolled in more than one of these courses only received one link to the survey. The courses were selected since author LLT had been a teaching assistant for each one, and was thus familiar with the course content and structure.

Each course was offered as a Statistical Science course and cross-listed in another department ("Linear Models with Matrices" as Biometry & Statistics, and the others as Social Statistics). They were all considered upper-level statistics courses where the typical enrollment included undergraduate majors in statistics and related fields, students in the Master's of Professional Studies in Applied Statistics program, and doctoral candidates in other departments. Course content included a mix of theory and application as well as some programming in R.

In addition, each course was known to have included some element of group work, so that all students surveyed had at least some experience with the practice. In "Linear Models" students could attend optional weekly programming labs in which they applied the theoretical methods from the lecture to real datasets using R. These labs were conducted with an active learning approach in which groups of two to four students were asked to review and run example code and answer associated questions. Group formation was mixed; originally groups were assigned randomly but because of low attendance "orphan" students often worked with whoever was seated nearby. Apart from some fluctuations in attendance, students

worked with the same teammates throughout the semester. In "Applied Design" students were given the option of working on homework assignments in groups of up to three members. These groups were self-selected; on most assignments about two-thirds of students chose to work with at least one teammate. Group membership stayed fairly consistent over the course of the semester. In "Multivariate Analysis" students were required to complete a group project analyzing and reporting on a dataset of their choice. The project was large in scale, spanning several weeks and incorporating both an in-class presentation and a formal written report. Students were assigned to groups of two to three individuals based on instructor assessment of their command of the material and presentation skills, with the instructor aiming for groups of diverse skill levels and language backgrounds. These groups had not yet been formed at the time of survey administration but students were aware of the upcoming project and had been asked to do some preparatory work, such as looking for a dataset.

Survey Methodology

The survey was administered online with access provided via a link emailed to students at their college account. It was made clear that the survey was both voluntary, anonymous, and students were free to skip parts of the survey. It should be noted that, inevitably, this format introduced considerable response bias: the small proportion of students who responded may well have been those with particularly strong feelings about group work, or those with an inclination to do extra work.

The instructions stated that aggregated responses might be considered in the formation of project groups for "Multivariate Analysis" which had not yet been completed. Accordingly, nearly half of respondents reported that they were enrolled in "Multivariate Analysis" even though it had the smallest total enrollment of the three courses.

There were thirty-eight questions in the survey including a mix of multiple choice questions, short free response questions, and longer text response questions. Multiple choice questions included objective questions ("How often have you done group work for school where students chose their own groupmates?") and Likert-type opinion questions ("On average, how much did you enjoy working with groupmates you chose?"). Three questions were designed as matrices, with Likert-style responses to a series of questions with a common stem ("If you are working on an assignment in an upper-level statistics course, how important to you is..."). Short free response questions focused on demographic information ("When talking with friends, the language I prefer to use is:"). The longer text response questions were generally paired with a multiple choice question, and offered students a chance to explain or expand upon their answers ("What is one reason (or more!) why you feel this way?").

Survey questions focused on four main topics. First, we collected demographic information on students' status, personal classroom style, and language preferences. Secondly, we asked students' views on doing class assignments in groups in general, as opposed to working alone; we then investigated their opinions on group formation under various circumstances, contrasting self-selection of teammates against external assignment of teammates. Finally, we asked students' opinions on the goals of group projects and how effectively, in their opinion, these goals were met.

Results

Demographic Information

We received 31 completed responses from a total of 173 invitations. A response was considered complete if the student reached the end of the survey, even if they did not answer every individual question. 28 students answered the question "When talking about statistical subjects, I feel my level of English proficiency is...," and we based our analysis on these responses (Figure 1). We grouped responses into those of native English speakers (those who answered "native/first-language speaker", n=16) and those of non-native speakers (those who answered "some difficulty" or "fluent", n=12). No student self-identified as being "not proficient."

Students were asked to identify their personal style in the classroom, ranging from "very withdrawn" to "very outgoing." Most students considered themselves in the middle of this range. There were no notable differences between native speakers and non-native speakers in terms of the distribution of this response, though it is important to remember that these students included only those who felt sufficiently engaged in the course to respond to a voluntary survey.

One point on which native and non-native speakers did differ was student status. Most native speakers were upper-level undergraduates (in their third year or later), with a handful of graduate students and one lower-level undergraduate. The non-native speakers, by contrast, were predominantly enrolled in Masters' or professional programs, with some upper-level undergraduates and doctoral students.
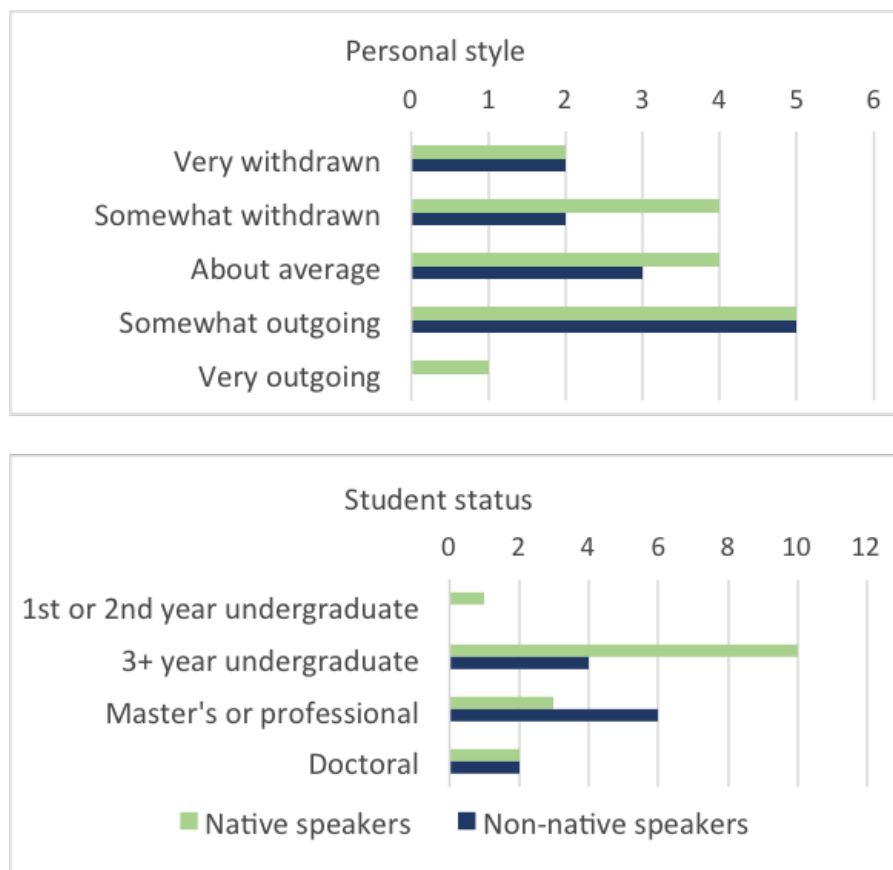
Figure 1: Student responses to demographic information questions pertaining to personal learning style and student status.

Group Work Versus Working Alone

We first addressed the straightforward question of whether students preferred to work alone or with others (Figure 2). In particular, we asked students to consider two types of assignments: small-scale assignments that took a short time and had a small impact on the course grade, and large-scale assignments that took place over a period of weeks and made up a large part of the course grade.

In the small-scale case, most students preferred working alone regardless of their language preferences. In text responses students cited the efficiency and flexibility of working alone; one student noted that "the effort it takes to coordinate group work is not worth it for a small assignment." Others considered working alone to be more effective in developing and assessing the student's own command of the material: "If I can do it myself," one student wrote, "then I know that I understand." Native English speakers, however, were more likely to say that their preferences depended on the difficulty of the assignment.

Among the perceived benefits of group work was the opportunity to "exchange ideas and have different perspectives." Especially for large-scale assign- ments students valued the ability to "bounce ideas off" teammates and draw on group members' "different strengths." For large-scale assignments student preferences shifted to group work. The main reason given was the opportunity to share work among group members, reducing the amount of work individuals would have to do and speeding up the overall process, particularly when components of the project could be "delegated" to group members. Other students saw group work as a form of quality assurance, a mechanism that took on greater importance for major assignments; one wrote, "I am risk averse. . . . I can be more confident that we get the things correct."

Students who maintained a preference for working alone even on large assignments expressed concern about "depending on someone else for part of [their] grade" and "the impact of a group member not pulling their weight." They were aware that a large project could be challenging to complete alone but did not feel that having teammates would necessarily make it easier; one wrote that "discussing the project with TAs and instructors can be much more informative than peer discussing."

Inspired by the "best and worst" approach of Bacon et al. (1999) we asked students to report their favorite and least-favorite part of working in groups,
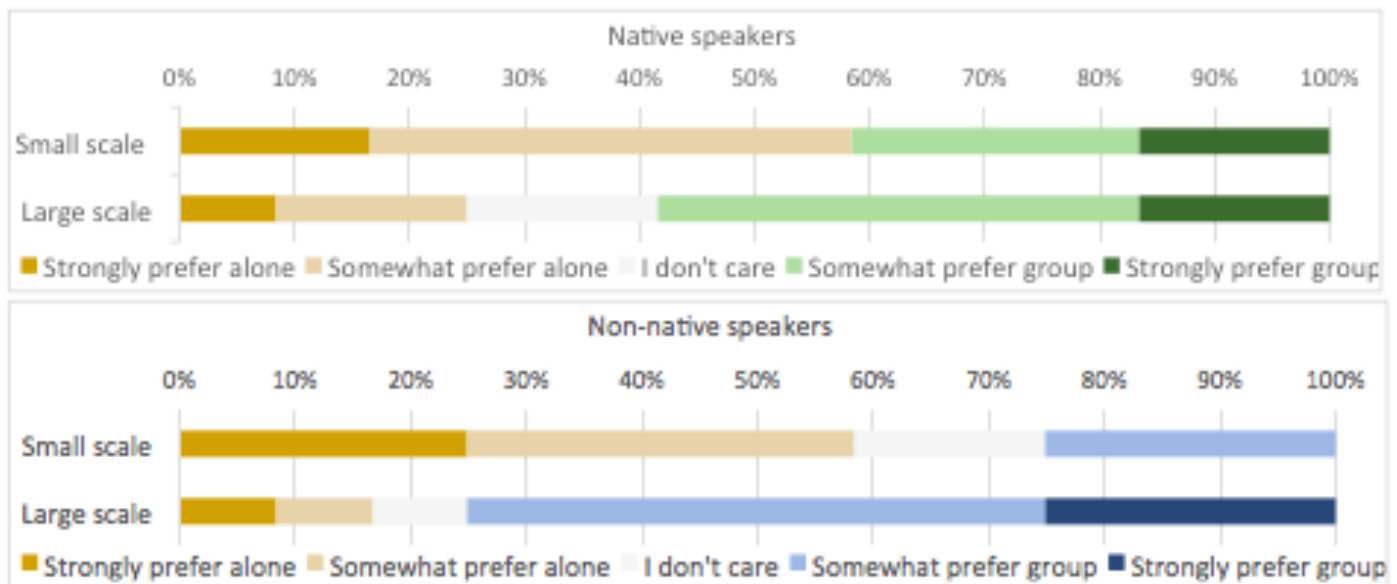
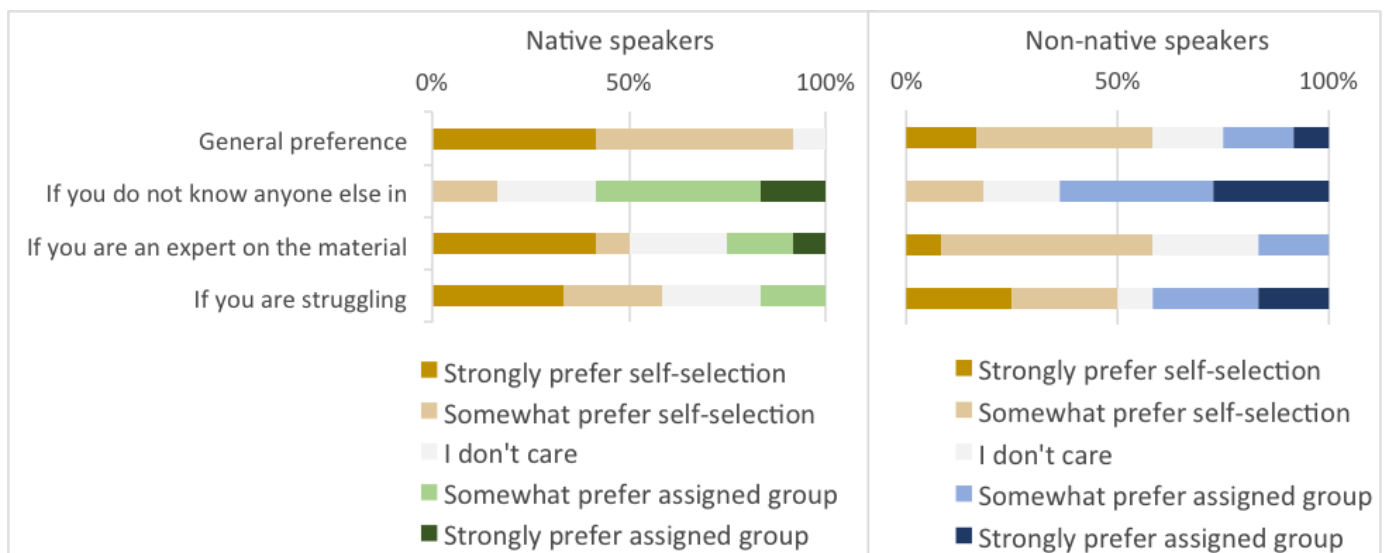Figure 2: Student preferences for group projects versus working alone.

Figure 3: Student preferences for self-selected versus externally-assigned partners when working on group projects.

as well as the most common problem they had encountered with group work in the past.

An interesting contradiction emerged here. When asked their least-favorite part of group work, students' top complaint, by far, was the "free-rider": a teammate who did not do their share of the work, whether through incompetence or unwillingness to put in effort. Students' responses were sometimes intensely emotional, with one student describing the situation as "like being held hostage." Other students used language of trust and fear; some preferred working alone to "having to rely on others." Yet, when asked the most common problem they had experienced, students' top response was more mundane: difficulty in finding times to meet. Such an imper-

sonal issue as scheduling may not elicit the strong emotional response of a "free-rider," and may not loom as large in students' expectations and fears of group work; but it nevertheless appears to be a major part of the experience. St. Clair and Chihara (2012) solved this issue by restricting group activities to class time, but this approach may not be feasible for large-scale projects.

A second contradiction appears as well: even as students decried teammates who did not put in an equal amount of work, they praised group work for its potential to decrease their own workload. The most commonly mentioned benefits of group work were efficiency and reduced workload, whether through "division of responsibility" or simply "helping hands."
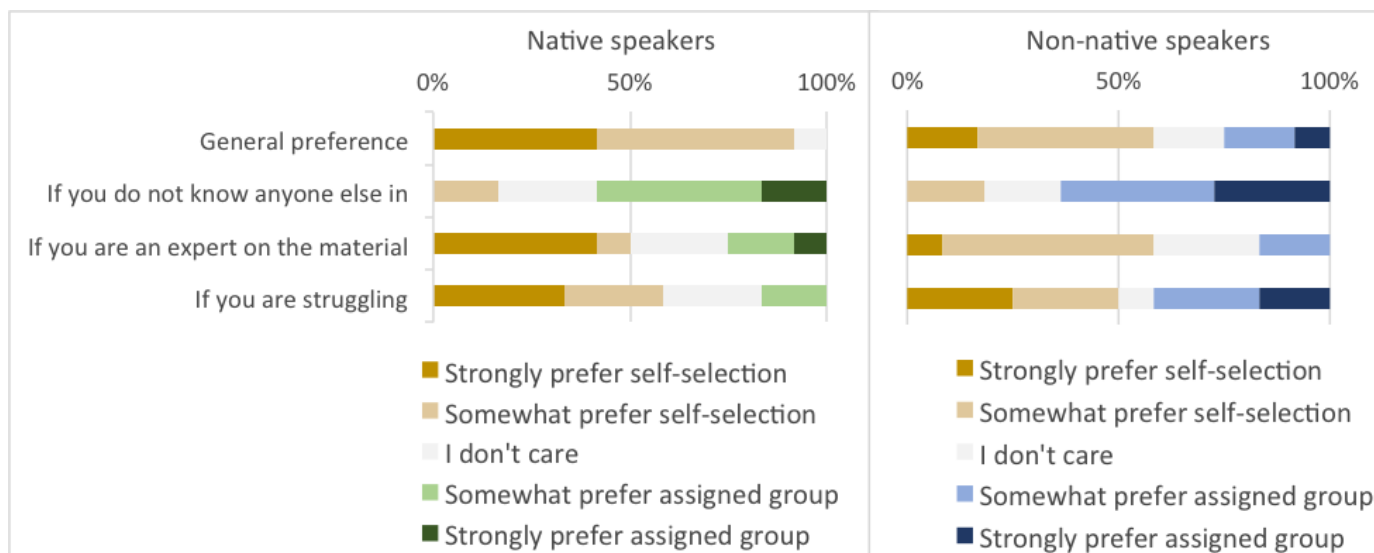
206

Figure 4: Student perception of the goals and benefits of participating in group learning experiences.

Some students explicitly stated an expectation that group members would, if necessary, compensate for a student's lack of input: "If someone gets busy," one wrote, "group members still get his back."

Self-Selected Versus Externally-Assigned Groups

A series of combined multiple-choice and free response questions investigated students' preferences for group formation, that is, whether students chose their own teammates or were assigned teammates externally (Figure 3). In the general case, most students stated a preference for self-selected groups. Notably, the only exceptions to this trend were non-native English speakers.

The top concern in group selection, according to text responses, was the students' "need to get along with group members." Students emphasized that their default preferences depended on whether they knew people in the class. Teammates' personality and work style were paramount.

Students' opinions changed markedly when they were asked to imagine a situation in which they did not know any of their classmates. A few responded pragmatically that if they had no information about their classmates there would be, effectively, no difference between choosing their own teammates and having them assigned. More frequently, however, they expressed a dislike for the idea of having "to reach out to strangers" during the group formation process.

When asked to imagine struggling with the course material students' top concern was finding groupmates who were "smarter" or "could clarify" the material. Many felt that external assignment would aid in this goal: "If I am not good at the course material," one wrote, "there is a good chance that the people I know are not good at it as well." This response echoes

the findings of Magel (1998), who observed that, in activities when students were allowed to choose their own collaborators, students with "poor quantitative reasoning skills" often worked with classmates at the same level. Magel favored random group assignment, specifically to counter this tendency, when deliberate instructor assignment is not feasible because of class size.

Most students had less strong preferences about group assignment in a situation where they were individually comfortable with the course material. One student wrote, "No matter who are in my group, we'll do well." Some expressed a desire for self-selection so they could find "complementary" teammates. Interpersonal considerations were once again a priority when students did not need to rely on their group members' skill with the material.

When they praised self-selection of groups students typically assumed that they would work either with their friends or with "good" students. None expressed concern that they would not be able to work with their preferred teammates—either because of restrictions on group size or because their desired partners would not want to work with them. Even when specifically prompted to imagine that they were struggling in the course, students never raised the issue of their own potential undesirability. Their preferences did shift toward externally-assigned groups in this situation, but their text answers remained focused on which teammates they desired, without consideration of whether others would be willing to work with them.

Goals of Group Work

When prompted to consider broader goals, like development of teamwork and communication skills, students obligingly rated these concepts as important

parts of group assignments (Figure 4). But the text answers are telling: only two students ever mentioned this type of goal of their own accord in free-response questions.

No students rated "reducing the amount of work you have to do" as "very important" in the Likert scale questions. It is, however, reasonable to assume some response bias here, particularly given the frequency with which workload issues appear in students' text answers. Some students even specifically listed the potential to do less work on an assignment, especially when they were busy with other commitments, as their favorite part of working in groups. This contradiction may, in part, be a matter of phrasing. While students are reluctant to admit that they would like to do less work on an assignment due to working in a group, they are very open about not wanting to do more work on an assignment.

Conclusions

Because of the low response rate, probable response bias, and limited scope of the survey, generalization of our results to a broader environment is questionable. Nevertheless, we identified three key points for consideration based on our results.

Escher's Staircase

Student responses presented a clear paradox in that all students wanted to work with classmates who were at least as strong as they were. In text answers, students praised group work for its potential to reduce their workload and compensate for areas where they struggled. Yet they expressed distaste for working with those who were less hardworking or capable. Logically, of course, this ideal is impossible: unless every student is exactly the same, some students will have to work with weaker or less diligent teammates.

Regardless of whether working with weaker classmates is in fact harmful to student performance the issue is an emotional one for students and demands explicit attention. Most immediately, instructors can set policies to protect stronger students from "free-riders"; for example, St. Clair and Chihara (2012) recommended only including group performance in a student's grade if that student's individual grades are above a certain threshold.

Some approaches to external group assignment make the best of skill differences between students by balancing the overall skill levels of the groups. In such cases, an appeal to fairness may be a strong response to student complaints about external group assignment. Instructors may also emphasize the multi-dimensional nature of each student: building Escher's staircase is more feasible if each teammate can be considered "stronger" in some ways but not in others. Students' satisfaction and investment may improve when they are encouraged to identify and value the particular strengths of their groupmates.

Vulnerable Students

Though most students expressed a preference for selecting their own groups the exceptions to this tendency are notable. Student preferences shifted toward externally-assigned groups in situations—actual or hypothetical—where we might consider students to be most vulnerable: when they were struggling with the material, unfamiliar with their classmates, or non-native English speakers.

The isolating effect of these situations may be exacerbated by self-selection of groups, as students are reluctant to approach their better-connected classmates. Left to their own devices such students may end up forming groups with others in a similar situation. In the case of struggling students the groups may thus become unbalanced in skill. In the case of poorly-connected students and non-native English speakers, even if the resulting groups are equal in ability, this academic and social isolation may work against an inclusive classroom environment.

Valuing the Process

Students had a course-focused view of group work, thinking in terms of its impact on their workload, grasp of material, and performance on the assignment at hand. Ancillary skills such as communication and relationship-building, which are of primary use in the workplace rather than the classroom, were almost never mentioned. Given such a narrow view of the goals and outcomes of group work students' preferences—for working with friends or stronger students, and for avoiding the overhead of group work—are entirely reasonable. Working together with weaker students will not, intuitively, increase the quality of the output.

There is, on a straightforward level, an opportunity to build on students' existing goal of better mastery of the course material. Research shows that peer teaching is an effective means to this end, this being one reason why educators favor group work in the first place. For example, Kalaian and Kasim (2014) conducted a meta-analysis of studies on small group workand found that collaborative learning techniques have a significant overall positive effect on student performance. Sharing this research with students and convincing them that the group process can in fact give them a deeper and more lasting grasp of the material may motivate them to participate more fully in the process.

A more fundamental shift in approach is to place more explicit importance on big-picture goals, holding up the process of group work with all its interpersonal challenges as a valuable experience in itself. When the process is emphasized a scholastically weak group member is no longer merely a dead weight lowering the group's productivity; instead, the situation becomes an opportunity to practice and demonstrate group management and peer teaching skills. When students feel that even their teammates' (and their own) weaknesses are contributing to the experience they may be more open to working with a broader range of their classmates, and they may try harder to keep their teammates involved in the group's work.

References

Bacon, D.R., Stewart, K.A., and Silver, W.S. (1999). Lessons from the best and worst student team experiences: How a teacher can make the difference. Journal of Management Education 23(5): 467-488.

Bjornsdottir, A., Garfield, J., and Everson, M. (2015). Evaluating two models of collaborative tests in an online introductory statistics course. Statistics Education Research Journal 14(1):36-59.

Carver, C., Everson, M., Gabrosek, J., Rowell, G.H., Horton, N., Lock, R., Mocko, M., Rossman, A., Velleman, P., Witmer, J., and Wood, B. (2016). Guidelines for assessment and instruction in statistics education. Gaise College Report (draft).

St. Clair, K. and Chihara, L. (2012). Team-based learning in a statistical literacy class. Journal of Statistics Education 20(1).

Garfield, J. (1993). Teaching statistics using small-group cooperative learning. Journal of Statistics Education 1(1).

Giraud, G. (1997). Cooperative learning and statistics instruction. Journal of Statistics Education 5(3).

Hamlyn-Harris, J.H., Hurst, B.J., von Baggo, K., and Bayley, A.J. (2006). Predictors of team work satisfaction. Journal of Information Technology Education 5.

Hansen, R.S. and Hansen, K. (2007). The student experience in speed teaming: A new approach to team formation. Journal of College Teaching and Learning 4(7).

Hilton, S. and Phillips, F. (2010). Instructor-assigned and student-selected groups: A view from inside. Issues in Accounting Education 25(1): 15-33.

Kalaian, S.A. and Kasim, R.M. (2014). A meta-analytic review of studies of the effectiveness of small-group learning methods on statistics achievement. Journal of Statistics Education 22(1).

Keeler, C.M. and Steinhorst, R.K. (1995). Using small groups to promote active learning in the introductory statistics course: A report from the field. Journal of Statistics Education 3(2).

Parmelee. (2010). The impact of team-based learning on medical students' academic performance. Academic Medicine 85(11).

Lesser, L.M., Wagler, A.E., Esquinca, A., and Valenzuela, M.G. (2013). Survey of native English speakers and Spanish-speaking English language learners in tertiary introductory statistics. Statistics Education Research Journal 12(2): 6-31.

Using cooperative learning in a large introductory statistics class. Journal of Statistics Education 6(3).

Oakley, B., Felder, R.M., Brent, R., and Elhajj, I. (2004). Turing student groups into effective teams. Journal of Student Centered Learning 2(1): 9-34.

Sisto, M. (2009). Can you explain that in plain English? Making statistics group projects work in a multicultural setting. Journal of Statistics Education 17(2).

Tonn, J.C. and Milledge, V. (2002). Team building in an MBA gateway course: Lessons learned. Journal of Management Education 26(4): 415-428.
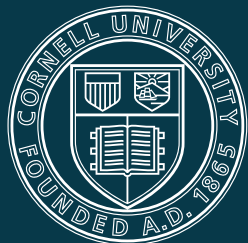
Center for Teaching Innovation
420 Computing and Communications Center
cornellcti@cornell.edu
teaching.cornell.edu
607-255-7224